Role of mammographic density and other risk factors in the development and detection of breast cancer

Dixa B Thakrar

University College

University of Oxford

Thesis submitted for the degree of

Doctor of Philosophy

Michaelmas term 2024

Cancer Epidemiology Unit

Nuffield Department of Population Health

University of Oxford

For

Mummy, Daddy,

Vedas & Sathya

Abstract

Mammographic density (MD) is the extent of radiologically dense breast tissue. It is an important risk factor for breast cancer and it reduces mammographic screening sensitivity. Menopausal hormone therapy (MHT) is known to reduce screening sensitivity due to its association with increased MD. The aim of this thesis was to elucidate the role of MD, alongside various other risk factors, in the development and detection of breast cancer. This was done through 1) a systematic review and meta-analysis of the association of breast cancer risk factors with MD; 2) a cross-sectional analysis of the association of breast cancer risk factors with MD in the Million Women Study (MWS), a UK-based screening cohort; 3) a prospective analysis of the effects of MHT, an important determinant of MD, on screening sensitivity in the MWS; and finally 4) developing overall breast cancer and subtype-specific (based on mode of detection and grade) risk prediction models using easily collectable risk factor information from questionnaires, and determining if given risk factors were better at predicting specific breast cancer subtypes, are were more likely to be missed at screening.

The systematic review, with 59 studies, suggested that higher alcohol intake and later age at first birth and menarche were associated with higher MD, while parity and smoking were inversely associated with MD. These results were largely in agreement with the analysis of the MWS, with MD data on 8190 women, in which current/recent MHT use, older age at first birth, alcohol consumption, physical activity, and benign breast disease were positively associated with MD, whereas age, parity, smoking and BMI were inversely associated with MD. The prospective analysis of MHT use, with 27,564 breast cancer cases, found that screening sensitivity was reduced in current/recent MHT users and that although this impact on screening sensitivity declined after stopping MHT use, it remained elevated 10+ years post-cessation, likely due to some persistence in the effects of MHT use on MD. The overall and subtype-specific risk prediction models using easily collectable questionnaire data (not including MD) performed poorly. MHT was only marginally predictive of cancers diagnosed out of screening, and none of the other risk factors were notably better at predicting specific breast cancer subtypes.

This thesis has confirmed and expanded upon existing knowledge regarding the role of MD and other breast cancer risk factors on breast cancer risk and detection at screening. Identifying those at greater risk of breast cancer or having their cancers missed at screening will ultimately guide future public education decisions, MHT prescribing practices, lifestyle advice, and – after further economic evaluation – possible personalised screening strategies.

Acknowledgements

I would like to express my sincerest gratitude to my supervisors, Professor Gillian Reeves and Associate Professor Isobel Barnes, for trusting me with this project and for their unwavering support and advice throughout my DPhil. Their doors were always open to me whenever I needed them. I also extend my thanks to Associate Professor Toral Gathani for her prior supervision. I am grateful to the women and staff of the Million Women Study; it was a privilege and an honour to use their data for my thesis. Additionally, I would like to thank Cancer Research UK for generously funding my studies.

There are several people within the Cancer Epidemiology Unit, without whom this work would not have been possible – Kirstin Pirie, who provided endless support with Stata or R; Sau Wan Kan, who kindly put together the datasets and helped me navigate the vast volume of data; fellow DPhil students Trishna Desai and Zoe Grenville for serving as my secondary reviewers in what was a mammoth of a systematic review; and fellow office comrades Maira Khan, Dr Yang Tao, and Dr Tonima Trisa for their sound boarding abilities.

I would additionally like to thank my Transfer of Status and Confirmation of Status assessors, Professor Sarah Darby, Professor Sarah Lewington, Professor Tim Key and Associate Professor Christiana Kartsonaki, who took the time to understand my thesis and provide feedback for the subsequent stages.

My time at Oxford would not have been the same without my college, Univ. The lodge staff and front-of-house team were always attentive to my countenance, often recognising when I seemed more stressed and never failing to check in on me. I also extend my thanks to my college advisor, Professor Nicola Sibson, for her rather entertaining words of wisdom, especially during college formals.

I would also like to thank my family, to whom this work is dedicated – my parents for their encouragement, commitment and faith throughout my education and the vicissitudes of life; Vedas, who is not only my brother but also a friend and confidant, and who even managed to tolerate my whims after his move to Oxford; and Sathya, a man of very few words, whose insights always provided some motivation.

Finally, I offer my loving gratitude to my guru, Sri Sathya Sai Baba, my source of love, light, and hope, who has guided me throughout my life and will continue to do so.

Role of the author in the preparation of this thesis

The research presented in this thesis is my own, conducted under the supervision of Professor Gillian Reeves and Associate Professor Isobel Barnes, who have overseen and supported my work, assisting with analyses, checking results, and reading and commenting on drafts of this thesis.

The Million Women Study (MWS) was conceived and started by the late Professor Dame Valerie Beral. Data collection, cleaning, updating and linkage has been conducted by several individuals, without whom this thesis would not have been possible.

The initial idea of investigating the role of mammographic density in relation to breast cancer risk factors and the associations of menopausal hormone therapy with screening sensitivity in the MWS was conceived by my supervisors and further developed in my discussions with them. With guidance from my supervisors, I conducted all statistical analyses, produced all relevant tables and figures, and drafted the text. To produce some of the figures, I used the R package Jasper, developed by Matthew Arnold, with the additional guidance of Dr Barnes and Kirstin Pirie (senior statistical programmer). Further sources of statistical code that were neither my own nor from my supervisors have been referenced, where used, within the thesis.

I have additionally benefited from contributions from colleagues at the Cancer Epidemiology Unit, including Kirstin Pirie, whose Stata code I adapted to generate descriptive tables and conduct parts of the risk prediction in Chapter 6, Sau Wan Kan, who provided the MWS datasets, fellow DPhil students Trishna Desai and Zoe Grenville who acted as secondary reviewers in the systematic review, and Keith Shaw who provided linkage to the mammographic density data. The mammographic density data was provided by the Medical Physics Department at the Royal Surrey County Hospital, Guildford, UK.

Table of Contents

List of tables

List of figures

List of abbreviations

95% CI	95% confidence interval
AI	Artificial Intelligence
AUC	Area under the curve
BBD	Benign breast disease
BIRADS	The Breast Imaging Reporting and Data System
BMI	Body Mass Index
DBT	Digital breast tomosynthesis
CPRD	Clinical Practice Research Datalink
ER	Estrogen/oestrogen receptor
GP	General practitioner
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
ICD	International Classification of Diseases
MD	Mammographic density
MHT	Menopausal hormone therapy
MRI	Magnetic resonance imaging
MWS	Million Women Study
NHS	National Health Service
NHSBSP	NHS Breast Screening Programme
O + P	Combined MHT (oestrogen + progestogen)
OC	Oral contraceptive
O-only	Oestrogen-only MHT
OR	Odds ratio
PR	Progesterone receptor
PRS	Polygenic risk scores
ROC	Receiver operating characteristic
RR	Relative risk
SD	Standard deviation
SE	Standard error
WHI	Women's Health Initiative Trial

Chapter 1. Background

1.1 Introduction

There are many well-established risk factors for breast cancer. These include female sex, higher age, family history, inherited gene mutations, increased mammographic density (MD), exogenous hormone use (menopausal hormone therapy or hormonal contraception), lifestyle factors (alcohol, physical activity, smoking), and anthropometric (obesity and height) and reproductive factors (nulliparity, earlier age at menarche and menopausal status) amongst others. It is generally accepted that many of these risk factors (for example, reproductive factors and use of exogenous hormones) alter breast cancer risk through their influence on sex hormones, but the mechanisms underlying associations with other risk factors (for example, height, smoking, and physical activity) are still unclear.

After age and some pathogenic mutations, mammographic density has one of the largest effect sizes associated with breast cancer risk and is likely to have a unique role in the development and detection of breast cancer.[1] Given that hormones influence MD, it is plausible that many hormonally related risk factors for breast cancer may act, at least in part, through their effects on MD. However, relatively few large studies have had data on both MD and risk factors and so there is comparatively little evidence regarding the degree to which MD may mediate known risk factors for breast cancer. MD is also known to reduce the sensitivity of breast cancer screening and so can affect the risk of having breast cancer diagnosed, independently of its effect on incidence.[2]

The aim of this thesis is to study the role of MD and other breast cancer risk factors on pathways to breast cancer diagnosis both by investigating the relationship between MD and other established risk factors for breast cancer, and by exploring the potential influence of menopausal hormone therapy (MHT), on mammographic screening sensitivity. The thesis also aims to explore breast cancer risk prediction, particularly of breast cancer subtypes. The findings could help identify ways to improve prevention and early detection of breast cancer.

This chapter provides a background to the topics relevant to this thesis. It starts with both a general introduction to breast cancer epidemiology and an overview of the most established risk factors and their relative importance for breast cancer incidence. Next, it provides a more detailed description of MD, its measurement, and the evidence for its role as a mediator in known risk factor-breast cancer associations. It then outlines the main pathways to breast cancer diagnosis and the potential role of MD and other risk factors in detecting breast cancer through routine screening. Lastly, it provides a brief introduction to breast cancer risk prediction before going on to the aims of this thesis.

1.2 Breast cancer epidemiology

1.2.1 Incidence

Breast cancer is the most common cancer in women worldwide, accounting for 1 in 8 cancer diagnoses. In 2020, there were 2.4 million cases among women worldwide. There is considerable geographical variation in incidence. Breast cancer incidence is highest in more developed countries. Incidence in parts of Asia and Africa is less than half of that in parts of Europe and North America. Age-standardised rates per 100,000 females in 2020 were 36.1 across developing countries and 75.7 in developed nations.[3] In the UK, breast cancer is the most common cancer in women, with approximately 55,900 new cases each year, and rates are increasing.[4] This accounts for 15% of all UK cancer diagnoses. There are around 370 cases in men in the UK each year.

It is estimated that worldwide, there will be more than 3 million new cases each year by 2040, representing a 40% increase from current annual rates.[3] This is likely due to a combination of changing lifestyle habits, reproductive patterns, ageing populations, population growth, increased public awareness, and improvements in detection, screening, and reporting.

1.2.2 Mortality

Globally, an estimated 685,000 women died from breast cancer in 2020. This corresponds to one in every six cancer deaths in women. Age-standardised mortality rates per 100,000 females were 20.1 across developing countries and 13.4 among developed countries in 2020.[3] This disparity is likely due to variations in public awareness, access to treatment, and population-based screening services.

Mortality across developed countries has declined by 40% over the past 40 years, reflecting improvements in detection, public awareness through self-examinations, and advanced and personalised treatment options. However, global mortality is estimated to increase to 1 million annually by 2040. Factors associated with an increased risk of mortality from breast cancer include later stage of diagnosis, diagnosis at the extremes of age (very young or very old), triple negative-hormone receptor status, and socioeconomic status. Among women in the UK, breast cancer has the second highest mortality rate after lung cancer, with 11,400 deaths each year.[5]

The overwhelming incidence and associated mortality of breast cancer make it a significant and growing public health concern in the UK and worldwide. Therefore, the study of breast cancer prevention, early detection, and management is paramount.

1.2.3 Risk factors

There are several accepted risk factors for breast cancer. These include mammographic density, increased age, family history, genetic mutations, body mass index (BMI), nulliparity, age at menarche, age at menopause, age at first birth, breastfeeding, hormone use, alcohol consumption, smoking tobacco, physical activity, and benign breast disease (BBD) (Table 1).[6,7] As seen in the table, MD is one of the largest risk factors associated with breast cancer, as women with very high MD have a more than a 4-fold increase in the risk of breast cancer compared with women with very low MD.[1] In this section, I will discuss breast cancer risk factors other than MD. A detailed discussion on MD is provided later in the chapter.

Increased age is a risk factor for breast cancer, with 80% of cases diagnosed in women aged over 50 years.[8] Age is a risk factor for several cancers, and this is likely due to increased cumulative exposure to oestrogens and carcinogens, which promote either cell proliferation or cell damage. The probability of cellular damage increases as age increases, leading to potential cancerous lesions.

Having a positive family history also confers an increased risk of breast cancer.[9] 13-19% of breast cancer patients report a first-degree relative (mother, sister, daughter) with breast cancer. Similar genetics, lifestyle habits, and other environmental exposures likely drive this.

There are some genetic mutations that are associated with an increased risk of breast cancer. These can be broadly divided into high penetrance variants, which significantly increase the risk of breast cancer, or low penetrance variants, which are associated with a small increase in breast cancer and are more widely present in the general population. There can also be pathogenic variants of lower penetrance in high risk genes. High penetrance genes include BRCA1, BRCA2, TP53, PTEN, CDH1, CHEK2, ATM and PALB2 mutations. They are associated with an increased risk of carcinogenesis as the normal cellular mechanisms for regulating cell division are suppressed, leading to cellular damage and the proliferation of cancerous cells.[8] Polygenic risk scores (PRS) estimate an individual's genetic susceptibility to breast cancer. They are based on the cumulative effect of multiple genetics variants called single nucleotide polymorphisms, derived from genome-wide association studies. They can explain over 30% of breast cancer heritability. Women with higher PRS scores have a greater risk of breast cancer than those with average scores.[10]

It is well acknowledged that endogenous hormone levels, oestrogen, progesterone and testosterone, affect breast cancer risk in both pre- and postmenopausal women.[11,12] It is therefore likely that events which affect levels of these hormones would alter the risk of breast cancer. However it is not entirely clear how hormones affect breast cancer risk, what is the role of relative concentrations are and how their effects vary over time. Factors affecting hormones include earlier age at menarche and later age at menopause. Each year decrease in age at menarche is associated with a 5% increase in the relative risk (RR) of breast cancer (RR 1.05; 95% CI 1.33, 1.52). [13] This is possibly because having more menstrual cycles reflects an increase in ovarian activity and, therefore, an increase in cumulative exposure to endogenous hormones. It could also be a marker of higher levels of hormones that lead to earlier menarche, or the effect of genetic and non-genetic factors that influence both timing of menarche

and risk of breast cancer. Breast cancer risk is also greater in premenopausal women than in postmenopausal women of the same age.[13] For example, the relative risk of developing breast cancer in pre- versus postmenopausal women of the same age (45-54 years) is 1.43 (1.33, 1.52). [13]

Like endogenous hormone exposure, exogenous hormone exposure is also associated with an increased risk of breast cancer.[14] This will be discussed in greater detail in Chapter 5 of the thesis. Menopausal hormone therapy (MHT), also known as hormone replacement therapy or postmenopausal hormone therapy, is a treatment used to relieve the symptoms of menopause by supplementing the oestrogen and progesterone that decline after menopause. Its use stimulates breast glandular tissue proliferation, which increases the risk of breast cancer and reduces the sensitivity of breast screening. The increased risk associated with MHT subsides to some extent after cessation, although some increased risk persists. Oral contraceptive use is also associated with an increase in breast cancer risk, but this risk appears to subside completely 10 years after cessation.[15]

Higher parity and earlier age at first birth are associated with a reduced risk of breast cancer. Parous women have a 14% reduction in the risk of breast cancer compared to nulliparous women (0.84; 0.80, 0.89) [16] and each year increase in age at first birth is associated with a 4% increase in the risk of breast cancer (1.04; 1.02, 1.04). [17] This is thought to be due to the terminal differentiation of breast epithelial tissue, which reduces the proliferative potential of the breast tissue and reduces the probability of cancerous proliferation. Breastfeeding also reduces breast cancer risk (RR per 12 months of lactation: 0.96; 0.94, 0.97).[18] This is perhaps due to suppressing ovulatory cycles during lactation, which reduces the overall exposure to oestrogens and progesterone.[16,18]

Obesity is associated with breast cancer risk in postmenopausal women.[19] This is likely due to increased inflammatory states and the aromatisation of adipose tissue to oestrogens, which increases the concentration of endogenous oestrogens and so increases cell proliferation in the breast.[20] the association between adiposity and breast cancer in premenopausal women is less consistent.[21]

Several lifestyle factors are associated with breast cancer risk. Alcohol consumption increases the risk of breast cancer.[22] Possible explanations for this include the effect of alcohol on increased oestrogen production and metabolism. While the evidence that smoking is a risk factor for breast cancer is inconsistent, the carcinogens found in tobacco can penetrate the breast tissue and increase the risk of mutations in important cell regulatory genes such as TP53.[8] Physical activity is thought to be protective against breast cancer through its effect on hormones, insulin levels and inflammation, although this is not fully understood.[23–25]

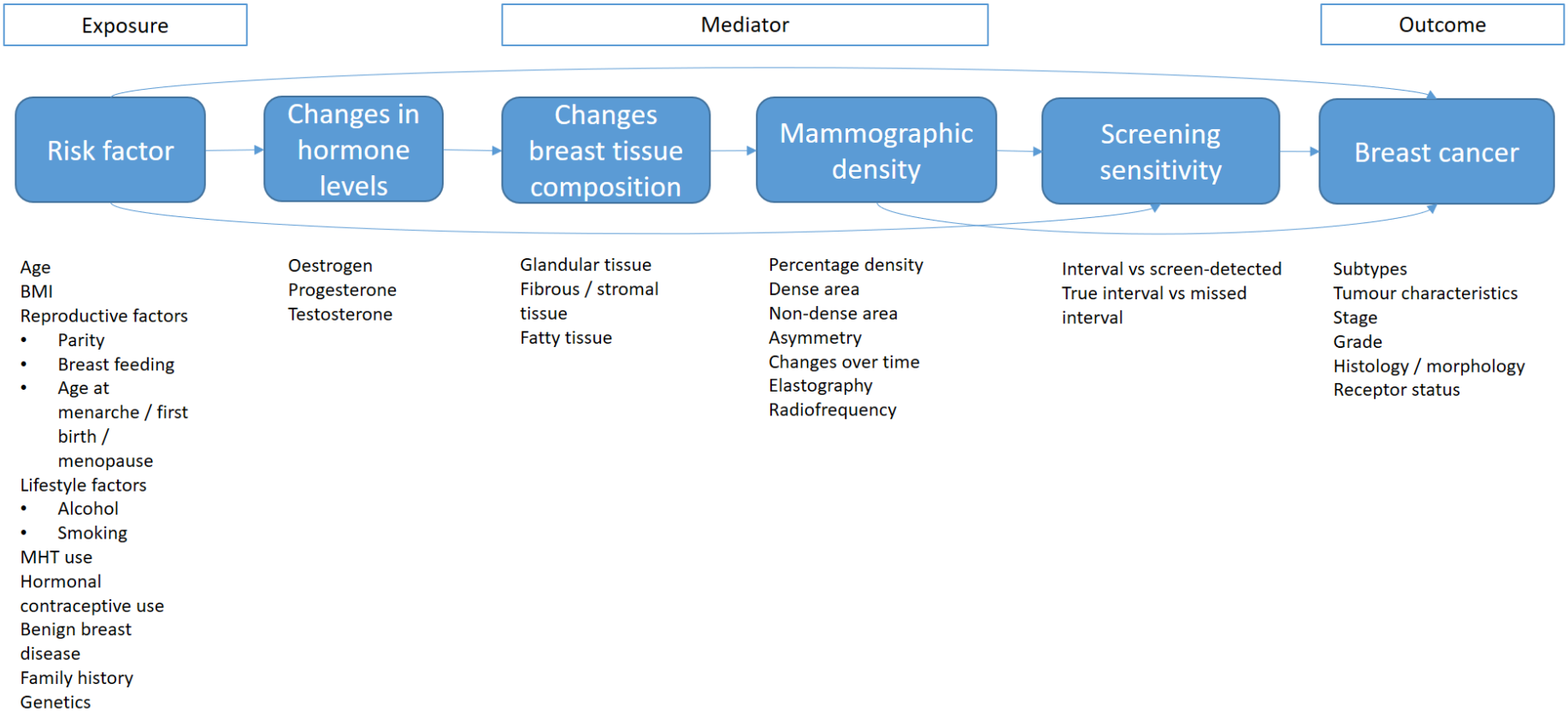
Non-cancerous breast lesions, namely proliferative benign breast diseases (BBD), such as ductal or lobular hyperplasia or intraductal papillomas, are also associated with breast cancer.[26,27] The association between BBD and breast cancer may arise because the supposedly benign condition is potentially a precursor to malignant transformation. An alternative and more likely theory is that the pathways which contribute to BBD development overlap with those which contribute to breast cancer development. Thus, having BBD reflects an environment conducive to breast cancer development. Atypical hyperplasia is a benign breast condition where cells in the breast increase in number and develop unusual morphology. Atypia is associated with a greater risk of breast cancer compared to other benign breast diseases (RR 4.24; 3.26, 5.41).[26]

Figure 1 details the pathways that will be explored in this thesis. Essentially, the figure depicts that a risk factor could influence breast cancer risk through different pathways: (i) through MD, (ii) through screening sensitivity, therefore delaying diagnosis to a later stage, or (iii) directly on breast cancer risk, or a combination of the three. While the aetiology and pathophysiology of breast cancer have been extensively studied, the causal pathways from risk factors to breast cancer could be better understood to help efforts in prevention and early detection. It is evident that the risk posed by MD is greater than the other risk factors (Table 1), suggesting that MD may, to some extent, mediate some risk factors. This will be addressed in section 1.3.

Table 1. Relative risks of breast cancer for various breast cancer risk factors

Risk factor		Relative risk (95% CI)
Mammographic density[1]	>75% vs <5% density	4.64 (3.64, 5.91)
Menopausal status [13]	Pre- vs postmenopausal women (45-54 years old)	1.43 (1.33, 1.52)
BMI (postmenopausal)[19]	>30 vs <23.7	1.60 (1.43, 1.79)
MHT use[14]	Ever vs never	1.33 (1.31, 1.35)
Oral contraceptive use[15]	Current vs never	1.24 (1.15, 1.33)
Age at first birth[17]	Per year increase	1.04 (1.02, 1.05)
Parity[16]	Parous vs nulliparous	0.84 (0.80, 0.89)
Age at menarche[13]	Per year decrease	1.05 (1.33, 1.52)
Breastfeeding[18]	Per 12 months of lactation	0.96 (0.94, 0.97)
Alcohol[22]	Per 10 g/day increase	1.07 (1.06, 1.09)
Smoking[22]	Ever vs never	1.03 (0.98, 1.07)
Benign breast disease[26]	Yes vs no	1.56 (1.45, 1.68)
Atypical hyperplasia [26]	Yes vs no	4.24 (3.26, 5.41)

Figure 1. Pathways of breast cancer development in relation to mammographic density explored in this thesis

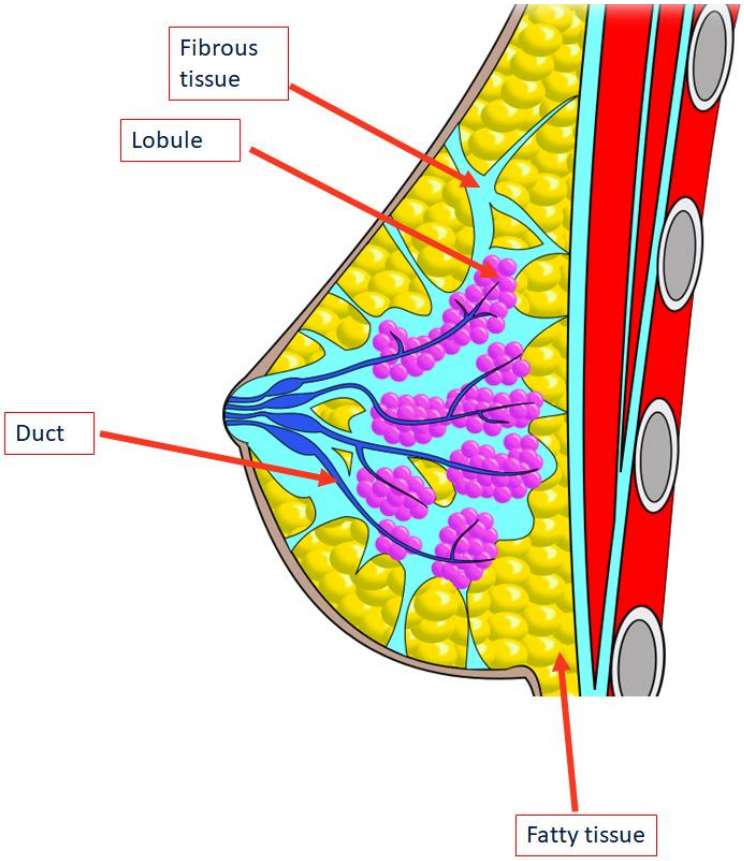


1.3 Mammographic density

1.3.1 Definition of mammographic density

Mammographic density is a radiological sign and is defined as the extent of radiologically dense breast tissue. To best understand MD, it is first essential to appreciate the various tissues present within the breast (Figure 2). The breast is composed of the parenchyma, which is the glandular epithelial tissue (ducts and lobules), fibrous connective tissue or stromal tissue, which supports the glandular tissue, and fatty tissue. Dense tissue comprises of the fibrous and glandular epithelial tissue. It is radio-opaque, appearing white on an x-ray mammogram. This contrasts to fatty tissue, which is radiolucent and hence appears dark on a mammogram. Women with denser breasts will therefore have larger areas in white on their mammograms when compared to women with less dense breasts, who will have larger areas in black.

Figure 2. Cross-section of the breast



(Image created by modifying "Breast anatomy" by © Sheldahl (https://commons.m.wikimedia.org/wiki/File:Breast_anatomy.png) Licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)) [28]

1.3.2 Measuring mammographic density

There are multiple methods of measuring and quantifying MD. These largely fall into categorical or continuous measures. Categorical measures of MD include BIRADS score, Boyd scale, Tabar scale, or Wolfe scale (Table 2).[29] These measures are usually determined by a reader, such as a radiologist, radiographer or qualified technician, who will visually assess the image and assign a score. The earliest of these was the Wolfe scale in 1976, which categorised variations of the breast parenchyma into four different patterns.[30] This scale was used historically but has been widely replaced by BIRADS. The Boyd scale attempted to assign a quantitative value to make the assessment more reproducible and is often used in epidemiological research. In this scale, MD was divided into six categories from 0% to >75% dense tissue.[31] The Tabar classification, a modification of the Wolfe scale, has five categories to account for the anatomical distribution or architecture of tissues within the breast seen on a mammogram.[32] This is used more in pathological risk assessment rather than routine screening. The Breast Imaging Reporting and Data System (BIRADS) classification was an attempt to standardise the reporting of mammograms. [29] The classification has undergone changes since the 1990s but broadly describes four categories of breast density patterns, also using a quantitative criterion of quartiles of percentage glandular tissue. The BIRADS categories, as viewed on an x-ray mammogram, are shown in Figure 3.

Continuous measures have been used since the 1990s and include percentage density, dense area, or non-dense area. These are measured using either computer-automated (often using artificial intelligence) software or computer-assisted tools, whereby a reader traces around the dense area of the mammogram to calculate the absolute dense area. Percentage density is calculated by dividing the dense area by the total breast area. Non-dense area is calculated by subtracting the dense area from the total breast area. As mammograms are almost exclusively digitised nowadays, computer-calculated measurements of density are widely used but require specific training and machine learning algorithms to process and to read the x-ray mammogram. Furthermore, AI-based or automated devices will need European Conformity (CE) marking, indicating that the device meets European safety

and quality standards for medical devices and can be sold within the European market for breast density analysis.

Percentage MD refers to the proportion of the breast composed of MD relative to the total breast area, on a mammogram. It accounts for differences in breast size, making it a relative measure of density. Absolute MD is the total area or volume of dense tissue in the breast, regardless of overall breast size.

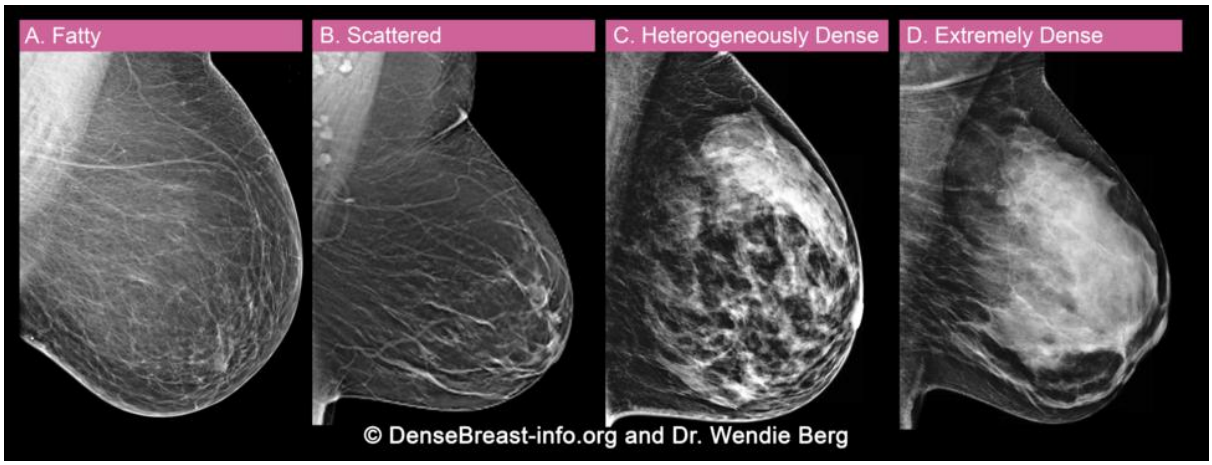
Given that many forms of MD estimates require a visual assessment, they will be subject to measurement error as readings will vary by assessor, time and image quality. Readings are not always reproducible.[33] A systematic review on the reproducibility of BIRADS scores found that 13-19% of women moved between “dense” and “non-dense” categories.[34] Therefore, comparisons across different breast units may be challenging. Whilst there are newer automated systems and models, it is not clear who effective they are at evaluating MD in relation to breast cancer risk, compared to qualitative methods, and will still be limited by image quality. Whilst these methods may eventually aim to standardise measurements, they will require more comparative research.

Whilst x-ray mammography has been the typical tool to visualise the breast and measure MD. Other tools include using MRI which produces radio-frequency pulses to generate detailed images of the breast, and ultrasound elastography to assess tissue stiffness.[35,36]

Table 2. Classifications of categorical measures of mammographic density

Density measure	Definition [29]
BIRADS Score	1/a: Almost entirely fat (<25% glandular) 2/b: Scattered fibroglandular densities (25% - 50% glandular) 3/c: Heterogeneously dense (51% - 75% glandular) 4/d: Extremely dense (>75% glandular)
Boyd Scale	0%: No dense tissue <10% dense tissue 10-25% dense tissue 26-50% dense tissue 51-75% dense tissue >75%: Extremely dense
Tabar Scale	I: Symmetry of all components with slightly greater fibrous tissue II: Predominance of fat tissue III: More fat tissue with fibrous tissue in the retroareolar region VI: Predominantly nodular densities V: Predominantly fibroglandular
Wolfe Scale	N1: Completely fatty with a few fibrous connective tissue P1: Fatty with prominent ducts ≤4 mm in diameter P2: Higher concentration of prominent triangular ducts in the central portion DY: Homogenous density with few ductal prominence

Figure 3. BIRADS categories of mammographic density as viewed on an x-ray mammogram



(Image from https://densebreast-info.org/. Image reproduced with permission from DenseBreast-info, Inc.) [37]

1.3.3 Mammographic density as a risk factor for breast cancer

1.3.3.1 Association of mammographic density with breast cancer risk

Higher MD has consistently been shown to increase breast cancer risk. This increased risk persists even after controlling for other known risk factors such as age, family history and hormone use, and after accounting for the effects on screening sensitivity.[38,39]

Since the 1970s, studies have reported the association between MD and breast cancer risk.[30,40] Women with high mammographic density have a 4-6 times higher risk of breast cancer compared to women with low density.[39,41] The increased risk of breast cancer seems to remain for 8-10 years after the date of the original mammogram to determine MD.[42] However, it is not clear how long MD remained high for, after the original mammogram. But it does suggest that an isolated MD reading is able to provide some indication of future breast cancer risk.

A meta-analysis from 2006 of over 200,000 women from 42 studies found that those with extremely dense breasts (≥75%) had 4.64 times the risk of developing breast cancer than those with very little dense tissue (<5%).[1] The association was stronger among studies conducted in the general population than among symptomatic women. Associations were also stronger in studies using percentage MD rather than a categorical measure of MD. In a recent meta-analysis of nine studies published in 2022, having extremely dense (BIRADS D) breast tissue resulted in over a 2-fold increased breast cancer risk

compared to having scattered dense (BIRADS B) breast tissue (OR 2.11; 95% CI 1.84, 2.42), and the same women had almost a 4-fold increased risk compared to women with little or no (BIRADS A) density (3.89; 2.47, 6.13). The authors speculated that the lower relative risks in the more recent review may have been due to the use of digital mammography, or different classifications of MD.[38]

A study from three Canadian breast cancer screening programmes reported that 26% of all breast cancers and 50% of cancers detected within a year of a negative breast screen were attributable to having >50% percentage MD.[43] The screening study additionally reported that MD was associated with both cancers detected at screening and cancers detected in the interval between screens and that these associations persisted for 8 years. This suggests that MD increases the risk of breast cancer, independently of its effect of screening sensitivity.

Whilst MD declines with age, some women may have persistently high MD due to genetics, reproductive or hormonal factors. A meta-analysis of nine observational studies aimed to evaluate the association between MD changes over time and the risk of breast cancer.[44] They found that increased MD over time was associated with higher breast cancer risk (HR 1.61; 95% CI 1.33, 1.96) whereas decreased MD over time was associated with lower breast cancer risk (HR 0.78; 95% CI 0.71, 0.87), in cohort studies. In case-control studies increased MD over time was also associated with higher breast cancer risk (OR 1.85; 95% CI 1.29, 2.65). This therefore suggests that MD changes over time are important to monitor when determining breast cancer risk as this likely represents ongoing hormonal exposure and the tissue environment within the breast. Understanding this can help guide supplemental screening for women who have ongoing higher MD.

If changes in MD affects breast cancer risk, then perhaps modifying MD may also modify breast cancer risk. As discussed earlier, MHT increases MD and increases breast cancer risk. Tamoxifen is a selective oestrogen receptor modulator used primarily for the prevention and treatment of ER positive breast cancer. It works by blocking oestrogen receptors in breast tissue, reducing oestrogen-driven cancer growth.[45] Tamoxifen is indicated in some women with a high risk of breast cancer. Tamoxifen has been

shown to reduce MD.[46,47] A nested case-control study within a randomised prevention trial of tamoxifen versus placebo, with 123 breast cancer cases and 943 controls, was conducted.[48] Women who took tamoxifen and experienced a 10% or greater reduction in MD experienced a 63% reduction in breast cancer risk (OR 0.37; 0.20, 0.69). Women who did not have a 10% or greater reduction in MD did not have a reduction in breast cancer risk, neither did the women in the placebo group. These findings do suggest that modifying MD does modify breast cancer risk.

Whilst the exact mechanism for this association is not fully understood, there are suggested hypotheses. Given that MD is a reflection of the proportion of fibroglandular tissues from which most breast cancers arise (known as adenocarcinomas), a higher MD may reflect increased proliferation and hence greater risks of mutations which lead to breast cancer. As women with higher MD have more glandular epithelial cells within the breast, there is a greater chance that a cell could undergo malignant transformation. Women with higher MD also have higher levels of endogenous sex hormones.[49,50]

Whilst mammographic density is effectively a radiological phenomenon, it is also important to consider what it represents. Higher MD reflects an increased exposure to oestrogens and progestogens, increased epithelial proliferation (mitogenesis), extracellular matrix stiffness, and a tumour-promoting microenvironment if there is genetic damage to proliferating cells caused by mutagens (mutagenesis).[51] These increase breast cancer risk. [51,52] Therefore, whilst it is the whiteness on the mammogram is the objectively being measured to produce a density estimate, it in reality is representing a complex interplay between hormones, cellular pathways and tissue composition, which is on the pathway to breast cancer development.

As mentioned earlier, MD can also be classified in terms of absolute volume or area. This will likely represent the total degree of fibroglandular tissue rather than the relative amount. It is perhaps more relevant for women with high BMI, who may have lower percentage MD but higher absolute MD. Both methods are predictive of breast cancer risk. [53] Some studies have suggested that percentage dense

area is a stronger risk indicator than absolute dense area[54] whilst others have suggested an absolute measure is more predictive of breast cancer risk.[55]

1.3.3.2 Association of mammographic density with other breast cancer risk factors

Several breast cancer risk factors are associated with MD, as shown in Figure 1. Menopausal status, age, MHT use, and BMI are among the most established of these risk factors, with consistent evidence for these associations and clear biological plausibility to explain the associations.

Premenopausal women have denser breasts than postmenopausal women. This is likely because of the reduced levels of circulating reproductive hormones after menopause when the ovarian function declines, which leads to breast tissue involution. A collaborative pooled analysis of cross-sectional data from 22 countries worldwide reported a decrease in the square root of percentage MD of 0.46% (95% CI -0.53, -0.39) when comparing postmenopausal to premenopausal women.[56] A study of 24,840 screening mammograms in California reported that premenopausal women had a 3.50% (0.62, 6.38) higher in mean percentage MD compared to postmenopausal women.[57]

Increasing age is associated with a reduction in MD in both pre- and postmenopausal women. The decrease in MD with age is likely due to the reduction in endogenous hormones, the loss of which leads to a reduction in the production and proliferation of fibroglandular tissue and the loss of utility of breast function, as well as the shrinkage of breast tissue with age.[58] The collaborative pooled analysis cited above reported a decrease in the square root of percentage MD per 10-year increase in age of 0.24% (-0.34, -0.14) in premenopausal women, and of 0.38% (-0.44, -0.33) in postmenopausal women.[56] Similarly, an occupational cohort of 1700 women in Mexico showed a decrease in percentage MD per 10-year increase in age of 3.80% (-4.02, -3.58) in premenopausal women, and 2.50% (-2.73, -2.27) in postmenopausal women.[59]

MHT use is associated with an increase in MD. This is likely due to the effect of the exogenous hormones stimulating breast tissue proliferation. A systematic literature review of 22 studies, including six trials and 16 observational studies, reported that MHT use has been consistently found

to increase MD.[60] There was no pooled summary reported in the review. In the Women's Health Initiative trial in the United States, 435 women were randomly assigned to either receiving oestrogen-only MHT or placebo. After one year of follow-up, mean percentage MD increased by 1.6% (0.8, 2.4) among women receiving oestrogen MHT and decreased by 1.0% (-1.7, -0.4) in the placebo group.[61] They also randomised 413 women to combined MHT or placebo. The following year, mean percentage MD increased by 6.0% (4.6, 7.5) among women receiving combined MHT and decreased by 0.9% (-1.5, -0.2) in the placebo group.[62]

Higher BMI is associated with lower percentage MD. This is likely due to an increased deposition of fatty tissue in the breast as BMI increases, which in turn reduces the overall proportion of fibroglandular or dense tissue within the breast. A large observational study with 24,556 women in the San Francisco Mammography Registry, which was designed to examine the association between BMI and MD, found that each BMI increase of 1kg/m^2 was associated with a 2.03% reduction in percentage density (-2.09, -1.98).[58] A study reported that dense area is also inversely associated with BMI.[63] However, another study reported that higher BMI was positively associated with absolute dense volume, therefore indicating that women with higher BMI still retain a higher degree of glandular tissue, which may explain why higher BMI is associated with an increased breast cancer risk.[64]

There is evidence that genetics can also determine MD.[65] Women with a family history of breast cancer have higher MD.[66] Furthermore, genome-wide association studies have identified several loci associated with MD.[67]

There are several studies examining the association of MD with reproductive and lifestyle factors. However, the results are inconsistent. Small sample sizes perhaps affected the reliability of these results. These factors include alcohol, smoking, parity, age at first birth, and age at menarche. There remains a lack of consensus regarding the relationship between these risk factors and MD in the literature. This will be explored later in this thesis.

Studying and recognising the determinants of MD can help identify women at high risk of breast cancer who may benefit from more frequent or sensitive screening practices or preventative initiatives. Furthermore, understanding these determinants can elucidate the pathophysiology and mechanisms of breast cancer development and causal pathways.

1.3.3.3 Mammographic density as a mediator of known risk factor associations

Given that MD is associated with breast cancer risk factors and breast cancer risk, it has been suggested that MD mediates the association between these risk factors and breast cancer and is not just an independent risk factor for breast cancer (outlined in Figure 1). MD is likely on the causal pathway. This is because of underlying tissue biology within the breast. High MD reflects a greater proportion of fibroglandular tissue, from which breast cancers can arise. Having more fibroglandular tissue will likely increase the probability of a breast cancer. It therefore follows that risk factors which affect breast tissue composition will also influence MD and breast cancer risk. A mediator differs from a confounder in that a mediator lies on the causal pathway between exposure and outcome. A confounder, however, influences both the exposure and the outcome outside of the causal pathway.

There is a paucity of data on MD as a mediator or intermediate phenotype. The data that is available suggests that MD may partially mediate some of the associations between some risk factors and breast cancer risk. A study of 1290 breast cancer cases and 3422 controls found that MD partially mediated the association for biopsy-confirmed BBD in all women, adolescent BMI in premenopausal women, and MHT use in postmenopausal women.[68] A further study of 3392 breast cancer cases and 8882 controls found that MD partially mediated the association between previous breast biopsy, nulliparity, age at first birth, current MHT use, and breast cancer risk in postmenopausal women.[69] These studies do suggest though that the influence of these risk factors on breast cancer risk may arise partially through changes in the composition of breast tissue. As with most mediation analyses, however, the causal interpretation warrants caution and can only be interpreted as statistical mediation rather than causal as there may be other underlying mechanisms which cannot be

measured or captured in a mediation analysis. Moreover, given the limited data on MD as a mediator, further studies in larger populations are necessary before drawing and acting upon conclusions.

Although a formal mediation analysis was not possible in this thesis as MD was only available for a limited number of women, investigating whether MD is a mediator is important in delineating the aetiology of breast cancer. If all the effect of a given risk factor on breast cancer is mediated through MD, then the excess fibroglandular, in denser breasts, is the primary intermediary step to breast cancer. If, however, only some of the effect of a risk factor is mediated through density, then these risk factors are also acting through different pathways to develop breast cancer. Determining these mechanisms can help identify pathways through which risk factors influence breast cancer development and potential targets for breast cancer prevention. Furthermore, as MD is likely an intermediate marker or surrogate of breast cancer risk, it could potentially be used as a surrogate marker for breast cancer risk in intervention trials for breast cancer prevention, if measuring MD at the start of the trial and after the intervention.[70] An important first step in determining whether MD is a mediator, is investigating associations between breast cancer risk factors and MD.

1.4 Diagnosing breast cancer

In this section, I will introduce population-based breast screening in the UK and discuss other ways in which breast cancer can be detected.

1.4.1 Breast screening programme

Population-based breast cancer screening programmes have been introduced in several countries over the last four decades, including in the UK.[71,72] Breast cancer screening programmes aim to facilitate breast cancer diagnosis at an earlier stage when treatment offers better prognosis.[73]

The UK NHS Breast Screening Programme (NHSBSP) was implemented in 1988.[73] Every three years, women who are registered with a GP are invited to attend for a routine screening mammogram. Initially, women aged 50-64 were invited for screening. The upper age limit was extended from 64 to 70 years in 2004. X-ray images are taken of the breast from two views: craniocaudal (view taken from

above the breast) and mediolateral oblique (view taken from a diagonal angle).[74] Mammograms were initially stored and read on a film. Since the introduction of digital mammography in the UK in 2008, however, they have been stored and read digitally due to technical advances and practical advantages.[75,76]

Mammograms are reported separately by two radiologists or radiographers with specialist training. If a consensus between the two readers cannot be found, a third reader is involved. Women with abnormal mammograms are recalled for further assessment, which may include a physical assessment, further mammograms, ultrasound scans, or tissue biopsies. Women who are diagnosed with breast cancer are then referred to a specialist team for investigation and management. Where no further tests are needed, the women will return to routine recall and be invited for screening three years later – provided they have not reached their 71st birthday.[77] Women aged over 70 years can self-refer for screening every three years.[78]

1.4.2 Routes to breast cancer diagnosis

Breast cancer can either be diagnosed through screening or following the development of symptoms such as a breast lump, axillary lump, breast skin changes, or nipple changes. The women in the latter group present to a general practitioner (GP) and then are referred to secondary-care breast services for further assessment and investigations. Urgent referrals are made when the GP suspects the presenting symptom may be due to cancer, and the patient is therefore seen by breast services within two weeks of the referral. Routine referrals are made when the GP does not suspect a breast symptom to be due to cancer.

Cancers detected through screening are termed “screen-detected,” and cancers diagnosed between screens, when symptoms may have developed, are termed “interval cancers”. The “interval period” is the time period between two consecutive screens. Interval cancers can be further divided into “true interval” or “missed interval”. True interval cancers are those which were never present at the preceding screen. Missed interval cancers are those which were missed at the screen and picked up

on a retrospective review of mammograms, after a diagnosis has been made. A summary of the routes to a breast cancer diagnosis among women who attend routine screening are given in Table 3.

Table 3. Routes to breast cancer diagnosis among women who attend routine breast screening in the UK

Route [79]	**Description**
Screen-detected	Detected via the NHS Breast Screening Programme
Routes to interval cancer diagnosis	
Urgent referral	Urgent referral with a suspicion of cancer made by GP or other health care professional
Routine Referral	Referrals where the patient has a breast symptom but was not referred urgently as cancer is not suspected

1.4.3 Breast cancer prognosis by route of detection

Prognosis refers to the likely outcome of the disease.[80] Factors which influence breast cancer prognosis differ by route of detection. These factors include stage and grade of breast cancer, as well as hormone receptor status status.[81]

The stage of a cancer is a combined measure of the tumour size, nodal status, and presence or absence of distant disease (Figure 4). Higher stages of cancer are associated with a worse prognosis. The stages of breast cancer and their corresponding survival rates are shown in Table 4.

Figure 4. Stages of breast cancer

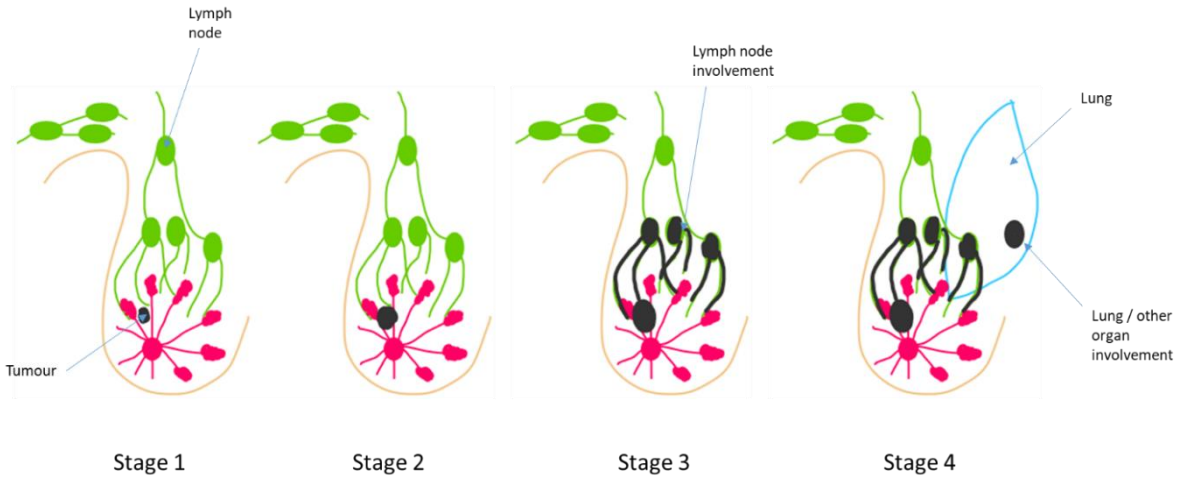


Table 4. Stages of breast cancer and survival in England in 2018[82,83]

Stage	1-year survival % (95% CI)	5-year survival % (95% CI)
1 (Early)	100.0 (99.8, 100)	97.9 (97.3, 98.5)
2 (Early)	98.9 (98.7, 99.0)	89.6 (89.0, 90.2)
3 (Locally advanced)	95.5 (95.1, 95.9)	72.0 (70.5, 73.5)
4 (Metastatic)	66.0 (65.0, 66.9)	26.2 (24.8, 27.7)

The TNM staging for breast cancer is another method of classifying breast cancer based on the size of the tumour (T), spread to lymph nodes (N) and the presence of metastases (Table 5). This can be used to determine prognosis, guide treatment and compare treatment results.[84]

Table 5. TNM staging of breast cancer

Tumour size	Nodes	Metastasis
	N0: no lymph node involvement	M0: No metastasis
T1: <2cm	N1: involvement of 1-3 axillary lymph nodes	M1: distant metastasis present
T2: 2-5cm	N2: involvement of 4-9 nodes or internal mammary lymph nodes	
T3: >5cm	N3: involvement of 10+ axillary lymph nodes or infra/ supraclavicular lymph nodes	
T4: extends to skin or chest wall		

The grade refers to the degree to which the cancer cells resemble the cells of the surrounding breast tissue under a microscope. A lower grade indicates the cells are similar and are therefore termed highly differentiated. A higher grade indicates that the cancer cells look significantly different from the normal breast tissue and are, therefore, termed poorly differentiated. Cancers of a higher grade grow faster and spread more rapidly than those of a lower grade. Hence, the former are associated with a poorer prognosis and are often detected outside of screening.[85]

A further factor which can determine prognosis is the expression of hormone receptors: oestrogen (ER), progesterone (PR), and human epidermal growth factor (HER2).[86] Cancers which are ER-, PR- and HER2- (triple negative breast cancers) are usually considered the most aggressive and have a poorer prognosis than hormone receptor positive cancers, and are also more likely to be detected outside of screening.[87]

There is evidence to suggest that interval breast cancers are associated with a poorer prognosis when compared with screen-detected cancers due to a combination of later-stage and more aggressive tumour biology.[88–90] A meta-analysis of three randomised screening trials reported that patients with interval cancers had a 53% greater risk of breast cancer mortality than patients with screen-detected breast cancers.[91] Several epidemiological studies have found that women diagnosed with an interval cancer have a poorer prognosis than women with a screen-detected cancer.[91–101] This, however, has been refuted by other observational studies.[102–104]

1.4.4 The role of mammographic density and other risk factors on mammographic screening sensitivity

The ability to detect a cancer at screening depends on the screening sensitivity (Figure 1). Any factors that affect the sensitivity of a breast screen will affect the risk of an interval cancer, where prognosis may be unfavourable. Treatment for early breast cancer has improved considerably over the past 40 years, which is reflected in the reduction in mortality from breast cancer.

X-ray mammogram is the most utilised method for breast cancer screening and imaging breast tissue. However, this imaging modality has its limitations. Given that an x-ray mammogram is a two-

dimensional imaging modality, different layers of tissues within the breast cannot be distinguished. This will limit the sensitivity of the screen.

There are several factors which can affect screening sensitivity, the main one being MD. Increased MD masks radiographic evidence of a tumour, as both density and potential cancerous lesions appear white on an x-ray mammogram. This makes it challenging to distinguish between the tissues radiologically.[2] Understanding the impact of MD on screening sensitivity can inform future screening practices and early detection. Women with high MD are more likely to have false negative screens and have their breast cancers diagnosed at a later stage during the interval period where prognosis may be unfavourable, or treatment options may be more invasive. Women with high MD may therefore, benefit from more frequent screening or more sensitive imaging options, such as an ultrasound or magnetic resonance imaging (MRI), to ensure that any potential lesions are detected.[105] Digital breast tomosynthesis (DBT), also known as 3D mammography, is increasingly considered for younger women, particularly those with dense breasts or have breast symptoms. DBT has been shown to have better sensitivity than x-ray mammography.[106,107] Tomosynthesis has been approved for use in the NHSBSP as an optional extra tool in the assessment of screen detected soft tissue breast abnormalities but not necessarily for women with dense breasts.[108] Another modality is contrast-enhanced mammography. This involves the intravenous injection of a dye containing iodine in combination with a standard digital mammogram. This is being considered in the US for women with dense breasts as it improves the sensitivity of breast screening.[109,110]

Other factors affecting screening sensitivity include prior breast surgery or scarring due to radiation, breast implants, image quality, and radiologist experience.[111,112] Scar tissue, like dense breast tissue, is radio-opaque and would therefore appear white on an x-ray mammogram. Breast implants interfere with the view seen on a mammogram. At present there is no separate routine screening modality for women with breast implants in the UK.

At present, the Royal College of Radiologists in the UK does not recommend supplemental screening for women with dense breasts and women are not informed of their density after their screen.[113] This is likely because there are no clinical pathways to manage dense breasts detected at screening, in the UK. Therefore, if women were informed, they would have to pay privately to have further investigations which would create unequitable care. Women who could not afford this are likely to experience anxiety over the possibly of having cancer, which therefore has ethical implications.[33] Furthermore, the inconsistencies across the way MD is measured, leading to lack of reproducibility, would need to be addressed and harmonised before guidelines can be recommended. Given that the UK has a publicly funded health care system, the government would need to find the resources to fund new pathways to allow women with dense breasts to have further investigations or supplemental screening with any of the more sensitive modalities mentioned earlier. The US Food and Drug Administration now requires mammography facilities to inform women if they have dense breasts.[114] These women will then be able to discuss future screening and surveillance measures to ensure potential breast cancers are not missed. Such practices may need to be implemented across other screening programmes worldwide. However, this may also require cost-benefit evaluations as publicly funded programmes would need to find the resources to support potentially costly further investigations.

Ultimately, as MD is associated with reduced mammographic screening sensitivity, and is independently a risk factor for breast cancer, any factor which affects MD will affect both the risk for breast cancer and the risk of having it detected outside of screening – namely during the interval period. It is important to understand which factors affect screening sensitivity, as this can help identify women at risk of a missed cancer. These women can then be educated based on their modifiable risk factors or offered more personalised screening measures. Of particular importance is MHT. Exogenous hormones will stimulate breast tissue proliferation, therefore affect MD and potentially mammographic screening sensitivity. Understanding whether MHT affects screening sensitivity may guide personalised screening pathways, or guide how women should be counselled on the health

implications of MHT use. This will become increasingly more relevant as use is rising in the UK, with over 2.3 million MHT users.[115] This will be investigated as part of this thesis.

1.5 Breast cancer risk prediction

Breast cancer risk prediction models estimate an individual's likelihood of developing breast cancer over time based on various factors such as genetic predisposition, family history, lifestyle, hormonal factors, reproductive factors, environmental exposures, and mammographic features – including MD. Their purpose is to identify those at greater risk of developing breast cancer and, therefore, to guide preventive and early detection strategies.

There are several breast cancer risk prediction models.[116] Examples of these models include Gail/BCRAT, Tyrer-Cuzick, Claus, BRCAPRO, and Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA).[117] Some aim to predict the risk of breast cancer (Gail, Tyrer-Cuzick, Claus), others to predict the probability of carrying a high penetrance gene such as BRCA1 or BRCA2 genes (BRCAPRO), and others aim to do both (BOADICEA). Models use a range of risk factor, genetic, questionnaire, family history, imaging and MD data. Further details on risk prediction models are given in Chapter 6 of this thesis.

What is currently lacking in the literature are subtype-specific risk prediction models as only few exist. As different breast cancer subtypes may have different underlying biology, responses to treatment, and prognoses, it is important to determine which risk factors are associated with which subtype and to identify the women at increased risk of various subtypes of cancers that are not detected at screening, namely interval cancers and high grade cancers. This may help guide more tailored screening pathways for these women.

1.6 Aims of this DPhil

The overarching aim of my thesis is to explore the role of mammographic density and other risk factors in breast cancer risk and detection among women who attend routine breast cancer screening. The main hypothesis of this thesis is that MD is a key mediator of breast cancer risk and detection,

influenced by various risk factors, and can be used to guide risk prediction and screening pathways, particularly in postmenopausal women.

The main aims of the thesis are to:

1. Systematically review the evidence on associations of breast cancer risk factors with MD;
2. Conduct analyses of breast cancer risk factors and MD in postmenopausal women in the Million Women Study;
3. Investigate how patterns of MHT use affect mammographic screening sensitivity; and
4. Develop risk prediction models for breast cancer overall and for specific breast cancer subtypes.

In Chapter 2, I shall provide a detailed overview of the data sources that I shall be using and the population that I will be studying. This is the population of postmenopausal women in England who attend routine breast cancer screening, linked to data from the NHS Breast Screening Programme, NHS England cancer registries, and MD data.

In Chapter 3, I will present a systematic review and meta-analysis of the associations of alcohol, smoking, parity, age at first birth, and age at menarche with continuous and categorical measures of MD.

In Chapter 4, I will provide a detailed characterisation of MD within a subset of my study population and present cross-sectional analyses of breast cancer risk factors and MD.

In Chapter 5, I will present a background literature review and meta-analysis of MHT use and present a prospective analysis of the risk of interval versus screen-detected breast cancers, as a proxy for mammographic screening sensitivity.

In Chapter 6, I will develop a model for predicting breast cancer risk using risk factor information from questionnaires. I will then assess how well the general model predicts breast cancer subtypes based

on the mode of detection and grade of breast cancer before developing subtype-specific models. I will also determine which risk factors, if any, are more predictive of specific subtypes.

In Chapter 7, I will discuss the findings of the thesis, its strengths and limitations, the potential implications for screening and clinical practices, and provide recommendations for future work.

Chapter 2. Data sources

2.1 Introduction

The sources of data which I shall be using in this thesis are the Million Women Study (MWS) and linked screening data from the NHS Breast Screening Programme (NHSBSP), including mammographic density (MD) data derived from digital screening mammograms. This chapter introduces the MWS and describes its linkage to other sources of data for this thesis. The chapter also summarises some of the statistical methods used in this thesis.

2.2 Million Women Study

The MWS is an open-ended prospective study of the health of 1.3 million women in England and Scotland. The initial aim was to investigate the risk of breast cancer from menopausal hormone therapy (MHT) use.[118] The MWS has since examined the effect of other exposures, including social, behavioural, lifestyle, demographic and anthropometric factors, on additional health outcomes such as other cancers, cardiovascular disease, neurological disease, psychological disease, musculoskeletal disease and mortality.[119]

2.2.1 Study design

2.2.1.1 Recruitment and participants

Between 1996 and 2001 (median 1998), women in England and Scotland, aged 50-64, who were registered with a GP in regions covered by participating 66 breast screening centres, were invited to join the study. Participating centres are shown in Figure 5. These centres covered approximately half of the UK population. Approximately three million invitations to join the study were sent out. Alongside their routine breast screening invitation, participants received a recruitment questionnaire. They were asked to bring the completed questionnaire to their screening appointment or to post it back to the Million Women Study Coordinating Centre in the Cancer Epidemiology Unit at the University of Oxford. Recruitment continued until the end of 2001. Approximately 70% of those receiving an invitation for routine breast screening attended screening, and 70% of these women consented to participate in the study. There were 1.3 million women in the study at the end of recruitment. This represented 1 in 4 women in the UK in the eligible age group, born between 1935

and 1950.[120,121] Study participants provided written consent for re-contact and follow-up through linked NHS medical records. More information on the study is available on the MWS website (https://www.ceu.ox.ac.uk/research/the-million-women-study).

Ahead of the MWS, a pilot study of 6400 women in Oxfordshire and West London from 1994 to 1996, who were invited for routine breast screening, concluded that including the study questionnaire with a routine postal screening invitation did not adversely affect screening uptake.[122] Women were randomised to receive the screening invitation either alone or with the study questionnaire. Screening attendance was 71% in each group.

Figure 5. Locations of the 66 NHS breast screening centres through which participants of the Million Women Study were recruited.

2.2.1.2 Questionnaires

The recruitment questionnaire asked for details on name, address, NHS number, breast screening number, date of birth, age, place of birth, height, weight, level of education, smoking, alcohol consumption, physical activity, number of children, breastfeeding, previous breast screening attendance, previous breast surgery, previous cancer, mother or sister with breast cancer, comorbidities, history of hysterectomy, oophorectomy, sterilisation, oral contraceptive use, duration of oral contraceptive use, MHT use, type of MHT, duration of MHT use, age at menarche, and age at menopause if appropriate.

Since recruitment, subsequent questionnaires have been sent to surviving participants to update information on factors that may have changed since recruitment. These include anthropometric factors, medical history, lifestyle factors, and MHT use. The subsequent surveys additionally asked about activities of daily living, social activities, sleep, occupation, pets, greying of hair, specific medications, family medical history, and diet. Re-surveys have been sent at 3-, 8-, 12-, 15- and 20-years following recruitment. Through the recruitment surveys and subsequent re-surveys, information on 1400 variables has been collected. Copies of the questionnaires are available on the MWS website (https://www.ceu.ox.ac.uk/research/the-million-women-study).

2.2.1.3 Validation

To examine reporting errors and changes over time for certain characteristics, self-reported information has been validated in samples of the participating women against objective measures or routinely collected health records. Good agreement has been found between the reported information on hormone use, anthropometric factors, and reproductive factors and the validated objective measures or health records.[118]

2.2.1.4 Linked outcomes

Participants of the MWS are followed up annually by electronic record linkage to routinely collected healthcare data in England and Scotland. Follow-up remains virtually complete after 20 years, with about 1% loss to follow-up, primarily due to emigration.

In England, data on deaths, emigrations, cancer registrations, hospital admissions and hospital day cases, breast screening data and cancer outcome data are provided by NHS England. Data on cancers includes information on cancer site, cancer morphology and date of registration. Information on cases from pathology departments, oncology departments and death certificates are received to ensure a cancer is unlikely missed. The MWS is also linked to data on derived tumour characteristics for breast cancer, such as grade, stage, size, lymph node status and hormone receptor status (ER/PR/HER2). Primary care data on consultations and prescriptions are provided through linkage to Clinical Practice Research Datalink (CPRD). Data in Scotland on deaths, cancer registrations, hospital admissions, and primary care is provided by Public Health Scotland. The MWS participants are linked to these datasets by name, date of birth and NHS number.

Breast screening data

The NHSBSP, which began in 1988, invites all women registered with a GP, aged 50-70, for routine breast screening every three years. At the time of recruitment, women aged 50-64 were invited. This was extended to 70 years between 2001 and 2004. The NHSBSP uses a central database to invite women to breast cancer screening, which contains information on individual screening episodes. A screening episode is defined as the period during which all screening-related activities occur, including invitation and diagnostic assessment. The data includes the date the episode started, the type of episode (routine, self-referred) and screening attendance. In 2013, the MWS cohort was linked to records on screening episodes for women in England, which had started by 31st December 2012. Each MWS participant was linked via participant identification numbers, matched to their name, date of birth, and NHS number on screening records. The database currently only has records from women recruited and living in England.

Mammographic density data

MWS also has linkages to MD data. Since ~2010, mammograms from the NHSBSP have been stored digitally rather than on an x-ray film. This has allowed for computer-automated measures of mammographic density to be calculated. In 2023, the MWS began collecting digital mammograms and

screening outcomes for some of the screening centres participating in the MWS. The processes used to collect and store these digital mammograms have been developed and implemented as part of the Cancer Research UK funded "OPTIMAM" project, which aimed to improve the efficiency and accuracy of the NHSBSP. Currently, MD data is only available for approximately 10,000 women who attended screening in the Oxford screening centre between 2011 and 2018.

2.2.2 Ethics

The study has ethical approval from the East of England-Cambridge South Research Ethics Committee (Ref 97/5/01). Ethics approval for the recruitment into the study in 1996 was initially obtained by region. In 1997, approval for a multi-centre cohort study was granted. The analyses contained in this thesis are all in keeping with the original aims and restrictions of the original ethics approval for the MWS.

2.2.3 Population characteristics

Characteristics of the study population at recruitment are shown in Table 6. Mean age at recruitment was 56.2 years. 33% of the population were current MHT users at recruitment.

The study population of the MWS are broadly representative of the target population within the general UK population, at the time of recruitment (Table 7).[118,123]

Table 6. Selected characteristics of women in the Million Women Study at recruitment

Characteristics at recruitment	Total
Number of women	1,364,313
Median year of birth (IQR)	1942 (1938-46)
Age at recruitment	56.2 (4.9)
Age at menopause	47.5 (5.8)
Height (cm)	162.0 (6.6)
Weight (kg)	68.9 (12.7)
Body mass index (kg/m2)	26.2 (4.7)
Current smokers	262,848 (19.4%)
Alcohol drinks per week	4.1 (5.3)
Parous	1,215,160 (89.2%)
Number of pregnancies in parous women	2.2 (1.3)
First degree relative with breast cancer	128,512 (10.1%)
Ever oral contraceptive user	797,115 (59.1%)
Ever breastfed among parous women	652,868 (67.9%)
Ever had breast surgery	182,345 (13.5%)
Deprivation quintiles	
Q1, least deprived	272,358 (20.1%)
Q2	271,739 (20.1%)
Q3	270,132 (19.9%)
Q4	270,380 (20.0%)
Q5, most deprived	269,677 (19.9%)
Region of recruitment	
Oxford	86,318 (6.3%)
East Anglia	69,370 (5.1%)
South West	283,630 (20.8%)
Thames	183,487 (13.4%)
West Midlands	119,262 (8.7%)
North Yorkshire	151,975 (11.1%)
Trent	159,153 (11.7%)
Mersey	92,318 (6.8%)
Manchester/Lancashire	101,632 (7.4%)
Scotland	117,168 (8.6%)
Current MHT users	445,233 (33.0%)
Results shown are N (%) or mean (SD) or median (IQR)	

Table 7. Comparison of MWS study participants with the general population

Characteristics at recruitment	MWS participants (average age 56)	General population* (women aged 55-64)
Deprivation (% in least deprived quintile)	20	20
Education (% with tertiary qualifications)	13	15
Height (mean cm)	162.0	159.9
BMI (mean kg/m^2)	26.2	27.7
Alcohol (mean units/week)	4.1	5.5
Ever smoker (%)	49	49

**based on 1996 Health Survey for England and 2001 Census*

2.3 Exposure and adjustment variables

Exposure and adjustment variables used in the analysis of this thesis are taken directly from the self-reported recruitment questionnaires, except for deprivation. Each results chapter explains the exact coding and categorisation of the variables.

The variables that were taken from the surveys were MHT use (never, past, current), MHT duration of use (years), MHT type (oestrogen-only, combined, other), ever oral contraceptive use (yes/no), units of alcohol consumed per week, smoking status (never, past, current), benign breast disease (yes/no), mother or sister with breast cancer (yes/no), BMI, age at menarche (years), age at first birth (years), parity (number of children), breastfeeding (yes/no), physical activity (sessions per week), bilateral oophorectomy (yes/no), hysterectomy (yes/no), age at menopause and menopausal status. Benign breast disease (BBD) was determined from the question "Have you ever had a breast lump removed or any operations on your breast(s)?" Age at menopause was determined using the reported age at menopause if known, or the age at bilateral oophorectomy, as the removal of the ovaries would trigger an automatic surgical menopause. Time since menopause was derived from the age at menopause variable.

2.3.1 Deprivation

Deprivation at recruitment was determined using quintiles of the Townsend index.[124] This is a widely used measure of deprivation in the UK first described by Peter Townsend in 1988. It is a measure of material deprivation within a population derived from the area where the participant was recruited.

The index incorporates four variables: unemployment, non-car ownership, non-home ownership, and household overcrowding. MWS participants were allocated a score based on their region of recruitment.

2.3.2 Region of recruitment

There were nine regions from which a participant could reside. These corresponded to the areas covered by the cancer registries: Oxford, East Anglia, South West, Thames, West Midlands, Northern and Yorkshire, Trent, North West [Mersey], and North West [Manchester/Lancashire]. Scotland was a separate region but as MWS only has linked screening data for England, women in Scotland were excluded from the analyses in this thesis. Furthermore, the analysis of breast cancer risk factors in Chapter 4 did not adjust for region as MD data was only available for women screened in Oxford.

2.4 Outcome ascertainment

Detailed descriptions of the outcomes relevant to this thesis are given in the results chapters (Chapters 3 to 6). Brief descriptions are provided here.

2.4.1 Mammographic density data

As mentioned earlier, MWS was linked to MD data for women who were screened in Oxford. MD data was provided by the Medical Physics Department at the Royal Surrey County Hospital, Guildford, UK. Automatic software (Volpara Health Technologies Ltd: Version 1.5.1, New Zealand) [125] was applied to a subset of digital mammograms, in women who were not part of the MWS, to determine their MD. The group in Surrey then used machine learning methods to develop an algorithm for determining categorical MD using the subset of Volpara measurements as the ground truths. This algorithm was then used to estimate volumetric breast density from which the BIRADS score (A-D) could be determined for all digital mammographic images available within the MWS. This method had been previously validated by determining the correlation between the ground truths from Volpara and the prediction with deep learning algorithms (Pearson correlation coefficient (r) = 0.96).[126] In my thesis, BIRADS was dichotomised to give low MD (A/B) and high MD (C/D). This was so the results could be compared to the literature as the majority of previous studies dichotomised MD. As there were fewer

than 10,000 women with MD data, BIRADS needed to be dichotomised in order to provide enough power in the individual categories.

2.4.2 Breast cancer

Participants were followed up for incident invasive breast (International Classification of Diseases 10th Edition code C50) through linkage to the NHS England cancer registries. The ICD-10 code refers purely to malignant neoplasms of breast tissue and not carcinomas in situ (D05) nor malignancies of the skin of the breast (C43.5 or C44.5). Cases were defined as women who developed invasive breast cancer after the date of recruitment and before the end date of follow-up.

Invasive breast cancer was further classified by mode of detection (screen-detected and interval breast cancers) and grade (low and high breast cancers). Mode of detection is used in Chapters 5 and 6. Grade, which shall be used in Chapter 6, was classified as: grade 1, well differentiated; grade 2, moderately differentiated; and grade 3, poorly differentiated. Grade 1 suggests the cells are slower growing and look more like normal breast cells. Grade 3 suggests the cells looks very different from normal breast cells and usually grow and spread faster. Grade 2 suggests the cells appear and grow at a pace in between grades 1 and 3. Information on other tumour characteristics was not sufficiently complete to allow classification by other tumour characteristics such as stage, size, lymph node status and hormone receptor status.

2.5 Statistical analysis

Detailed methods are given in the individual results chapters (Chapters 3-6). Some common features of the statistical analyses are outlined here.

Multivariable regression models were used for the majority of the analyses in Chapters 4 to 6. Logistic regression was used in Chapter 4 to examine the associations between breast cancer risk factors and MD. Cox regression and logistic regression were used in Chapter 5 to determine the associations between MHT use and interval and screen-detected breast cancer, as a proxy for mammographic screening sensitivity. Conditional logistic regression was used in Chapter 6 to generate risk prediction

models for breast cancer and breast cancer subtypes, as this was more computationally manageable than a cohort study. Continuous variables, such as BMI or age at menarche, were grouped into ordinal categories.

2.5.1 Exclusions

Detailed exclusions are given in the individual chapters as they were analysis dependent. Consistent exclusions were:

- Women living in Scotland, as screening and MD data were not available for these women (N=117,168).
- Women who returned their recruitment questionnaire after receiving the results of their screen were excluded from my analyses. This was because those who were diagnosed with breast cancer may have been more likely to recall certain risk factors, especially those thought to be associated with breast cancer, than women who were not diagnosed with breast cancer, thus introducing an element of recall bias (N= 64,817).

2.5.2 Missing data

For most of the exposure or adjustment variables used in the thesis, approximately 5% or less of the variable was missing. In Chapters 4 and 5, "missing" was included as a separate category if the variable was included as an adjustment variable in the regression models. However, age at menopause, and therefore time since menopause, was missing for 47% of women in the MWS, partly due to the substantial proportion of women who had a hysterectomy before natural menopause (16%). This is because following a hysterectomy, women immediately cease menstruating and are not able to determine the point at which they experience natural menopause. As age at menopause was an important risk factor for breast cancer, it was crudely adjusted for by including women who had a hysterectomy without a bilateral oophorectomy as a separate category. Sensitivity analyses were completed to assess the effect of crudely adjusting for age at menopause (Chapter 5) or time since menopause (Chapter 4). In Chapter 6, multiple imputation was used to handle missing data. This was to reduce bias in model development due to missing data and to improve the power of the prediction

models to reliably estimate coefficients. Further details are given in the methods section of Chapter 6.

Analyses were conducted in Stata versions 17 and 18 (Stata Corp, LP, College Station, TX) [127,128] and figures were created in Stata and R, using the Jasper package in R-Studio. [129]

2.6 Conclusions

As described in this chapter, this thesis will utilise various data collected over the past 20-30 years. The linkage of data across databases, platforms and registries has aided in collecting such information. This will allow me to examine the role of MD and other breast cancer risk factors on the development and detection of breast cancer in this thesis. Over the following four chapters, I shall present the results of the thesis, starting with a systematic review and meta-analysis of breast cancer risk factors and mammographic density.

Chapter 3. Breast cancer risk factors and mammographic density: Systematic review and meta-analysis

3.1 Introduction

In this chapter, I will explore the first part of the pathway shown in Figure 1 in Chapter 1 that is, the association between breast cancer risk factors and MD, by looking at the existing evidence.

Given that increased MD reduces breast screening sensitivity, understanding which factors are associated with MD may also identify women who would benefit from more frequent screens or more sensitive imaging, to avoid delaying diagnosis. As some factors may be modifiable, implementing lifestyle modifications may reduce breast cancer risk or improve detection. Understanding the determinants can also help elucidate the pathophysiological mechanisms of breast cancer development, as this can highlight the breast cancer risk factors for which MD may be on the causal pathway.

As discussed in Chapter 1, several risk factors for breast cancer have been thought to be associated with MD. Of these risk factors, menopausal status, age, MHT use and BMI are among the most established, with consistent evidence for these associations and clear biological plausibility explaining the associations. Therefore, I will not be examining these risk factors in this chapter.

Numerous studies have examined the association of MD with other reproductive and lifestyle factors, including parity, age at first birth, age at menarche, alcohol and smoking. However, the results are inconsistent, perhaps due to small sample sizes. Therefore, I conducted a systematic review on the associations of these risk factors with MD. Synthesising the results from all available literature will provide higher statistical power to draw more robust conclusions on these associations. Furthermore, pooling the results may allow for more detailed subgroup analyses, for example, exploring these associations by menopausal status or MHT use.

3.2 Aims

This systematic review aimed to provide a synthesis of the current published evidence on the associations of alcohol, smoking, parity, age at first birth and age at menarche, with MD, and where appropriate, conduct a meta-analysis.

3.3 Methods

3.3.1 Literature search

A systematic search of PubMed, EMBASE and Web of Science, was conducted to find published articles examining the associations of alcohol, smoking, parity, age at first birth and age at menarche, with MD, using specific search terms (Table 8). The exact search phrases used are shown in Appendix 1.

Table 8. Search terms used in electronic databases

Breast cancer risk factor/Outcome	Search terms used
Mammographic density	Breast density Mammographic density
Alcohol	Alcohol Lifestyle
Smoking	Smoking Tobacco Lifestyle
Parity	Parity Number of births Number of children Reproductive
Age at first birth	Age at first birth Age at first pregnancy
Age at menarche	Menstrual Menarche

3.3.2 Eligibility criteria

Studies were selected according to the inclusion criteria (Table 9). As methods of MD measurement have changed and improved over time, publication dates were restricted from 2000 to 2021. Studies reporting percentage MD were included, however studies only reporting dense or non-dense area or volume were excluded. This was because percentage MD was the most widely reported continuous measure. Therefore, to remain consistent and to capture as many studies as possible, percentage MD was used. Studies were excluded if they reported results based on transformed percentage MD, as

they could not be synthesised in the meta-analysis, with studies using non-transformed MD. The results were excluded if a study back-transformed their results to the original percentage MD scale. This was because the mean from a back-transformed estimate would be a geometric mean and not an arithmetic mean. The geometric mean is smaller than the arithmetic mean from an estimate that was not transformed. Furthermore, the CI is not symmetrical about the geometric mean.[130] Therefore, these studies could not be synthesised. Studies using ordinal logistic regression were also excluded as the results could not be pooled with the estimates from binary logistic regression. They were not analyses separately in order to make the number of analyses more feasible within the systematic review. As MD is strongly inversely associated with age, studies not reporting age-adjusted estimates were excluded.[56,131] To make the results more generalisable to the adult female population, studies exclusively investigating breast cancer patients were excluded.

No restrictions were placed on location. Reference lists of studies meeting the inclusion criteria and relevant reviews were screened for additional studies.

Table 9. Study inclusion and exclusion criteria

Inclusion criteria	**Exclusion criteria**
- Studies looking at smoking, alcohol, parity, age at first birth or age at menarche as exposures, and mammographic density as an outcome - Original research in peer-reviewed journal - Full text available online in English - Publication date 2000-2021 - Age-adjusted estimates reported - Mammographic density reported as percentage density or categorical measures of density (including BIRADS score, Boyd scale, Tabar scale or Wolfe scale), which were then dichotomised	- Reviews or narratives - Studies exclusively on breast cancer patients - Mammographic density only reported as dense area/volume or non-dense area/volume - Results based on transformed percentage MD or back-transformed MD - Ordinal logistic regression (instead of binary logistic regression)

3.3.3 Study selection and screening

Search results were imported into Covidence [132], an online platform for systematic reviews, which allows abstracts to be screened, full-texts to be reviewed and finally, data to be extracted. Covidence automatically removes duplicate articles. Two independent reviewers screened the abstracts and reviewed the full-texts. Discrepancies were discussed and resolved between reviewers. Any discrepancies which could not be resolved were discussed with my supervisors for the final verdict. Reasons for exclusion were recorded.[133]

3.3.4 Data extraction and harmonisation

A data extraction template, specific to this review, was created in Covidence. The specific data extracted are presented in Table 10. One reviewer extracted data from all studies. Due to the large numbers of studies involved, it was only feasible for the second reviewer to extract data from 40% of studies, but the extracted data from the remaining 60% of the studies, were cross-checked by the second reviewer. Discrepancies were discussed and resolved among reviewers. Any discrepancies which could not be resolved were discussed with my supervisors for the final verdict.

Table 10. Summary of data extracted from papers

Study information	Data extracted
Study characteristics	Authors Year of publication Name of study Type of study Number of participants contributing to estimates of associations
Study population	Age distribution (mean, median, range) Menopausal status Use of menopausal hormone therapy
Breast cancer risk factors	Definition/categorisation of alcohol consumption, smoking, parity, age at first birth or age at menarche
Outcome	Definition of MD (percentage or categorical measure) Method of assessment of MD Any transformation of percentage density
Measures of Association	Most adjusted measure of association reported 95% CI/standard error
Confounders	All variables adjusted for in the analysis

The measure of association that was extracted for a given study, and the need to harmonise the data by estimating an additional measure of association, depended on the way in which the risk factor and MD were classified.

Table 11 presents the measures of associations that were extracted for different combinations of risk factor classifications (binary, ordinal, continuous) and of MD classifications (categorical, continuous).

Table 11 also presents the additional measures of association that were estimated. For studies using a binary risk factor classification and reporting mean MD, I estimated a difference in the means. For studies using an ordinal risk factor classification, I fitted a trend through the extracted measures of association. The method used to fit trends depended on the classification of MD. For studies reporting mean MD, I used weighted least squares to fit trends. [134] In studies reporting ORs of high vs low MD, I fitted trends using generalised least squares as described by Greenland and Longnecker. [135] This was because the ORs were correlated as they had a common reference group. For studies reporting differences in mean MD, I fitted trends using an extension to the method of Greenland and Longnecker which accounted for the correlation between the differences in means.[136] This extension was needed to account for the correlation between the differences in means, as the reported differences were not independent of each other. Essentially, this method adjusts for the fact that comparisons share a common baseline, ensuring more accurate trend estimation.

Additional methods of data harmonisation included converting estimated trends for alcohol to increments of 10 g/day, smoking 10 cigarettes/day, and where a study reported separate results for pre- and postmenopausal women, calculating a weighted average to produce an overall estimate.

3.3.5 Data synthesis

Meta-analyses were conducted for the following four measures of association:

- difference in mean MD between binary risk factor categories
- difference in MD per unit or given increment increase (e.g. 10g/day) in risk factor
- OR for high vs low MD associated with high vs low risk factor status

- OR for high vs low MD per unit increase in risk factor

Pooled estimates were calculated under a fixed-effects model using the method of empirically weighted least squares, where the weights were the inverse of the study variance.[137] Heterogeneity was assessed using I^2 tests. [138,139] Analyses were conducted in Stata 17.[127] Forest plots were generated in R using the Jasper package.[129]

3.3.6 Sources of heterogeneity

Subgroup analyses by menopausal status, MHT use, levels of adjustment, location of study, and method of MD assessment were conducted to investigate sources of heterogeneity, where possible. For adjustments, categories included age-adjusted only, age + BMI, and age + BMI + MHT + menopausal status. The location categories were Africa/Asia/Oceania, Europe, and the Americas. The methods of MD assessment were computer-assisted, computer-automated, and visually assessed. Chi-squared test was used to determine any evidence of statistical heterogeneity between pre- and postmenopausal women.

3.3.7 Publication bias and critical appraisal

Publication bias was evaluated visually, using a funnel plot, and quantitatively, using Egger's test, where there were at least ten studies in the meta-analysis.[138,139] Funnel plots were generated in Stata 17. Two reviewers assessed study quality using appraisal tools from the Joanna Briggs Institute. These are used to evaluate the "trustworthiness, relevance and results of published papers".[140] As part of the appraisal process, reviewers assessed whether studies adjusted for or were stratified by important confounders such as BMI, menopausal status and MHT.

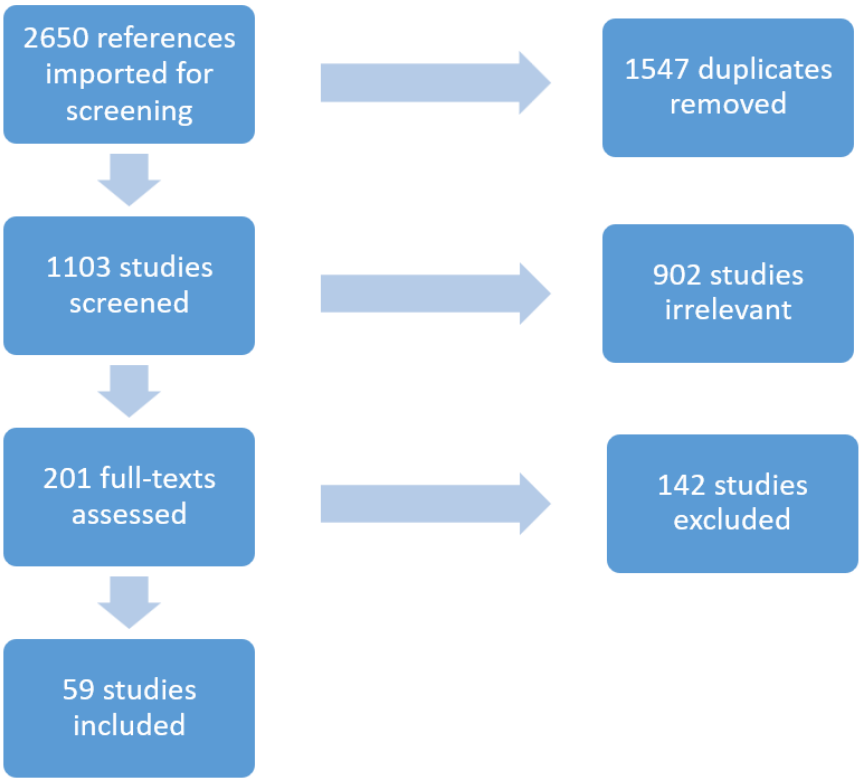
Table 11. Data extraction/harmonisation for measures of association by breast cancer risk factor classification and mammographic density classification

Risk factor classification/ MD classification	Measures of association extracted	Additional data extracted	Additional measure of association estimated	Measure of association used in meta-analysis
Risk factor: binary **MD: continuous**	Difference in mean MD between binary risk factor categories (and 95% CI) Means for MD by binary risk factor status (and 95% CI)	N/A N/A	N/A Difference in mean MD between binary risk factor categories	Difference in mean MD between binary risk factor categories Difference in mean MD between binary risk factor categories
Risk factor: binary **MD: categorical (high/low)**	OR for high vs low MD in high vs low risk factor status (and 95% CI)	N/A	N/A	OR for high vs low MD in high vs low risk factor status
Risk factor: ordinal **MD: continuous**	Means for MD by risk factor categories (and 95% CI) Difference in MD by risk factor categories (and 95% CI)	N/A No. of subjects by risk factor category	Difference in MD/unit increase in risk factor estimated by weighted least squares Difference in MD/unit increase in risk factor estimated by generalised least squares	Difference in MD/unit increase in risk factor Difference in MD/unit increase in risk factor
Risk factor: ordinal **MD: categorical (high/low)**	ORs for high vs low MD by risk factor categories (and 95% CI)	No. of high and low MD subjects by risk factor category	Relative difference in OR for high vs low MD /unit increase in risk factor estimated by generalised least squares	OR for high vs low MD /unit increase in risk factor
Risk factor: continuous **MD: categorical (high/low)**	Relative difference in OR for high vs low MD/unit increase in risk factor (and 95% CI)	N/A	N/A	OR for high vs low MD/unit increase in risk factor
Risk factor: continuous **MD: continuous**	Difference in MD/unit increase in risk factor (and 95% CI)	N/A	N/A	Difference in MD/unit increase in risk factor

3.4 Results

The literature search yielded 2650 studies, with 1544 duplicates which were automatically removed by Covidence. Following the abstract and full-text review, 59 studies remained (Figure 6). A summary of studies is presented in Appendix 2. Twenty-six studies were from North America, 18 from Europe, 11 from Asia, one from Africa, one from Oceania and one from South America. One study was conducted in both North America and Asia.[141] Study population sizes ranged from 102 to 7,265,584 women. Six studies were in postmenopausal populations, seven had premenopausal populations, and 46 were in all women.

Figure 6. PRISMA flowchart summarising the study screening process in the systematic review



3.4.1 Alcohol

Twenty studies examined the association between alcohol consumption and MD. The direction of the pooled results consistently showed alcohol was associated with higher MD (Figure 7).

Drinkers had a 15% increase in the odds of having higher MD, than non-drinkers: pooled OR 1.15 (95% CI 1.06, 1.25). Each 10 g/day increase in alcohol was associated with a 6% increase in the odds of having higher MD: 1.06 (1.02, 1.10).

The pooled estimate of the difference in mean MD among drinkers compared with non-drinkers was 1.71% (95% CI 0.29, 3.13). Each 10 g/day increase in alcohol was associated with a 0.03% increase in the pooled estimate of the difference in mean MD, although not significant (-0.01, 0.06). McDonald et al. was the only study to report a significant inverse association between alcohol and MD: -1.48% (-1.59, -1.37).[142]

There was significant heterogeneity among studies examining alcohol based on the I^2 values across the meta-analyses. This was further evidenced by the varying magnitude and direction of the results.

3.4.2 Smoking

Twenty-one studies examined the association between smoking tobacco and MD (Figure 8). There was no clear evidence of an effect in studies examining ever smoking on MD. However, amongst other classifications, smoking was associated with lower MD. Each 10 cigarettes/day increase was associated with a decrease of 7% in the odds of having high MD: pooled OR 0.93 (0.90, 0.96), and with a decrease of 0.30% mean percentage MD (-0.32, -0.29). The pooled difference in mean percentage MD among current smokers was 1.09% (-1.74, -0.43) less than never smokers, and among past smokers was 0.77% (-1.35, -0.20) less than never smokers.

There was again significant heterogeneity among some of the meta-analyses for smoking (current vs never smokers: p=0.024).

3.4.3 Parity

Thirty-seven studies examined the association between parity and MD (Figure 9). Consistently, across pooled results, parity was associated with lower MD. Parous women had a 32% reduction in the odds of having high MD compared to nulliparous women: pooled OR 0.68 (0.65, 0.71). Each birth was associated with a 12% decrease in the odds of having higher MD: 0.88 (0.875, 0.884).

The pooled mean percentage MD among parous women was 1.25% lower than nulliparous women (-2.29, -0.21). Each birth was associated with a 0.77% decrease in the pooled difference in mean percentage MD (-0.80, -0.74).

There was statistical heterogeneity among the studies. Although the direction of association was more consistent than the lifestyle factors, the magnitude differed.

3.4.4 Age at first birth

Twenty-four studies examined the association between age at first birth and MD (Figure 10). The pooled estimates suggested that a later age at birth was associated with higher MD. An increase of one year in age at first birth was associated with a 4% increase in the odds of having high MD: 1.04 (1.04, 1.05), and with a 0.14% increase in the pooled difference in mean percentage MD (0.13, 0.14).

There was evidence of heterogeneity where there were a large number of studies in the meta-analysis. Visually, whilst there was largely consistency in the direction of the results, the magnitude differed.

3.4.5 Age at menarche

Twenty-nine studies examined the association between age at menarche and MD (Figure 11). The pooled effect was more consistent in the meta-analyses with more than two studies. An increase of one year in age at menarche was associated with a 4% increase in the odds of having higher MD: 1.04 (1.03, 1.04), and with a 0.22% increase in the pooled difference in mean percentage MD (0.21, 0.24).

There was less evidence of heterogeneity in the meta-analyses for age at menarche, except for the difference in mean percentage MD per year increase in age at menarche. Studies did, however, differ in terms of magnitude and direction, and many showed a null association.

A summary of the pooled results is given in Table 12.

3.4.6 Sources of heterogeneity

Results by menopausal status (forest plots) are given in Appendix 3. In addition to the studies exclusively in pre- or postmenopausal women, six studies reported findings by menopausal status for alcohol and MD, three for smoking, 13 for parity, six for age at first birth, and nine for age at menarche.

There was statistical evidence that results differed by menopausal status for alcohol, smoking, parity and age at menarche. Generally, the effects of the risk factors on MD were greater in premenopausal women than in postmenopausal women. The OR for high vs low MD for drinkers vs non-drinkers was 1.78 (1.30, 1.24) for premenopausal women and 1.13 (0.92, 1.38) for postmenopausal women (p-value for heterogeneity=0.017). The difference in mean percentage MD for current vs never smokers was -6.80% (-11.0, -2.58) for premenopausal women -1.15% (-2.21, -0.09) for postmenopausal women (p=0.011). The effect of each birth on the OR for high vs low MD was greater in postmenopausal women than premenopausal women: OR 0.88 (0.88, 0.88) vs 0.90 (0.88, 0.92), respectively (p=0.048). However, the effect of each birth on the difference in mean percentage MD was greater in premenopausal women than postmenopausal women: -0.78% (-0.82, -0.75) vs -0.55% (-0.67, -0.43), respectively (p<0.001). The OR for high vs low MD per year increase in age at menarche was 1.05 (1.04, 1.06) for premenopausal women and 1.03 (1.02, 1.04) for postmenopausal women (p=0.006).

Only four studies reported results by MHT status, so I did not stratify by this. This included three examining alcohol[143–145] and three examining smoking.[143,145,146] One study examined parity, age at first birth and menarche, by MHT use.[145] Jacobson et al. reported no evidence of effect modification by ever MHT for the effects of alcohol and smoking on MD.[144,146] Brand et al. (2013) reported the effect of alcohol on MD was seen in current MHT users but not in never users. This was not seen for smoking.[143] Yaghjyan et al. (2012) reported that the effects of alcohol and parity on MD were greater in ever MHT users than in never users.[145]

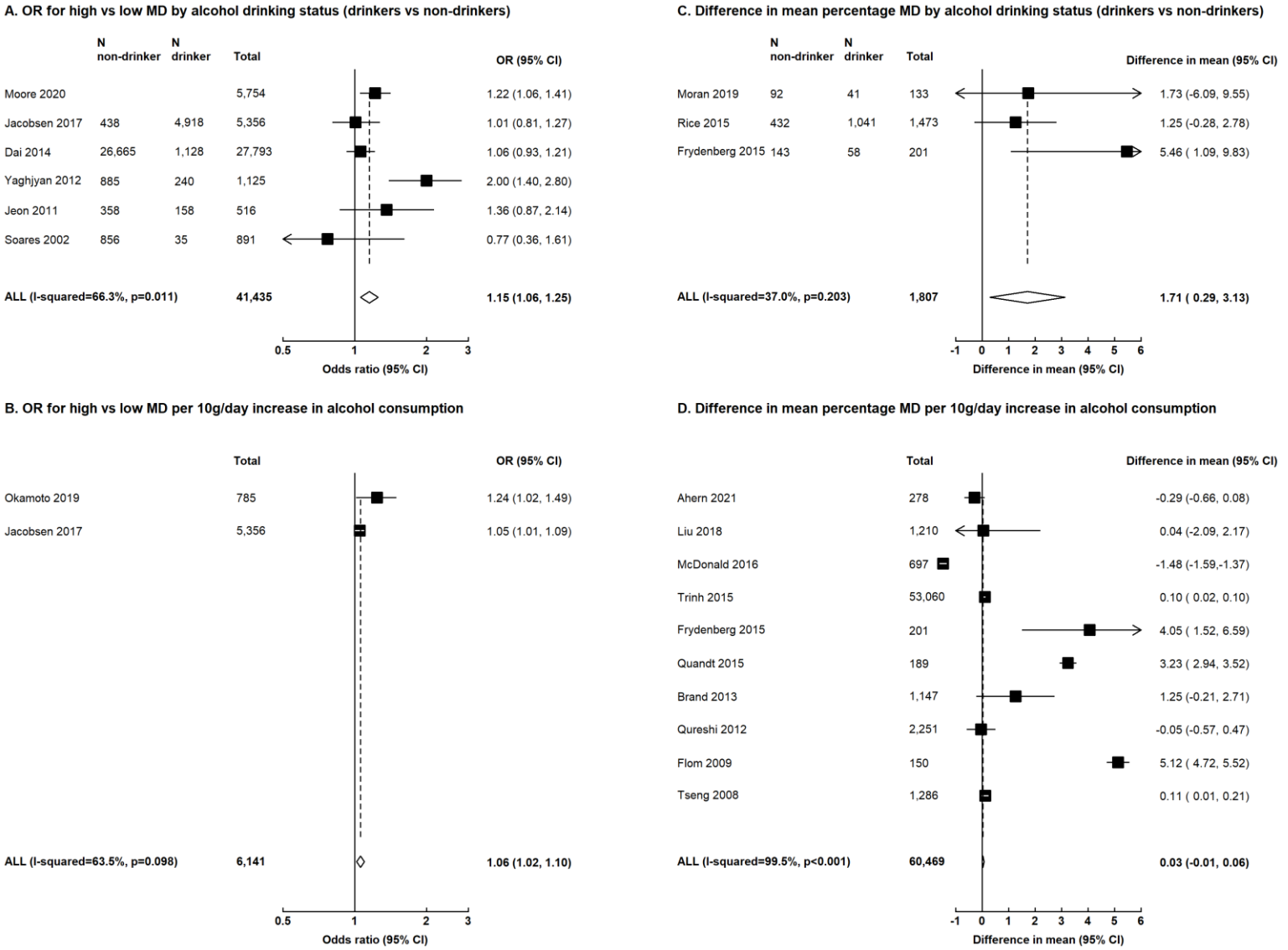
The effects of adjustments, location, and method of MD assessment are shown in Appendix 4. Five studies did not adjust for BMI, nine studies did not adjust for menopausal status, and 15 studies did not adjust for MHT. For measuring MD, 32 used a visual assessment method, 21 studies used computer-assisted methods, and 6 used computer-automated technology. Among studies using visually assessed or computer-assisted methods to measure MD, 19 studies used a single reader to interpret the mammogram. This could contribute to misclassification of outcome. The remaining

studies had two readers interpreting mammograms. Given the relatively few studies within each category (levels of adjustment, location, and assessment of MD), it was unclear whether these factors explained the heterogeneity.

3.4.7 Publication bias and critical appraisal

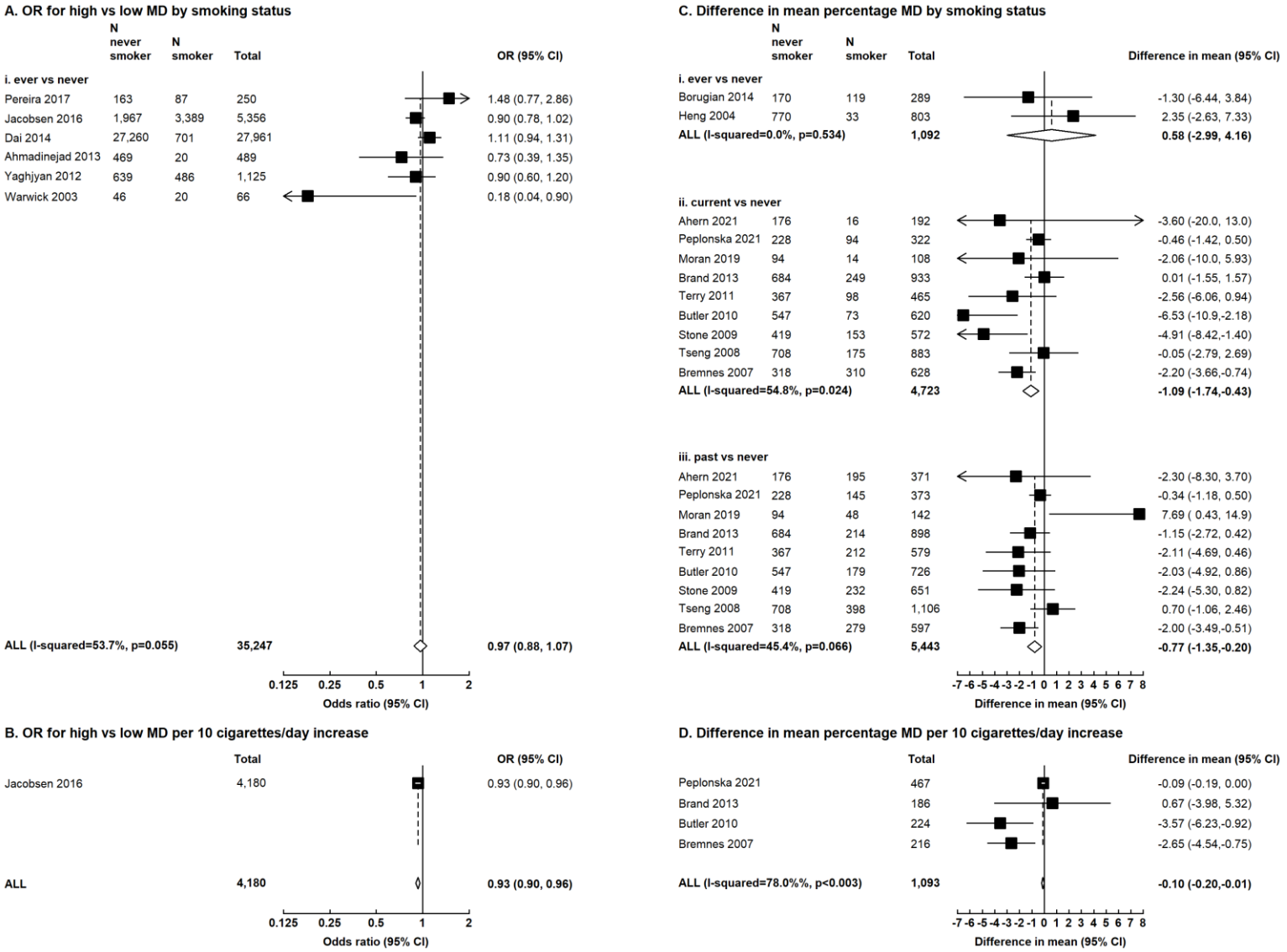
Funnel plots are given in Appendix 5. There was a suggestion of publication bias, as shown by the asymmetrical funnel plots, among studies examining alcohol, parity, and age at first birth. There was no obvious evidence of funnel plot asymmetry among studies examining age at menarche. The Egger's test for funnel plot asymmetry only revealed statistical evidence of publication bias for parity, and not for any other risk factor. Publication bias for smoking could not be tested as there were fewer than 10 studies in each of the meta-analyses. No studies were removed based on the critical appraisal tools from the Joanna Briggs Institute, as they were deemed suitable for inclusion unless they did not meet the inclusion criteria specified in the methods.

Figure 7. The relationship of alcohol intake with continuous and categorical measures of mammographic density



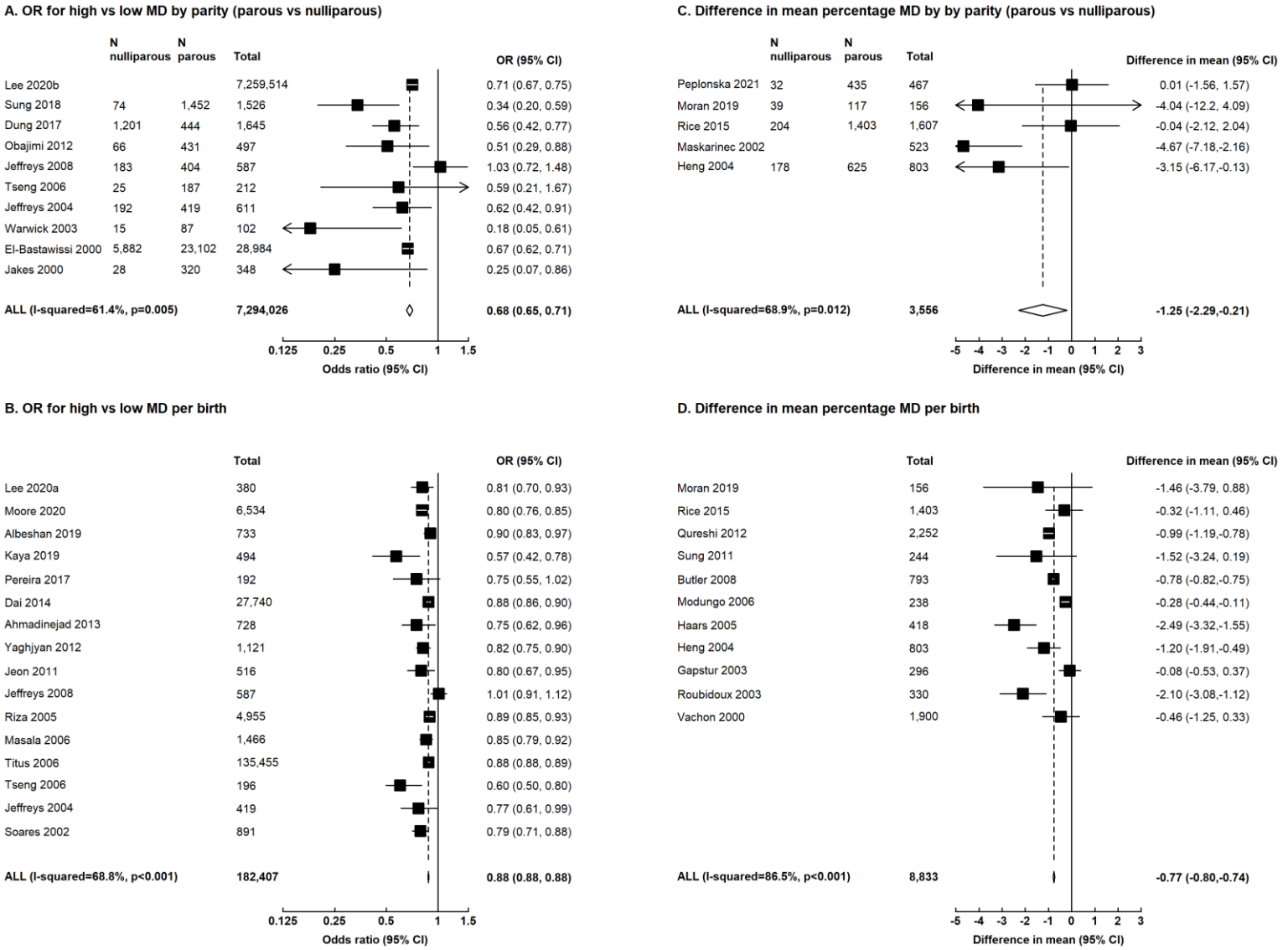
This figure shows four fixed-effects meta-analyses investigating the relationship between alcohol use and mammographic density for the studies included in the systematic review: Plot A shows the OR for having high vs low mammographic density comparing ever drinkers vs non-drinkers; Plot B shows the OR for high vs low MD per increment of 10g of alcohol/day; Plot Ci shows the difference in mean percentage MD comparing ever drinkers vs non-drinkers; Plot D shows difference in mean percentage MD per increment of 10g of alcohol /day.

Figure 8. The relationship of smoking with continuous and categorical measures of mammographic density



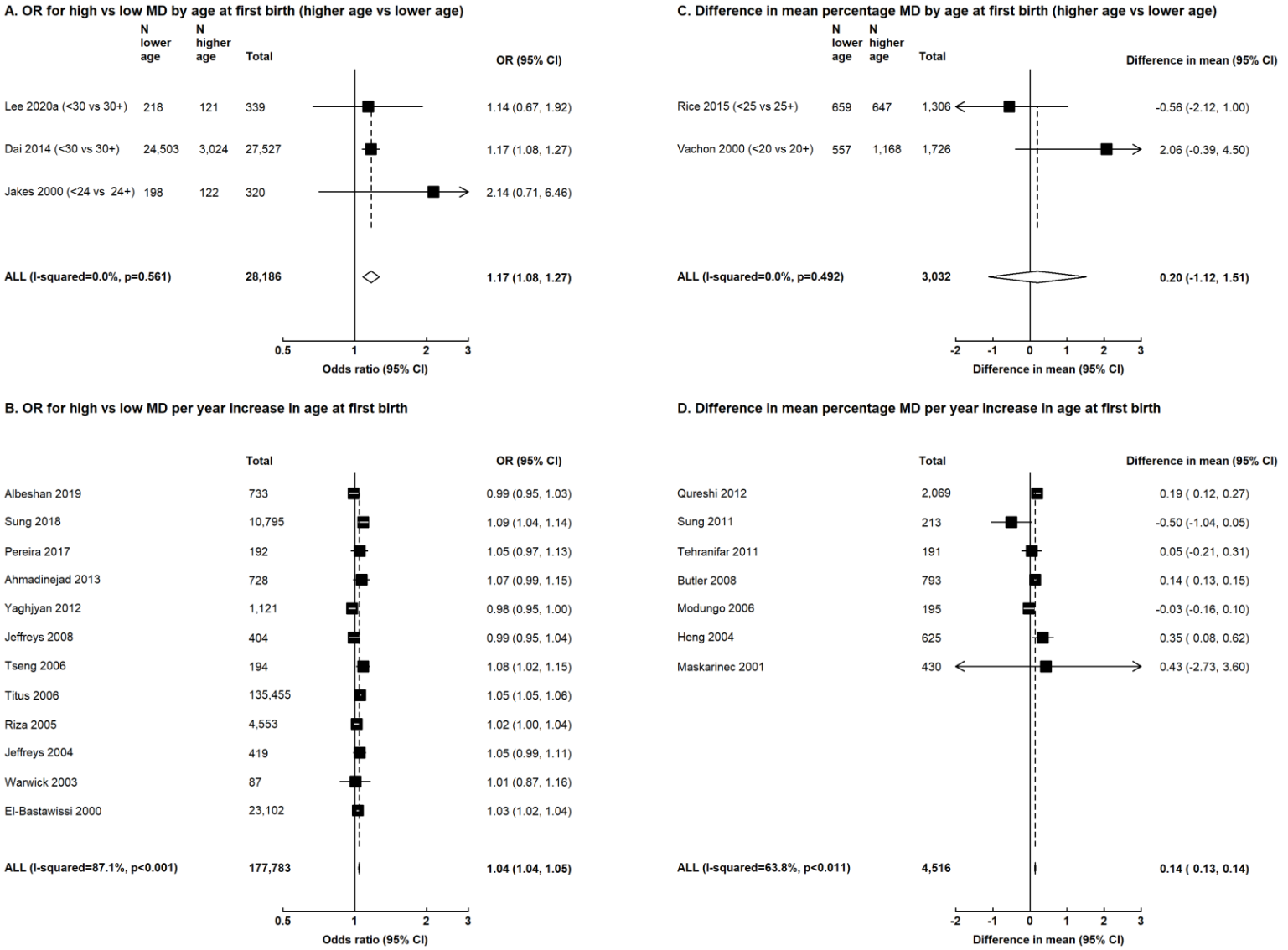
This figure shows six fixed-effects meta-analyses investigating the relationship between smoking and mammographic density for the studies included in the systematic review: Plot A shows the OR for having high vs low mammographic density comparing ever smokers vs never smokers; Plot B shows the OR for high vs low MD per increment of 10 cigarettes/day; Plot Ci shows the difference in mean percentage MD comparing ever vs never smokers; Plot Cii shows the difference in mean percentage MD comparing current vs never smokers; Plot Ciii shows the difference in mean percentage MD comparing past vs never smokers; Plot D shows difference in mean percentage MD per increment of 10 cigarettes/day.

Figure 9. The relationship of parity with continuous and categorical measures of mammographic density



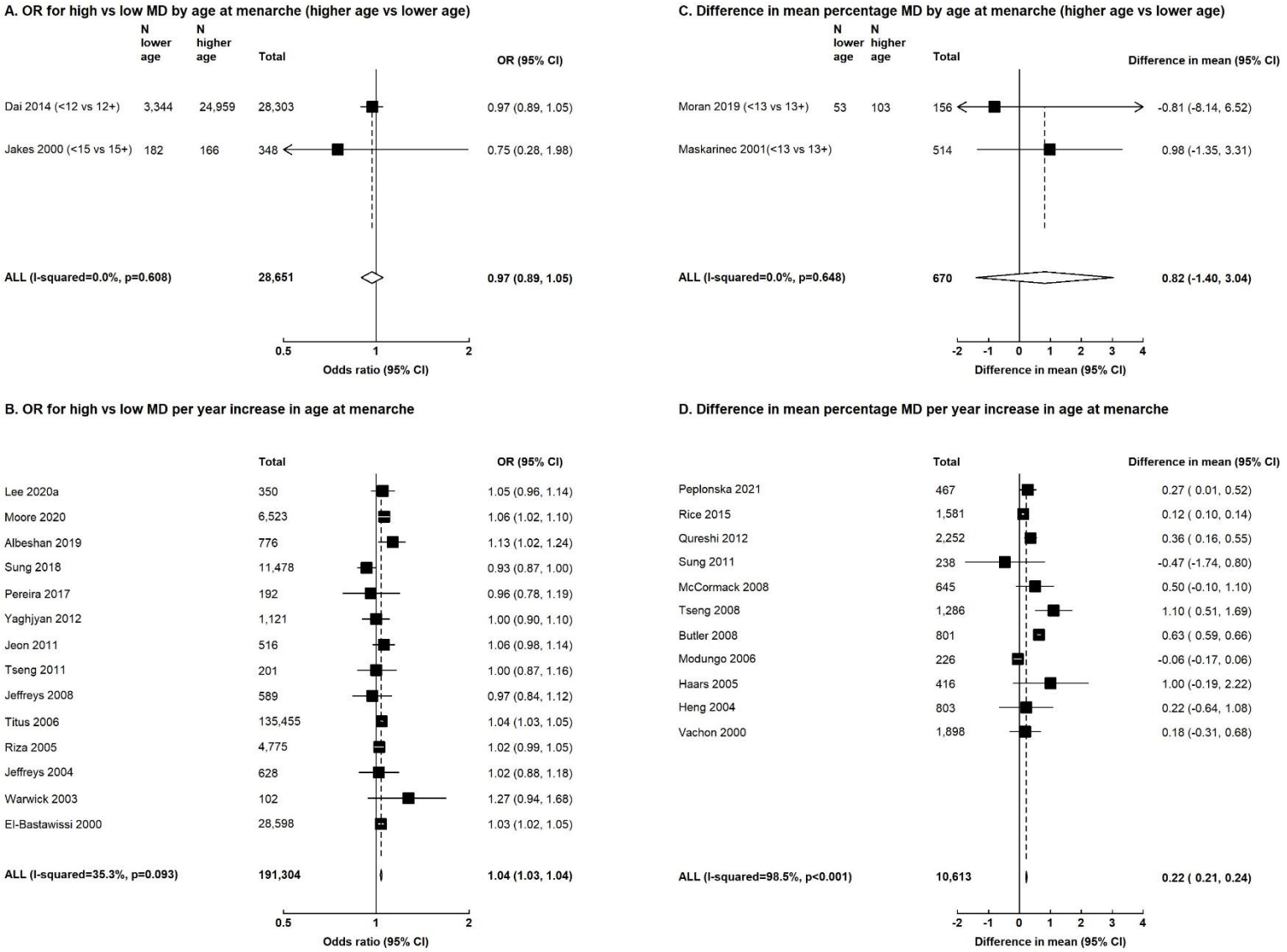
This figure shows four fixed-effects meta-analyses investigating the relationship between parity and mammographic density for the studies included in the systematic review: Plot A shows the OR for having high vs low mammographic density comparing parous vs nulliparous women; Plot B shows the OR for high vs low MD per birth; Plot C shows the difference in mean percentage MD comparing parous vs nulliparous women; Plot D shows difference in mean percentage MD per birth.

Figure 10. The relationship of age at first birth with continuous and categorical measures of mammographic density



This figure shows four fixed-effects meta-analyses investigating the relationship between age at first birth and mammographic density for the studies included in the systematic review: Plot A shows the OR for having high vs low mammographic density comparing higher vs lower age at first birth; Plot B shows the OR for high vs low MD per year increase in age at first birth; Plot C shows the difference in mean percentage MD comparing higher vs lower age at first birth; Plot D shows difference in mean percentage MD per year increase in age at first birth.

Figure 11. The relationship of age at menarche with continuous and categorical measures of mammographic density



This figure shows four fixed-effect meta-analyses investigating the relationship between age at menarche and mammographic density of the studies included in the systematic review: Plot A shows the OR for having high vs low mammographic density comparing higher vs lower age at menarche; Plot B shows the OR for high vs low MD per year increase in age at menarche; Plot C shows the difference in mean percentage MD comparing higher vs lower age at menarche; Plot D shows difference in mean percentage MD per year increase in age at menarche.

Table 12. Pooled meta-analyses estimates of the associations between breast cancer risk factors and mammographic density

Risk factor	N	OR (95% CI)	I^2 (p-value)	Risk factor	N	Diff in mean (95% CI)	I^2 (p-value)
Alcohol							
A. Drinkers vs non-drinkers	6	1.15 (1.06, 1.25)	66.3 (0.011)	**C. Drinkers vs non-drinkers**	3	1.71 (0.29, 3.13)	37.0 (0.20)
B. per 10 g/day	2	1.06 (1.02, 1.10)	63.5 (0.098)	**D. per 10 g/day**	10	0.03 (-0.01, 0.06)	99.5 (<0.001)
Smoking							
A. Ever vs never	6	0.97 (0.88, 1.07)	53.7 (0.055)	**Ci. Ever vs never**	2	0.58 (-2.99, 4.16)	0.0 (0.53)
B. per 10 cigarettes/day	1	0.93 (0.90, 0.96)		**Cii. Current vs never**	9	-1.09 (-1.74, -0.43)	54.8 (0.024)
				Ciii. Past vs never	9	-0.77 (-1.35, -0.02)	45.4 (0.066)
				D. per 10 cigarettes/day	4	-0.10 (-0.20, -0.01)	78.0 (<0.001)
Parity							
A. Parous vs nulliparous	10	0.68 (0.68, 0.71)	61.4 (0.005)	**C. Parous vs nulliparous**	5	-1.25 (-2.29, -0.21)	68.9 (0.12)
B. per birth	16	0.88 (0.88, 0.88)	68.8 (<0.001)	**D. per birth**	11	-0.77 (-0.80, -0.74)	86.5 (<0.001)
Age at first birth							
A. Higher vs lower age	3	1.17 (1.08, 1.27)	0.0 (0.56)	**C. Higher vs lower age**	2	0.20 (-1.12, 1.51)	0.0 (0.48)
B. per 1 year increase	12	1.04 (1.04, 1.05)	87.1 (<0.001)	**D. per 1 year increase**	7	0.14 (0.13, 0.14)	63.8 (<0.001)
Age at menarche							
A. Higher vs lower age	2	0.97 (0.89, 1.05)	0.0 (0.61)	**C. Higher vs lower age**	2	0.82 (-1.40, 3.04)	0.0 (0.65)
B. per 1 year increase	14	1.04 (1.03, 1.04)	35.3 (0.093)	**D. per 1 year increase**	11	0.22 (0.21, 0.24)	98.5 (<0.001)

3.5 Discussion

This review aimed to consolidate the evidence surrounding the associations between MD and various lifestyle and reproductive breast cancer risk factors. BMI, age, menopausal status, and MHT use were not included, as the effects of these on MD were already well-established and accepted. This review suggested a possible positive association between alcohol, age at first birth and age at menarche with MD, and an inverse association of smoking and parity with MD.

There was a consistent positive association of alcohol with MD across the groupings of analysis. There are plausible mechanisms which may explain the pooled association between alcohol and MD. Alcohol promotes aromatase activity, which increases the synthesis of oestrogens, stimulating breast epithelial proliferation.[144,147] However, given the heterogeneity in the individual meta-analyses, as indicated by the I^2, pooled estimates are difficult to interpret. A systematic review and meta-analysis of alcohol and MD was published in 2017.[148] The authors included 20 studies published until 2015, of which 10 were included in this review. Like the meta-analysis in this chapter, they concluded that alcohol consumption may increase MD among studies reporting either OR for high versus low MD or continuous measures of MD. They reported a difference in mean percentage MD comparing the highest level of alcohol consumption with the lowest level, of 0.84% (95% CI 0.12, 1.56) and an OR of 1.81 (1.07, 3.04). Where an ordinal measure of alcohol was reported for percentage MD, the authors extracted the value associated with the highest category (versus the lowest category) of alcohol consumption rather than generating a trend as I did. They, therefore, did not make use of the full data within individual studies. As a result, I cannot compare the magnitude of their results to those in my meta-analysis.

Except for the analysis of ever smokers, smoking was associated with a possible reduction in MD. A possible mechanism for an inverse association, is the anti-oestrogenic effect of smoking, which inhibits breast epithelial tissue proliferation.[149,150] However, there was again evidence of heterogeneity in terms of magnitude and direction. The association seen in this review is interesting as the evidence between smoking and breast cancer risk has long been debated. The Generations Cohort of 102,927

women in the UK, found that ever smokers had a 14% increased risk of breast cancer, compared with never smokers (HR 1.14, 95% CI 1.03, 1.25).[151] The Collaborative Group on Hormonal Factors in Breast Cancer performed a meta-analysis of 53 studies and found no association between ever smoking and breast cancer (HR 1.03, 0.98, 1.07).[22] It is unclear whether these studies were conducted on women who attended screening. Perhaps the reduction in MD may unmask cancers among smokers at screening, improving the sensitivity of mammographic screening, and therefore, smokers may appear to have an increased risk of breast cancer at screening compared to never smokers.

Given that alcohol consumption is associated with smoking tobacco,[152] it is possible that the effects of these on MD may cancel each other out. For example, inadequate adjustment of smoking might lead to an underestimation of the positive effect of alcohol on MD. Conversely, inadequate adjustment of alcohol may lead to an underestimation of the inverse association of smoking with MD.

Parity was associated with a reduction in MD and increasing age at first birth was associated with higher MD, although small. An inverse association of parity with MD is consistent with what is understood about changes in breast lobular tissue during pregnancy, known as terminal differentiation. These changes reduce cell proliferation.[153] Therefore, the more children a woman has, the further the proliferative potential of her breast tissue declines. Similarly, by having children earlier, proliferative potential declines at an earlier age, leading to reduced MD.[154] Whilst there was evidence of heterogeneity, the direction of individual studies was more consistent than the lifestyle factors considered in this systematic review.

Higher age at menarche was associated with higher MD, although the association was small. This is contrary to the understanding that the later onset of menarche is associated with a lower risk of breast cancer.[155] This paradoxical association with MD is not fully understood but may be explained by the relationship between early-life adiposity and age at menarche. Higher early-life BMI is associated with earlier age at menarche. Since early-life BMI is correlated with later-life BMI, this could give rise to an indirect positive association between age at menarche and MD. A study published after my systematic

review by the International Consortium of Mammographic Density found that in over 10,000 women in 22 countries, later age at menarche was associated with higher MD.[156] They found that the magnitude of the association attenuated after adjusting for BMI. They also found that the estimates were greatest among women with average or below median BMI. Only three studies in my review of age at menarche did not adjust for BMI. It is important to note, though, that in my review, there was heterogeneity among the studies, and many reported small or null associations.

Due to the large degree of heterogeneity among studies, results need to be interpreted with caution. Sub-group analyses by menopausal status were performed, where possible, to explore a potential source of heterogeneity. There was some evidence that the effects of alcohol, smoking, parity and age at menarche differed by menopausal status, suggesting this may account for part of the heterogeneity across the studies. The effects of alcohol, smoking, and age at menarche were greater in premenopausal women than in postmenopausal women. This was less clear for parity as the results differed by classification of MD and it is not clear why this is. Menopausal status is a strong determinant of MD. Premenopausal women have denser breasts than postmenopausal women. This is because menopausal status reflects contrasting physiological states, such that premenopausal women have higher levels of circulating hormones, which lead to more epithelial proliferation within the breast than postmenopausal women. As a result, the effects of breast cancer risk factors on density are likely to differ by menopausal status. The results here suggest that being premenopausal has an additive effect on the associations seen. It would be prudent for future studies examining the associations of risk factors and MD to look separately by menopausal status.

MHT is composed of female reproductive hormones, oestrogen with or without progestogen, which stimulate epithelial proliferation in the breast. Only four studies reported their results by MHT status, and therefore, it was not possible to explore this as a potential source of heterogeneity. Two studies did suggest that the effects of alcohol were greater in women ever MHT users than in never MHT users and one study suggested the effects of parity on MD were also greater in women on MHT.[143,145]

Therefore, studies examining the associations of risk factors and MD, should restrict their populations to those not taking MHT, to ensure changes are not masked by the already higher MD in women in MHT or, conversely as suggested by two studies here, influenced by the additive effect of MHT on the relationships.

Level of adjustment, location, and method of MD assessment were also investigated as sources of heterogeneity. However, given the relatively few studies in each category, it was difficult to say with certainty whether there was real variation in associations according to these factors as the patterns were not clear.

The literature describing the association of MD on breast cancer risk, usually categorises percentage MD in 25% categories of MD, for example <5%, 5-24%, 25-49%, 50-74% and >75%. A previous meta-analysis reported that the relative risk of breast cancer comparing women with 5-24% MD to the lowest category (<5%) was 1.79 and >75% to the lowest category was 4.64.[1] Therefore, the changes of <2% seen in many of the meta-analyses are unlikely to represent a substantial change in breast cancer risk or affect screening sensitivity, at least not in isolation. Furthermore, the evidence suggests that the associations of age, menopausal status, BMI, and combined MHT use with MD, as detailed in Chapter 1, are greater than those for the risk factors studied in my systematic review. For example, premenopausal women have a 3.50% (0.62, 6.38) higher mean percentage MD compared to postmenopausal women,[57] and a 1kg/m^2 increase in BMI is associated with a 2.03% reduction in percentage density (-2.09, -1.98).[58]

The small associations seen may not be clinically or pathologically significant enough to influence breast cancer risk or affect screening sensitivity. Whilst the results here may represent the true associations, I would have expected larger effects, if MD truly mediated the associations between these risk factors and breast cancer risk. However it is possible that MD partially mediates the associations.

This systematic review does have its strengths. As individual studies were likely underpowered to show significant associations between the risk factors and MD, the pooled analysis provided greater power to test the associations. Furthermore, despite classifications of MD and risk factors differing substantially across studies, I was able to harmonise and synthesise many of the results so that more studies could be included to give my meta-analysis power to detect differences in MD.

The review does have some limitations. Several studies had to be excluded because they did not meet the inclusion criteria even though they examined age-adjusted associations between the risk factors of interest and MD. The main reasons for this were either because MD was transformed (usually log or square-root/cube-root transformed) or because ordinal logistic regression was used. These studies could not be pooled with the other studies in the meta-analysis. Seven studies had to be excluded as the results were transformed. Seven studies had to be excluded as they utilised ordinal logistic regression and, therefore, could not be included in the meta-analysis with binary logistic regression. The results of these studies largely agreed with the pooled results in the present review and are unlikely to have changed the conclusions of this chapter. McBride et al. (2020) conducted a study in 23,456 women attending screening in California, USA.[157] Alcohol consumption was positively associated with MD (P_{trend}=0.01). They also found that smoking was inversely associated with MD (P_{trend}=0.0008). Alexeeff et al. (2019) reported that parity was associated with reduced MD and later age of first birth or menarche was associated with higher MD among 24,840 women in California.[57] In a population of over 3000 screening women in Spain, there was a positive association with alcohol (P_{trend} =0.045) and a negative association with smoking (P_{trend} =0.017).[158] In the same population, there was no evidence of an association between age at menarche and MD, but there was an inverse association with parity (P_{trend}<0.001) and a positive association with age at first birth (P_{trend} =0.02).[154,159]

Observational studies are also subject to biases due to design and conduct. Selection bias is likely as some studies used screening cohorts (N=30). There may be systematic differences between those who attend screening, compared to those who do not. Furthermore, as reporting of risk factors relied on

questionnaires, they could be prone to misclassification. Women may underestimate smoking or alcohol habits, which would overestimate the association. Alcohol consumption was categorised based on current drinking habits rather than historical use. Hence, historical alcohol consumption amongst past drinkers could not be ascertained.

Further limitations include several sources of heterogeneity across the studies, such as population demographics and behaviours, as these are likely to affect MD measures and patterns of MD, respectively. Each study used different methods and studied different populations, making comparison challenging. Whilst most studies adjusted for BMI, some did not adjust for MHT use or menopausal status. This may have led to some residual confounding. MHT use and being premenopausal may have modified some of the associations with other risk factors, as suggested by a few studies. As 19 studies used a single reader to determine MD, there may have been a greater chance of MD misclassification than among studies with two readers. As mentioned, despite an attempt to investigate the sources of heterogeneity, the few studies within categories made it challenging to determine a pattern. However, menopausal status could account for some of the heterogeneity seen. Ideally in a systematic review a secondary reviewer would extract all the relevant data from the studies included in the review. However, given the sheer size of this review, it was not feasible for a second person to extract all the information. Therefore, a portion of the data was cross-checked rather than exacted. This may have introduced errors, likely non-differential. As that the review was completed at the start of my DPhil, it will need to be updated to include studies published since 2022. During this process, I will be able to ask a secondary review to extract data from all of the included studies.

To make use of all data, a trend was fit across categorical groups of exposures. Consequently, I made assumptions about the distributions within categories, which may have differed from what would have been seen if the authors had done the analysis themselves. Whilst the funnel plot asymmetry suggested publication bias, other possible causes of asymmetry include insufficient statistical power,

bias in small studies, artefacts, and chance. Egger's test, however, only showed evidence of publication bias for parity. Given the heterogeneity across studies and the relatively few numbers of studies per meta-analysis, any tests for funnel plots asymmetry may be underpowered and, therefore, need to be interpreted with caution. [139,160]

Despite limitations, this is a robust review of many studies, which have been synthesised to utilise all available data. The review has added to the body of evidence that lifestyle and reproductive factors are associated with MD. The outstanding questions following this review are whether these associations are seen in larger populations, restricted by menopausal status and MHT use, and where important confounders are controlled for.

The review justifies the need to examine these associations in a large study, like the MWS. In the MWS, I will be able to look at the associations, where analyses can be restricted to postmenopausal women and those not using MHT, to avoid any effect modification by them. Due to the population demographics, I will not be able to explore these associations in premenopausal women. Such an analysis will contribute to the understanding of how these factors influence breast cancer risk, and how they may impact screening sensitivity.

3.6 Conclusions

The results suggest there is an association between lifestyle and reproductive risk factors and MD. It is yet to be determined whether this corresponds to clinically relevant changes in breast cancer risk and the absolute impact on screening sensitivity. Given the small magnitudes of associations, it is not clear whether these will have clinical relevance. The associations seen for alcohol, parity and age at first birth are likely to have plausible explanations. Those seen for smoking and age at menarche are less clear. An analysis in the MWS (Chapter 4) will allow me to address some of the limitations in this review. I will be able to investigate the associations more robustly in a large population, where I can restrict the analysis to postmenopausal women who are not on MHT and control for important potential confounders.

Chapter 4. Breast cancer risk factors and mammographic density in the Million Women Study

4.1 Introduction

In this chapter, I present an analysis of breast cancer risk factors and mammographic density in the MWS. As with Chapter 3, this will allow me to explore the first part of the pathway shown in Figure 1 of Chapter 1, but this time by conducting a cross-sectional analysis.

As evident from Chapter 3, there are many studies which have examined the associations between breast cancer risk factors and mammographic density. The meta-analysed evidence suggested that alcohol consumption, increasing age at first birth, and increasing age at menarche were associated with higher MD. In contrast, cigarette smoking and increasing parity were inversely associated with MD. The meta-analysis also suggested that these associations for alcohol, smoking and age at menarche were greater in premenopausal women than in postmenopausal women.

However, as discussed at the end of Chapter 3, the systematic review had several limitations. This included few small studies in the meta-analysis for some of the risk factors, therefore limiting its power. There was also a significant degree of heterogeneity across the studies, which made the interpretation of the pooled results challenging. Most of the studies in the systematic review included women on MHT, with most not presenting their analyses separately by MHT. As MHT is known to increase MD, this may have masked the lesser effects of some of the other risk factors on MD, or conversely compounded these associations, as suggested by two studies in the systematic review.[143,145] Several of the studies in the review did not adjust for BMI. BMI is an important determinant of MD as it will directly affect the volume of fatty tissue within the breast, which reduces the overall proportion of fibroglandular tissue. Therefore, the results may have been subject to confounding by BMI and possibly other factors. A further limitation was that some studies combined populations of pre- and postmenopausal women. The former would have denser breasts. As suggested in the review, there were differential effects of the risk factors on MD by menopausal

status, and therefore associations should be examined separately in pre- and postmenopausal women.

To address these limitations, I conducted a novel analysis of postmenopausal women within the MWS, in which I was able to examine several breast cancer risk factors. As detailed in Chapter 2, the MWS is uniquely placed in that it was able to collect data on several anthropometric, lifestyle, medical, reproductive, familial and socio-demographic factors which contribute to breast cancer risk. These data will allow me to control for confounders when studying the various associations. Of particular importance are MHT use and BMI. In the present study, I will be able to finely adjust for BMI and examine the effects of other risk factors on MD, by categories of BMI, to understand any effect modification by BMI. I will also be able to explore these associations in non-MHT users.

Examining the determinants of MD can aid in understanding the pathophysiology and aetiology of breast cancer and potentially guide breast cancer risk assessment, preventative measures, and tailored screening programmes. As MD has not been studied within the MWS before, the present study will allow an examination of the distribution of MD across a UK-based screening cohort. I will start the analysis by characterising MD in the current population and then examine the associations between the risk factors and MD.

4.2 Aims

The aims of this chapter were to:

- Characterise MD in this population of UK women within the MWS
- Determine the association between patterns of MHT use and MD
- Determine the association between other breast cancer risk factors and MD, independently of MHT
- Assess the effect modification by BMI between breast cancer risk factors and MD

4.3 Methods

4.3.1 Participants and data

As described in Chapter 2, the MWS is a population-based prospective study that recruited 1.3 million women aged 50-64 years through 66 participating breast screening centres in England and Scotland in 1996-2001. Participants provided information on their anthropometric, lifestyle, reproductive, medical, familial, and exogenous hormone use, including MHT, and sociodemographic risk factors at recruitment. They were resurveyed approximately every three to five years after recruitment. In 2023, the study commenced the collection of information on mammographic imaging and MD data from several screening centres. The present analyses are based on preliminary data from the Oxford Breast Screening Centre, as this was the first centre for which data collection was complete.

MWS participants were eligible for inclusion in this analysis if they had at least one digital mammogram as part of their routine screening at the Oxford Breast Screening Centre. As digital mammography was introduced around 2010, over a decade after recruitment into the MWS began (mean 15.2 years [SD 5.1]), all the women in the present study were postmenopausal at the time of their first digital screening mammogram. For the women in Oxford in the MWS, digital mammograms were available for screens between May 2011 and January 2018. As some women attended more than one screen during this period, multiple digital mammograms were available for a subgroup of women, depending on their age in 2011.

Major changes in MD due to menopausal status typically occur around the menopausal transition. Given that the women in the MWS would have gone through menopause roughly over a decade before the introduction of digital mammography, the women might be expected to have relatively low MD, on average, which is stable over time. Therefore, as an initial part of my analysis, it was important to characterise the distribution of MD in the MWS to determine the degree of variation across the available population.

4.3.2 Exposures: breast cancer risk factors

The risk factors of interest in this study included MHT status (never, past, current/recent), last used MHT type (combined (O+P), oestrogen-only, other) amongst current/recent users, age at screen (<65, 65-69, 70+ years), parity (0, 1-2, 3+ births), age at first birth (nulliparous, <20,20-24, 25-29, 30+ years), age at menarche (<12, 12-13 14+ years), breastfeeding (nulliparous, yes/no), oral contraceptive use (yes/no), bilateral oophorectomy (yes/no), alcohol (<30, 3-7, 8+ units/week), smoking (never, past, current and never/ever), first degree relative (mother or sister) with a history of breast cancer (yes/no), BMI (<25, 25-29, 30+kg/m^2), exercise (<1,1, 2-3, 4+ sessions/week), benign breast disease (yes/no), and time since menopause (<20, 20-24, 25-29, 30+ years). MHT type was investigated to determine whether the associations differed by the presence of progestogens.

Information on MHT use, alcohol, smoking, exercise, family history, and BMI were taken from the latest returned questionnaire prior to their first digital screening mammogram, in order to get the most up to date exposure status and avoid misclassification. Information for the remaining risk factors, except for age at screen, were taken from the recruitment questionnaire at the start of MWS, as these were unlikely to have changed throughout the follow-up period. Current or recent MHT use was defined as MHT use reported within 4 years of the relevant screening mammogram. Where current use was reported more than 4 years prior to the screening mammogram, the status was set to "Missing" to avoid misclassification of exposure and, therefore, potential dilution of effect. Missing was not included as a separate category as it would not have represented a meaningful classification of MHT use. Past and never use was based on the latest available survey, prior to the digital screen, as past users and never users at this age (>60 years old) were unlikely to become users.[14]

4.3.3 Outcome: mammographic density

The main outcome was mammographic density from the first available digital mammogram. MD data were provided by the Medical Physics Department at the Royal Surrey County Hospital, Guildford, UK. Automatic software (Volpara Health Technologies Ltd: Version 1.5.1, New Zealand) [125] was applied to a subset of digital mammograms in women who were not part of the MWS to determine their MD.

For these images, a Volpara Density Grade was determined, which was based on BIRADS scoring. The definition of the BIRADS categories of fibroglandular tissue are as follows: A, almost entirely fat; B, scattered fibroglandular tissue; C, heterogeneous fibroglandular tissue; and D, extreme fibroglandular tissue.[161]

In the subgroup of women with categorical MD derived from Volpara and processed mammographic images, the group in Surrey used machine learning to develop an algorithm for determining categorical MD using the subset of Volpara measurements as the ground truths. This algorithm was then used to estimate volumetric breast density from which the BIRADS score (A-D) could be determined for all digital mammographic images available within the MWS. This method had been previously validated by determining the correlation between the ground truths from Volpara and the prediction with deep learning algorithms (Pearson correlation coefficient (r) = 0.96).[126]

To examine associations of MD with breast cancer risk factors, the BIRADS score was categorised as low MD (A/B) and high MD (C/D). This was so the results could be compared to the literature as the majority of previous studies dichotomised MD. Furthermore, as there were fewer than 10,000 women with MD data, BIRADS needed to be dichotomised in order to provide enough power in the individual categories.

4.3.4 Statistical analysis

A descriptive analysis was presented to characterise the distribution of all the available MD data in the MWS. MD was tabulated by age at first digital screen and MHT use. Changes in BIRADS for women with multiple screens were also tabulated, as well as MD changes, within the same women, by age at screen and by MHT use across women with at least two digital screens.

For analyses of the association of MD with breast cancer risk factors, women who returned their MWS recruitment questionnaire after receiving the results of their recruitment mammogram were excluded. This was because those who were diagnosed with breast cancer may have been more likely

to recall certain risk factors, especially those thought to be associated with breast cancer, than women who were not diagnosed with breast cancer.

As MHT is associated with higher MD and, therefore, may modify the effects of other risk factors on MD, I first examined associations between MHT and MD before investigating other factors. For the analysis of other risk factors, current MHT users were removed to avoid any effect modification from MHT use. Women with missing MHT status were also excluded from all analyses.

Binary logistic regression was used to estimate the association of breast cancers risk factors with MD, from the first available digital mammographic screen. The results are presented as odds ratios (ORs) and 95% confidence intervals (95% CIs) for high versus low MD. Three models were fitted: adjusted for age at screen, adjusted for BMI and age at screen, and a fully adjusted model, with all breast cancer risk factors in the model, including BMI and age at screen. As BMI is known to have a substantial effect on MD, by adjusting for it separately, I could compare the extent of the effects of risk factors on MD both prior to BMI adjustment and after BMI adjustment. Comparing the effect sizes after adjusting for BMI and after the fully adjusted model would indicate whether BMI confounded the associations seen to a greater or similar extent to that of other risk factors. Whilst it is possible that changes in estimates after adjusting for BMI could be due to non-collapsibility, it is more likely that BMI is a confounder.

Linear trends in log ORs were calculated for ordinal risk factors (age, BMI, alcohol, exercise, parity, age at first birth, age at menarche and time since menopause). This was done using the median value within each category of the risk factor.

Adjustment variables were coded as described in section 4.3.3, except for age at screen, which was included as finer categories (61, 62, 63, …, 79, 80+ years), and BMI (<22.5, 22.5-24, 25-27.4, 27.5-29, 30+), to minimise residual confounding by these factors. Missing values for adjustment variables were added as a separate category for each variable. These variables included parity (46 [0.5%]); age at first birth (173 [1.7%]); age at menarche (172 [1.7%]); breastfeeding in parous women (3037 [34.0%]);

time since menopause (6028 [60.7%]); oral contraceptive use (63 [0.6%]); bilateral oophorectomy (339 [3.4%]); alcohol (16 [0.2%]); smoking (113 [1.1%]); family history of breast cancer (552 [5.6%]); BMI (199 [2.0%]); exercise (1840 [18.5%]); and benign breast disease (52 [0.5%]).

It was important to adjust for time since menopause as this would directly affect the time since decline in ovarian function, and therefore levels of circulating endogenous oestrogens and progesterone, which would affect MD. Time since menopause could also be associated with other risk factors of interest, like BMI or MHT use. Time since menopause was missing for around 60% of women. This, in part, was because N=1777 (17.0%) had a hysterectomy before their natural menopause. In main analyses, time since menopause was adjusted for using a variable with separate categories for women with hysterectomy prior to natural menopause and for women who did not provide an age at menopause (<20, 20-24, 25-29, 30+, hysterectomy, missing [N=4251, 42.8%]). A sensitivity analysis, as described subsequently, explored the likely impact of this lack of complete adjustment for time since menopause.

Likelihood ratio tests were used to assess heterogeneity or trends in ORs across categories of ordinal risk factors.

4.3.4.1 Subgroup analysis

As described earlier, BMI is known to have a large effect on MD. Therefore, the association of each risk factor with MD was examined within subgroups of BMI. For MHT, the subgroups of BMI were <25 and 25+ kg/m^2, as there were few current MHT users. For all other risk factors, the subgroups of BMI were <25, 25-29 and 30+ kg/m^2. A likelihood ratio test was used to test for heterogeneity across subgroups.

4.3.4.2 Sensitivity analysis

As time since menopause was missing for roughly two-thirds of the women, a sensitivity analysis was conducted where only women with known time since menopause were included. Women with a hysterectomy were excluded from this analysis. The main analysis was repeated in this subset of

women, for which I initially did not adjust for time since menopause and then did adjust for time since menopause. These results were then compared. The results from the sensitivity analysis were also compared to the main analysis, where a crude time since menopause adjustment was used. These analyses were to determine the effect of having a large amount of missing data on time since menopause and the effect of adjusting for time since menopause on the overall interpretation of the results.

4.3.4.3 Plausibility of mammographic density as a mediator

A crude analysis was conducted to determine whether it was plausible that MD mediated the well-established associations between the risk factors considered here and breast cancer risk. This was attempted by checking whether the effect of a given risk factor on breast cancer risk might be fully explained through its effect on MD. This was based on the estimates of the associations of these factors with MD observed here, previously published estimates of the effects of MD on breast cancer risk, and published estimates of the associations of these risk factors with breast cancer, to determine what magnitude of effect needed to be explained.

The analysis was restricted to risk factors whose association with MD was in the same direction as its association with breast cancer risk and to risk factors where it was plausible that they acted through MD. The former was because if the association between a risk factor and MD was in the opposite direction of that between the risk factor and breast cancer, then MD could not be a mediator. BBD was excluded as this is more likely to be a consequence of MD rather than a cause of increased MD. For each risk factor, I first calculated an estimated mean MD for women within each risk factor category. To do this, each BIRADS category (A-D) was assigned a representative percentage MD using the midpoint percentage density of the category, based on the ranges: BIRADS A, <25%; B, 25-50%, C, 51-75%, D, >75%. Therefore, BIRADS A corresponded to 12.5%; B to 37.5%; C to 62.5%; and D to 87.5%. Then, for each risk factor, a mean percentage MD for women in each category was estimated based on the BIRADS specific percent densities and the frequency distribution of BIRADS among women in

that risk factor category. The difference in estimated mean percentage MDs between the highest and lowest risk factor categories was then calculated for each ordinal risk factor. For MHT use, the difference in mean percentage MD between current and never users was used. An example for parity (3+ births vs nulliparous) is given in Table 13.

The assumed magnitude of the associations between percentage MD and breast cancer were based on the findings from Yaghjyan et al. (2011).[162] The authors reported an OR for breast cancer by increasing categories of percentage MD. [162] In order to apply this to my data, I fitted a trend through the ordinal measures of the percentage MD, using generalised least squares as described by Greenland and Longnecker.[135] This method had to be used as the standard errors associated with the ORs were not independent of each other, as they would have been relative to a baseline category. This is the same method used in Chapter 3 for generating a trend across ordinal risk factors. From this, I was able to estimate the increase in breast cancer risk that would be expected to be associated with the corresponding difference in percentage MD across risk factor categories. These associations were then compared to the associations between the risk factor and breast cancer, using estimates published in the literature. [16,163–165] By doing so, I could compare whether the anticipated effect of the difference in MD associated with each risk factor on breast cancer risk was of a similar magnitude to the effect of the risk factor itself on breast cancer risk.

Table 13. Process to crudely calculate the mean percentage MD for risk factor categories using parity as an example

BIRADS	**A**	**B**	**C**	**D**	**Total**	**Weighted mean**
Assigned %MD	12.5	37.5	62.5	87.5		
Nulliparous (N)	106	374	218	86	784	
3+ births (N)	382	1502	423	67	2374	
N x %MD						
Nulliparous	1325	14,025	13,625	7525	36,500	46.6
3+ births	4775	56,325	26,437.5	5862.5	93,400	39.3
Difference (%)						-7.3

Analyses were conducted in Stata 18 [127] and forest plots were created in R-Studio using the Jasper package.[129]

4.4 Results

4.4.1 Characterisation of mammographic density in the study population

At least one digital mammogram was available for 9928 women from the Oxford screening centre. Of these, 5884 had a second digital mammogram, and of these, 914 had a third digital mammogram. Mean age at recruitment in the current population was 53.8 years (SD 3.4), whereas the mean age at first digital mammogram was 69.1 years (SD 3.7). Dates of the screens, where digital mammography was available, and mean age at screen are given in Table 14. Mean time between the first and second digital mammographic screen was 3.0 years and between second and third screen was 2.9 years. Mean time between menopause and first digital mammogram was 21.9 years (SD 5.1), amongst those with known age at menopause.

Table 14. Average age and dates at digital mammographic screens

Digital mammographic screen	**N**	**Mean age (SD) at the digital mammographic screen**	**Dates of the digital mammographic screen**
Screen 1	9928	69.1 (3.7)	12/5/2011 – 14/1/2018
Screen 2	5884	70.9 (3.3)	10/7/2012 – 19/12/2017
Screen 3	914	72.6 (3.3)	12/6/2015 – 23/12/2017

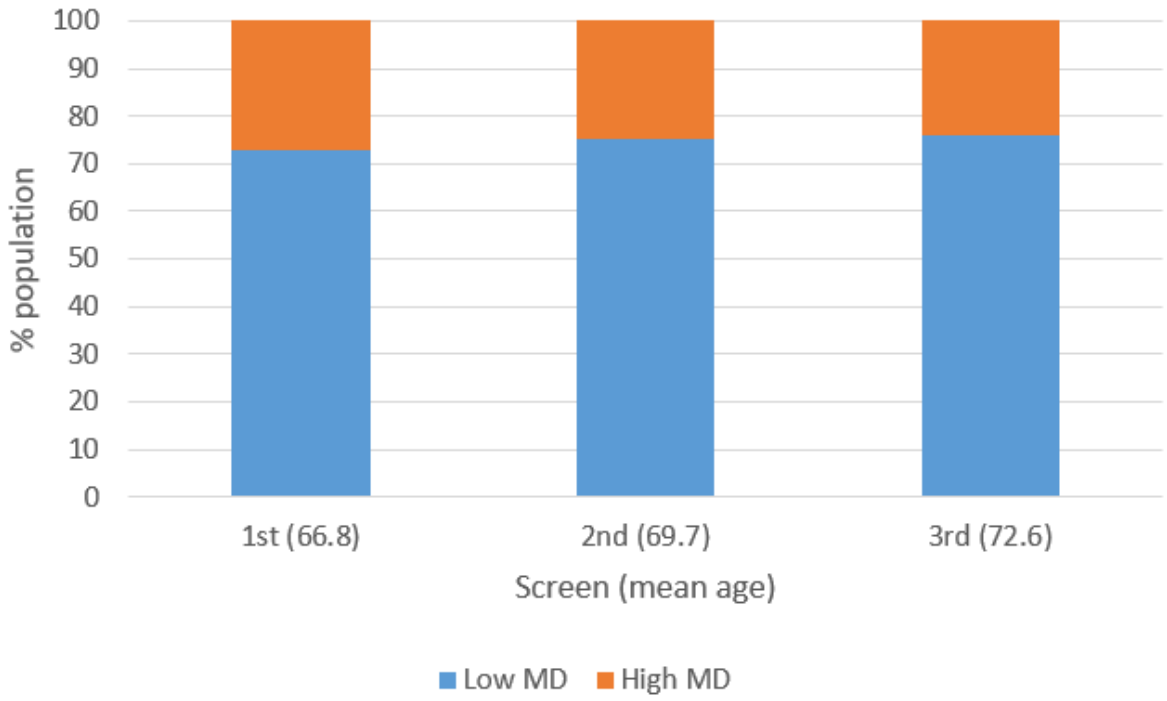
In the total of 9928 women with at least one digital screen, the distribution of high/low MD scores both overall, and by age at first digital screen and MHT use is shown in Table 15. Most women had BIRADS B (~60%). Roughly 5% of women had BIRADS D. The proportion of women with high MD decreased as age at first screen increased: 30.3% at <65 years and 23.1% at 70+ years. There was a greater proportion of women with high MD among current MHT users (36.6%) than among never (26.6%) or past users (24.4%).

Table 15. Distribution of BIRADS and high/low MD scores by age and MHT at first digital screen N(%)

Digital mammographic screen	**A**	**B**	**C**	**D**	**Low MD**	**High MD**	**Total**
N	1387 (14.0)	5943 (59.9)	2109 (21.2)	489 (4.9)	7330 (73.8)	2598 (26.1)	9928
Age (years)							
<65	126 (15.6)	438 (54.1)	190 (23.5)	55 (6.8)	564 (69.7)	245 (30.3)	809
65-67.4	420 (13.8)	1764 (57.8)	699 (22.9)	168 (5.5)	2184 (71.6)	867 (28.4)	3051
67.5-69	409 (14.2)	1720 (59.8)	604 (21)	145 (5.0)	2129 (74.0)	749 (26.0)	2878
70+	432 (13.5)	2021 (63.4)	616 (19.3)	121 (3.8)	2453 (76.9)	737 (23.1)	3190
MHT status							
Never	608 (14.3)	2522 (59.1)	913 (21.4)	222 (5.2)	3130 (73.4)	1135 (26.6)	4265
Past	614 (14.6)	2564 (61)	829 (19.7)	199 (4.7)	3178 (75.6)	1028 (24.4)	4206
Current	19 (8.2)	128 (55.2)	76 (32.8)	9 (3.9)	147 (63.4)	85 (36.6)	232

The following results examine changes across the three available digital screens amongst women with more than one digital mammographic screen. The distribution of BIRADS scores across these screens remained largely the same. Among women with three digital screens, there was a slight increase in the proportion of women with low MD with time (Figure 12).

Figure 12. MD classifications at each screen among women with three digital screens



Between consecutive digital screens (first to second or second to third screen) BIRADS scores largely remained the same (Figure 13). 80.5% of women who were BIRADS D at their first screen also were BIRADS D at their second screen. Whereas 76.1% of women who were BIRADS D at their second screen remained BIRADS D at their third screen, suggesting some but small changes across time.

Due to the relatively few women with multiple screens, the distribution of MD across the screens, by age at first digital screen and MHT use at first digital screen was examined in terms of high and low MD, rather than as a BIRADS score, and between the first and second screen.

Changes between first and second screen, by age at first screen remained largely the same across the different age categories (<65, 65-69, 70+). Across all age categories, 97% of women with low MD at their first screen, had low MD at their second screen, and 80-81% of women with high MD at the first screen, had high MD at their second screen (Figure 14).

There were 2436 never, 2297 past and 132 current MHT users with two digital screens (Figure 15). In never and past users, 97% of women with low MD at their first screen had low MD at their second

screen. 80-81% of women with high MD at their first screen had high MD at their second screen. Among current users, 85% of women with high MD at their first screen had high MD at their second screen. 95% of women with low MD at their first screen had low MD at their second screen.

Figure 13. Heat map of BIRADS score between first and second digital screen and between second and third screen (figures displayed are a percentage of the numbers at the earlier screen)

BIRADS at 2nd screen

BIRADS at 1st screen	A	B	C	D	Total
A	82.2	17.8	0.0	0.0	811
B	9.5	86.5	3.9	0.1	3517
C	0.0	22.5	74.1	3.4	1263
D	0.0	1.0	18.4	80.5	293
Total	1,002	3,473	1,128	281	5,884

BIRADS at 3rd screen

BIRADS at 2nd screen	A	B	C	D	Total
A	78.4	21.6	0.0	0.0	153
B	5.1	89.9	5.1	0.0	534
C	0.0	19.3	74.6	6.1	181
D	0.0	0.0	23.9	76.1	46
Total	147	548	173	46	914

Legend

>70%
30-70%
10-30%
<10%

Figure 14. Heat map of high/low MD between first and second digital screen by age at first screen (figures displayed are a percentage of the numbers at the first screen)

Age <65		**MD at 2nd screen**		
		Low MD	High MD	Total
MD at 1st screen	Low MD	97.0	3.1	393
	High MD	20.4	79.6	167
	Total	415	145	560

Age 65-69		**MD at 2nd screen**		
		Low MD	High MD	Total
MD at 1st screen	Low MD	96.8	3.2	2489
	High MD	18.7	81.3	898
	Total	2,578	809	3,387

Age 70+		**MD at 2nd screen**		
		Low MD	High MD	Total
MD at 1st screen	Low MD	97.0	3.0	704
	High MD	19.2	80.8	214
	Total	724	194	918

Legend
>70%
30-70%
10-30%
<10%

Figure 15. Heat map of high/low MD between first and second digital screen by MHT status at first screen (figures displayed are a percentage of the numbers at the first screen)

Never MHT		**MD at 2nd screen**		
		Low MD	High MD	Total
MD at 1st screen	Low MD	96.6	3.4	1771
	High MD	18.8	81.2	665
	Total	1,836	600	2436

Past MHT		**MD at 2nd screen**		
		Low MD	High MD	Total
MD at 1st screen	Low MD	97.2	2.8	2489
	High MD	18.7	81.3	898
	Total	2,578	809	3387

Current/ recent MHT		**MD at 2nd screen**		
		Low MD	High MD	Total
MD at 1st screen	Low MD	95.4	4.7	86
	High MD	15.2	84.8	46
	Total	89	43	132

Legend
>70%
30-70%
10-30%
<10%

4.4.2 Analysis: Breast cancer risk factors and mammographic density

Given that there was a degree of variation in MD in the present population, despite having a mean age of 69 years, further analyses by risk factors was justified.

From the population of 9928 women with digital mammograms, 579 were excluded as these were women who returned their questionnaires after their screen. Women whose MHT status was missing (N=67) and who reported current use more than 4 years prior to digital mammographic screen (N=1092) were also excluded. This left 8190 women for the MHT analysis. Table 16 displays participant risk factors, used in the analysis, by level of BIRADS at the first available digital mammogram. 74.4% of women had low MD (BIRADS A/B).

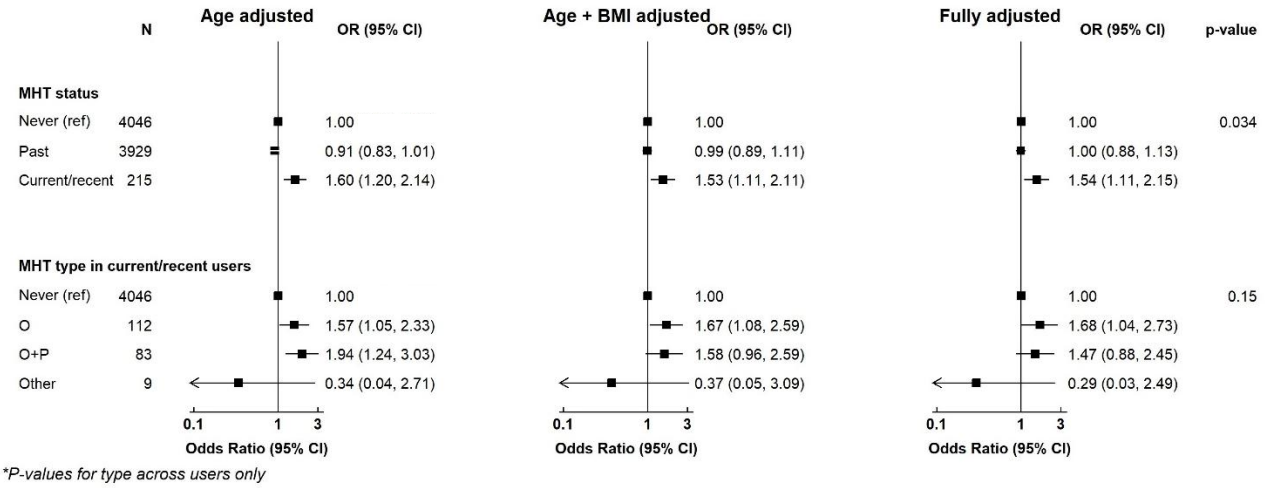
Table 16. Risk factors of women included in the analysis by BIRADS score at first digital screen

Risk factor	A	B	C	D
N	1,173 (14.3%)	4,921 (60.1%)	1,704 (20.8%)	392 (4.8%)
Age at first digital screen	68.9 (3.4)	69.3 (3.8)	68.8 (3.6)	68.5 (3.5)
Body mass index (kg/m^2)	30.7 (5.3)	26.9 (4.2)	23.8 (3.1)	21.9 (2.8)
Time since menopause (years)	20.3 (5.2)	22.0 (5.1)	21.5 (5.1)	21.3 (5.1)
Number of full-term pregnancies	2.2 (1.1)	2.2 (1.1)	1.9 (1.1)	1.6 (1.1)
Ever breastfed among parous women	546 (39.4%)	2,520 (42.4%)	906 (43.0%)	181 (37.0%)
Age at first birth (years)	23.6 (4.3)	24.0 (4.2)	25.1 (4.7)	26.2 (4.8)
Age at menarche (years)	12.6 (1.5)	12.8 (1.5)	13.0 (1.5)	13.3 (1.4)
Ever oral contraceptive user	376 (32.1%)	1,569 (31.9%)	484 (28.4%)	98 (25.0%)
Bilateral oophorectomy	79 (6.7%)	301 (6.1%)	79 (4.6%)	11 (2.8%)
Alcohol drinks per week	3.8 (6.1)	4.8 (6.5)	5.8 (6.6)	6.1 (6.5)
Current smoker	121 (10.3%)	452 (9.2%)	150 (8.8%)	33 (8.4%)
Family history of breast cancer	130 (11.1%)	651 (13.2%)	243 (14.3%)	63 (16.1%)
Benign breast disease	73 (6.2%)	559 (11.4%)	294 (17.3%)	83 (21.2%)
Current/recent MHT user	17 (1.4%)	119 (2.4%)	71 (4.2%)	8 (2.0%)

Results shown are N (%) or mean (SD)

In the fully adjusted model, there was evidence of an association between MHT use and MD (p=0.034). Current or recent MHT use was associated with high MD, compared to never MHT users (OR 1.54 [1.11, 2.15]). Past MHT use was not associated with high MD. These associations remained consistent in all three models of different levels of adjustment (Figure 16). There was no evidence of an association between type of MHT among current users (p=0.15). The age-adjusted association for combined (O+P) MHT was slightly attenuated by adjustment for BMI and the other risk factors of interest. The OR for oestrogen-only MHT became stronger after adjustments.

Figure 16. Association between MHT status/type and MD



P-values for type across users only

Fully adjusted: adjusted for age at screen, parity, age at first birth, age at menarche, breastfeeding, oral contraceptive use, bilateral oophorectomy, alcohol, smoking, first-degree relative with a history of breast cancer, BMI, exercise, benign breast disease, and time since menopause. Abbreviations: O, oestrogen-only MHT; O+P, combined MHT.

For the remaining risk factor analysis, current MHT users were excluded, leaving 7975 women. The results of this analyses are presented in Figure 17.

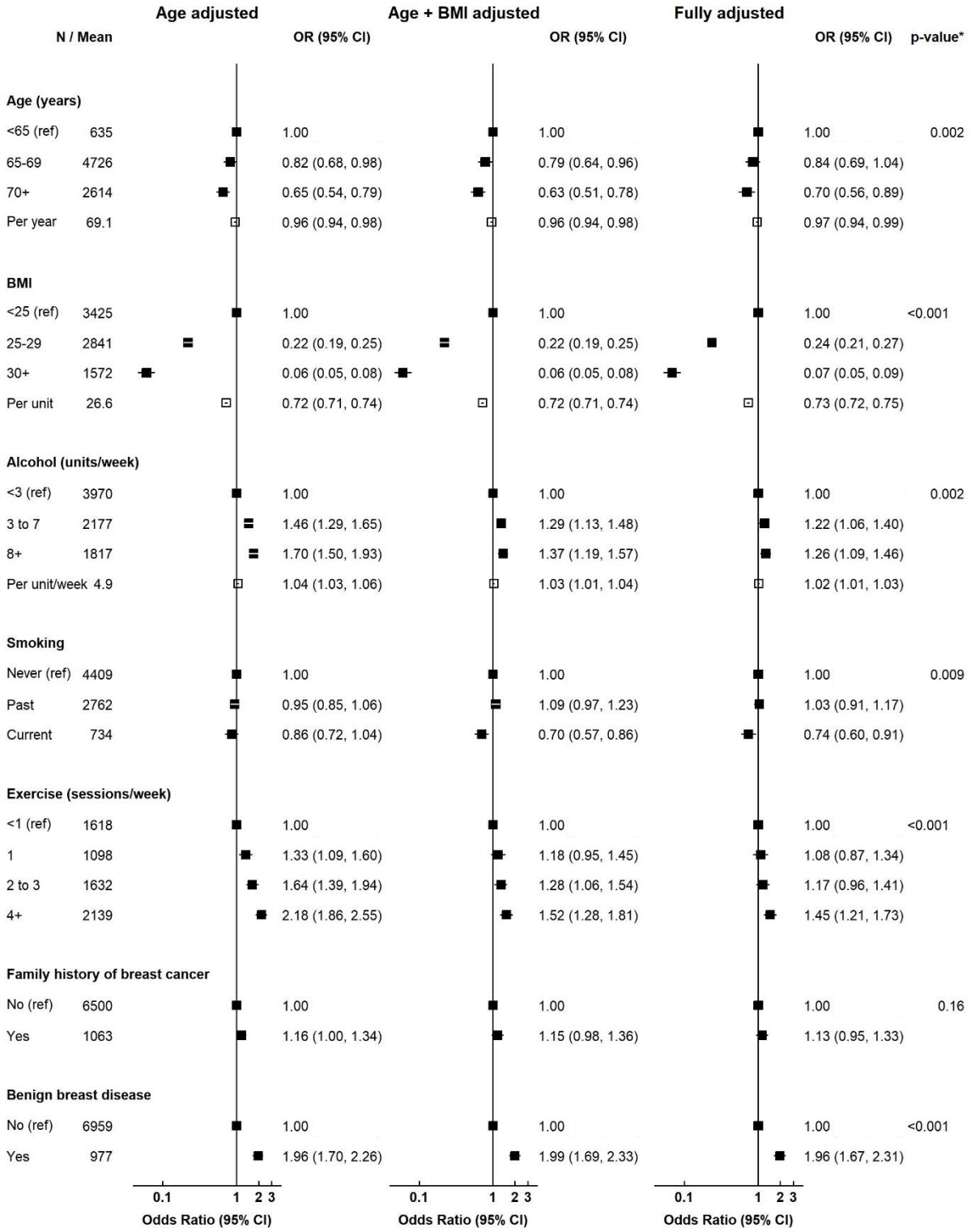
In the age-adjusted analysis, most risk factors were associated with MD. Increasing age at screen, increasing BMI, increasing parity and bilateral oophorectomy were inversely associated with MD. Increasing age at first birth, increasing age at menarche, oral contraceptive use, higher alcohol

consumption, exercise, family history of breast cancer and BBD were associated with higher MD. There was no evidence of an association between breastfeeding, time since menopause and smoking, and MD. After additionally adjusting for BMI, many of the associations were attenuated. However, the OR for current smoking became stronger after adjusting for BMI (age-adjusted: 0.86 vs age + BMI adjusted: 0.70).

In the fully adjusted model, the results remained much the same. Increasing BMI had the largest association with MD (OR for 30+ vs <25: 0.07 [0.05, 0.09], p<0.001). This was followed by BBD (yes vs no BBD: 1.96 [1.67, 2.31], p<0.001), parity (3+ births vs nulliparous: 0.43 [0.34, 0.53), p<0.001], exercise (4+ vs <1 sessions/week: 1.45 [1.21, 1.73], p<0.001), and age (70+ vs <65 years: 0.70 [0.56,0.89], p=0.002). Each unit increase in BMI was associated with a 27% reduction in the odds of having high MD (0.73 [0.72, 0.75]), and each additional birth was associated with a 25% reduction in the odds of having high MD (0.75 [0.70, 0.81]). Greater alcohol intake was associated with high MD (8+ vs <3 units/week: 1.26 [1.09, 1.42], p=0.002). Smoking was inversely associated with MD (current vs never smokers: 0.74 [0.60, 0.92], p=0.009). There was no evidence of an association between MD and breastfeeding, oral contraceptive use, bilateral oophorectomy, family history, and time since menopause in the fully adjusted model.

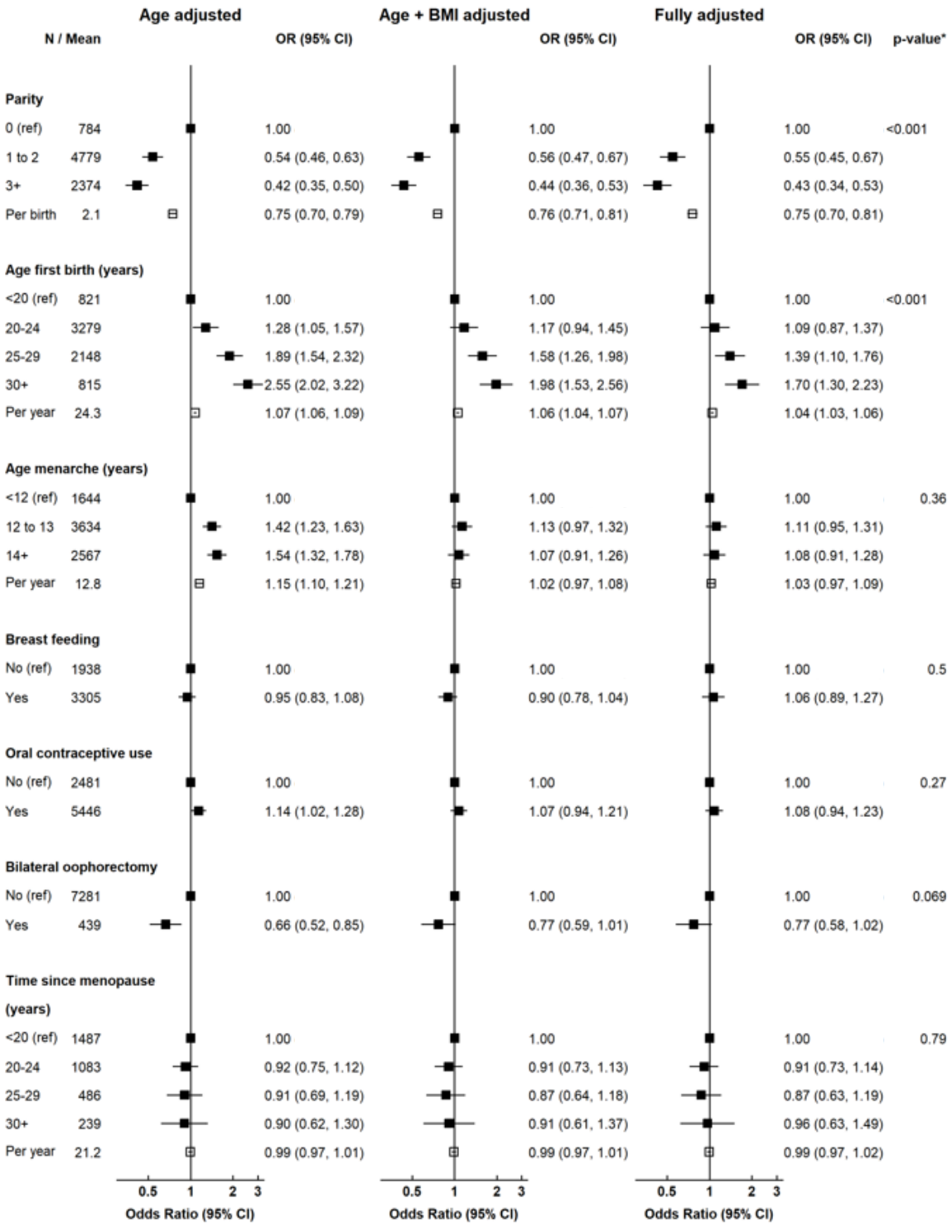
In the subgroup analysis, MHT (Figure 18) and other risk factors (Figure 19) were also examined by BMI categories. The analysis of other risk factors again excluded women who were currently or recently on MHT. There was no evidence of effect modification by BMI for all risk factors except for BBD (p=0.009). In the highest BMI category (30+), the magnitude for BBD vs no BBD was far greater than in the lower categories (4.50 [2.63, 7.72] for BMI 30+ vs 1.77 [1.45, 2.17] for BMI <20).

Figure 17. Analysis of breast cancer risk factors and MD among women not currently on MHT



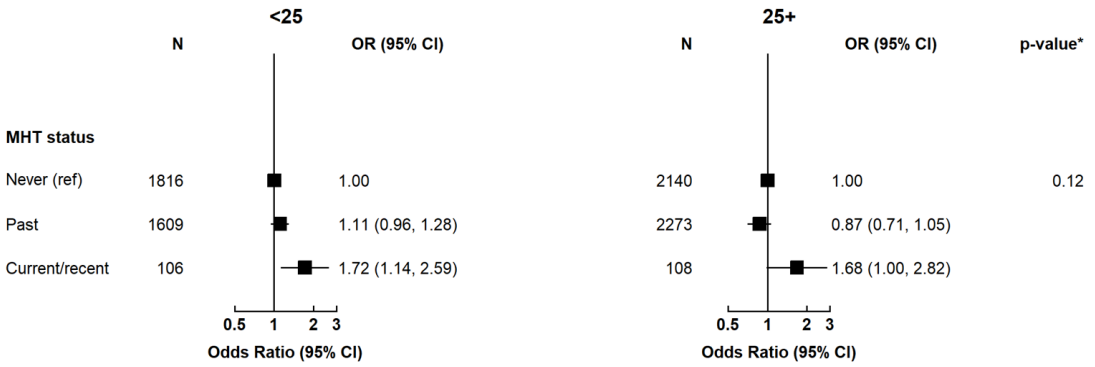
**P(trend) for continuous and ordinal variables*
Current MHT users excluded

(Figure 17 continued)



*P(trend) for continuous and ordinal variables
Current MHT users excluded

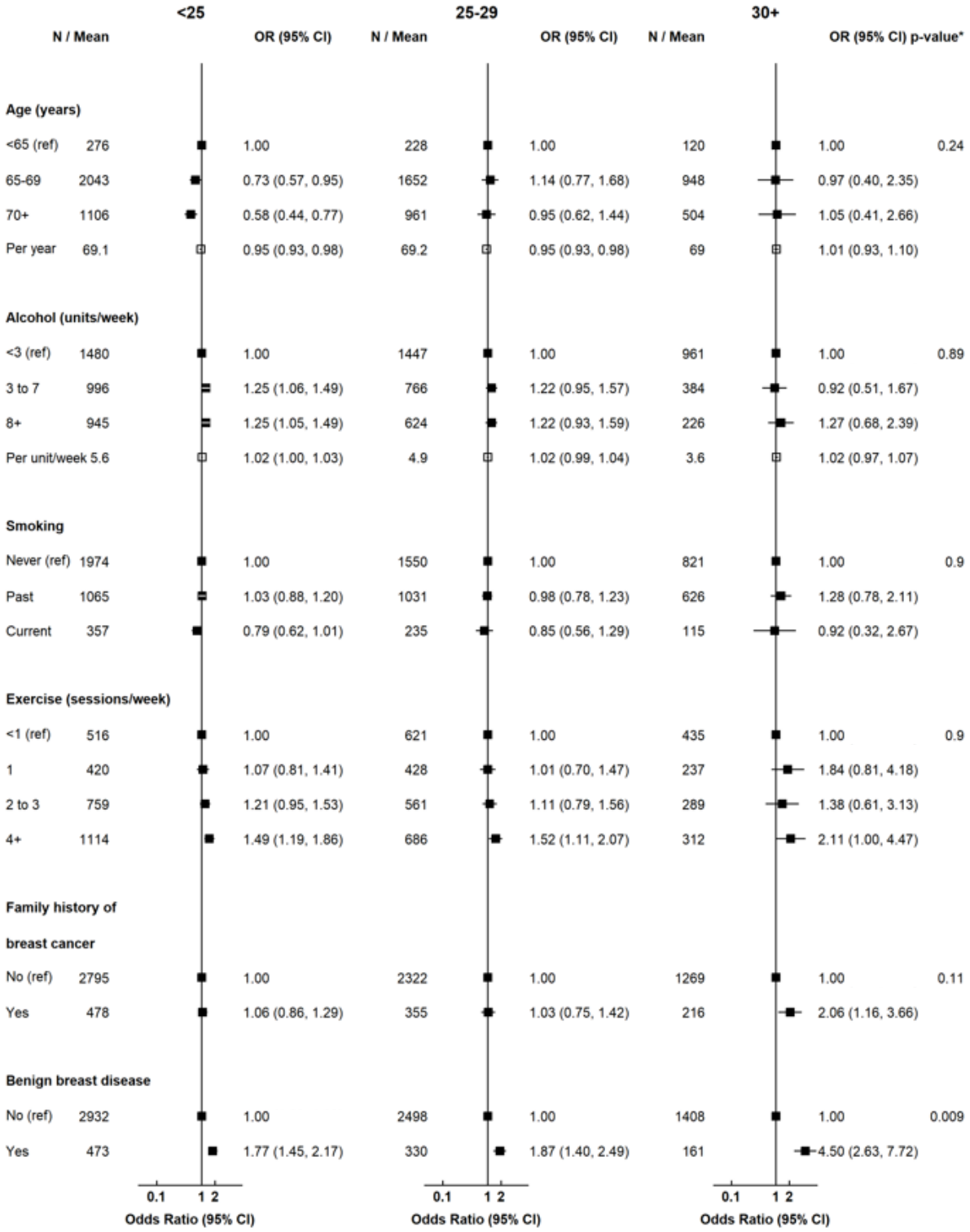
Figure 18. Association between MHT status and MD by BMI categories



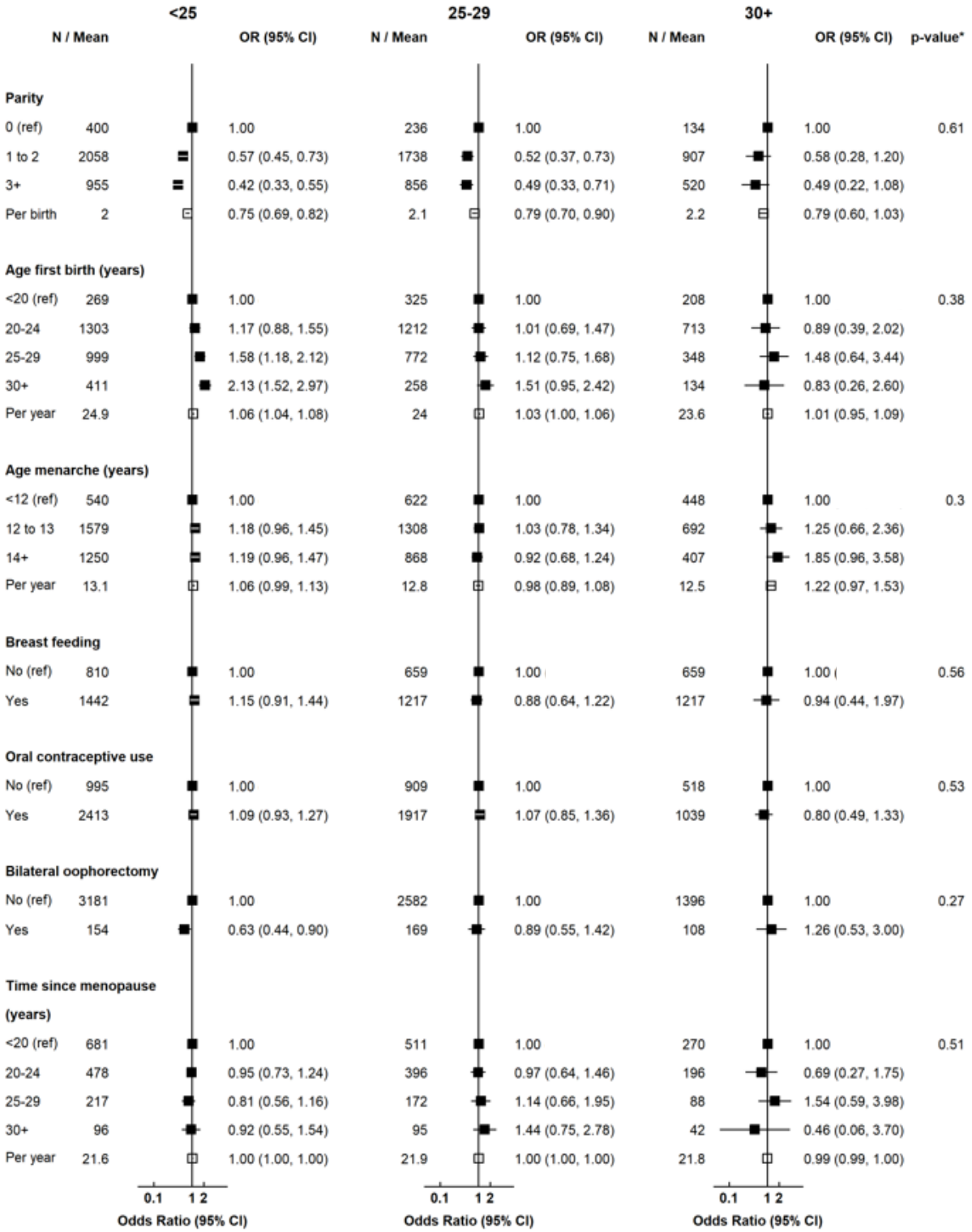
**P-value for interaction*

Adjusted for age at screen, parity, age at first birth, age at menarche, breastfeeding, oral contraceptive use, bilateral oophorectomy, alcohol, smoking, first-degree relative with a history of breast cancer, exercise, benign breast disease, and age at menopause. Abbreviations: O, oestrogen-only MHT; O+P, combined MHT.

Figure 19. Association between breast cancer risk factors and MD by BMI categories among women not currently using MHT



(Figure 19 continued)



4.4.2.1 Sensitivity analysis

In the sensitivity analysis, only women with known time since menopause were included (N=3295). The two sets of results in which time since menopause was not adjusted for and then was adjusted for were almost identical. For most of the risk factors, the magnitude and direction of the association in the sensitivity analysis were comparable with the main analysis (Appendix 6).

4.4.2.2 Comparison with systematic review

The results in the present analyses were relatively consistent with the results from the systematic review in Chapter 3 (Table 17).

Table 17. Comparison of systematic review results and MWS results

Risk factor	Review: OR (95% CI)	MWS: OR (95% CI)
Alcohol		
per 10 g/day	1.06 (1.02, 1.10)	-
Units/week	-	1.02 (1.01, 1.03)
Smoking		
Ever vs never	0.97 (0.88, 1.07)	0.96 (0.85, 1.08)
Current vs never	-	0.74 (0.60, 0.91)
per 10 cigarettes/day	0.93 (0.90, 0.96)	
Parity		
per birth	0.88 (0.88, 0.88)	0.75 (0.70, 0.81)
Age at first birth		
per 1 year increase	1.04 (1.04, 1.05)	1.04 (1.03, 1.06)
Age at menarche		
per 1 year increase	1.04 (1.03, 1.04)	1.03 (0.97, 1.09)

4.4.2.3 Plausibility of mammographic density as a mediator

The associations between MHT, alcohol, parity, and age at first birth and MD were in the same direction as the associations between these risk factors and breast cancer. For all risk factors considered, the expected magnitude of the increase in breast cancer risk associated with the estimated difference in percentage MD between the top versus bottom category of each risk factor

was smaller than the corresponding increase in breast cancer risk associated with being in the top versus bottom category of each risk factor. This suggests that the effects of these risk factors may be partly, but not necessarily entirely, mediated by MD (Table 18). For example, current or recent MHT users had on average, a 3.8% higher percentage MD than never users, which would be expected to give rise to an OR of 1.07 for breast cancer. However, the OR for breast cancer in current vs never users, based on the literature, was much higher, at 1.66 (1.58, 1.75). [163]

Table 18. Crude mediation analysis of breast cancer risk factors, mammographic density and breast cancer

Risk factor	Categories	Mean %MD	Difference in mean %MD	Association (OR (95% CI)) between risk factor and breast cancer	Associations (OR (95% CI)) between difference in mean %MD and breast cancer risk
MHT use	Never	41.8		ref	
	Current/recent	45.6	3.8	1.66 (1.58, 1.75) [163]	1.07 (1.05, 1.09)
Alcohol	<3	39.4		ref	
(units/week)	8+	44.5	5.1	1.22 (1.13, 1.32)[164]	1.09 (1.07, 1.12)
Parity	0	46.6		ref	
	3+	39.3	-7.3	0.79 (0.74, 0.84)[16]	0.88 (0.85, 0.91)
Age at first birth	<20	37.5		ref	
(years)	30+	45.1	7.6	1.29 (1.12, 1.48) [165]	1.14 (1.11, 1.18)

4.5 Discussion

In the present study of roughly 8000 postmenopausal women, several risk factors were associated with mammographic density when measured using a BIRADS score. Current/recent MHT use, older age at first birth, alcohol consumption, physical activity, and benign breast disease were positively associated with MD. Age, parity, smoking, and BMI were inversely associated with MD. There was no association between MD and breastfeeding, oral contraceptive use, bilateral oophorectomy, family history of breast cancer and time since menopause.

The initial aim of this chapter was to characterise the distribution of MD in postmenopausal women in the study. Over 70% of women in the study had low MD (BIRADS A/B). This is slightly higher, but comparable with the proportion of women with low MD in other studies of postmenopausal women included in the systematic review in Chapter 3 (54-68%). [146,166–168] Mean age at recruitment into the MWS in the current population was 53.8 years, whereas the mean age at first digital mammogram was 69.1 years. Therefore, women in the present analysis would generally have had their menopause at least a decade before their first digital screen. In fact, mean time between menopause and first digital mammogram was 21.9 years, amongst those with known age at menopause. This meant I did not expect to observe a large variation in MD across the population or across time points for those women with more than one screening mammogram available. Most women included here had low MD by the time of their first digital screen, and any changes in MD across subsequent mammograms were very slight, which is consistent with previous studies which show that declines in MD decelerate rapidly after menopause.[52,169]

The main aim of the chapter was to examine the associations between several breast cancer risk factors and MD. Reassuringly, the results in the present analyses were relatively consistent with those from the systematic review in the Chapter 3.

In the systematic review, later age at menarche was associated with higher MD (OR 1.04 [1.03, 1.04]) but was not associated with MD after adjusting for BMI in the MWS (1.03 [0.97, 1.09]). As mentioned

in Chapter 3, a study from the International Consortium of Mammographic Density, which was published after I completed the systematic review, found that in over 10,000 women in 22 countries, later age at menarche was associated with higher MD but that the magnitude of the association was attenuated after adjusting for BMI.[156] Early-life adiposity is associated with earlier menarche and early-life adiposity is also known to be correlated with midlife adiposity.[170] This would suggest a corresponding association between midlife adiposity and early menarche. Since midlife adiposity is associated with early menarche and with reduced MD, it follows that early menarche is also associated with lower MD if midlife BMI is not adjusted for. There are many genetic variants (including, LIN28B, FSHR, LEP) linked to earlier menarche which are related to metabolism and reproductive development. These may also influence BMI and therefore MD.[171,172]

Interestingly, smoking was inversely associated with MD. There was some evidence of this in the systematic review in Chapter 3. Whilst the evidence between smoking and breast cancer is inconsistent, as discussed in Chapter 1, any possible increase in risk associated with smoking, may be due to the reduction of MD in smokers, improving the sensitivity of mammographic screening and not necessarily because smoking causes breast cancer. As such, smokers may appear to have an increased risk of breast cancer at screening. [173] however, the interplay of risk and screening sensitivity is complex and would need to be studied further. Further analyses by duration or pack-years of smoking may highlight if this relationship has a dose-response.

Age was inversely associated with MD, which has been well documented. The International Consortium of Mammographic Density performed a pooled analysis of cross-sectional data from 22 countries worldwide. They reported a decrease in the square root of percentage MD per 10-year increase in age in both premenopausal and postmenopausal women of 0.24% (-0.34, -0.14) and 0.38% (-0.44, -0.33), respectively.[56] The reduction in MD with age can be explained by the decrease in endogenous hormones, which, in turn, leads to a reduction in the production and proliferation of

fibroglandular tissue and loss of utility of breast function, as well as the involution of breast tissue with age, after the reproductive years have ceased. [58]

In these data, current MHT use, but not past or never use, was clearly associated with higher MD. This association has been previously documented in randomised trials[174–176], observational studies,[177–179] and a systematic review [60]. The association between current MHT use and higher MD can be explained by the influence of exogenous hormones on breast tissues, causing the proliferation of the fibroglandular tissue, which corresponds to increased MD. The lack of association in past users suggests that the effect of MHT on MD is temporary. In a larger population, an analysis of time since last MHT use and MD would help guide how long the effects of MHT use on MD last.

In the fully adjusted model, there was a lack of evidence that the association between MHT and MD differed by MHT type. This is inconsistent with previous large studies which have found this association to be greater for combined MHT as opposed to oestrogen-only MHT. [180–183] In the Women's Health Initiative Trial, 435 women were randomly assigned to receive oestrogen-only MHT or placebo. After one year of follow-up, mean percentage MD increased by 1.6% (0.8, 2.4) among women receiving oestrogen MHT and decreased by 1.0% (-1.7, -0.4) in the placebo group. [61] In a separate trial, they also randomised 413 women to combined MHT or placebo. After one year of follow-up, mean percentage MD increased by 6.0% (4.6, 7.5) among women receiving combined MHT and decreased by 0.9% (-1.5, -0.2) in the placebo group.[62] A systematic literature review of 22 studies, including six trials and 16 observational studies, concluded that users of combined MHT generally had higher MD than users of oestrogen-only MHT.[60] Although I did not detect a difference by MHT type, this was likely driven by a lack of power. In the present study, I only had 112 combined users and 83 oestrogen-only users. The Women's Health Initiative Trial had over 200 women taking combined preparations and over 200 women taking oestrogen-only MHT.[61,62] Other observational studies that detected a difference between the preparations were also larger than the MWS, with over 400 women currently on MHT.[180–183]

BMI had by far the largest association with MD. Each unit increase in BMI was associated with a 27% reduction in the odds of having high MD (0.73 [0.72,0.75]). This is consistent with several previous observational studies.[58,184] For example, a large observational study of 24,556 women in the San Francisco Mammography Registry, which was designed to examine the association between BMI and MD, found that each BMI increase of 1kg/m^2 was associated with a 2.03% reduction in percentage density (-2.09, -1.98). [58] A further study with over 28,000 women reported that women with BMI <22.4 had a 7-fold increase in the odds of having high MD than women with BMI >29.8 (OR 7.1 [6.6, 7.6]). [184] The inverse association with MD is thoiught to be due to an increased deposition of fatty tissue in the breast as BMI increases, reducing the overall proportion of fibroglandular or dense tissue within the breast.

BBD had one of the largest positive associations with MD. The associations between BBD and MD have been reported previously. BBD in the MWS study was based on reporting having previous breast surgery for any benign breast pathology and was not obtained from any medical records. Therefore, there was no way of determining whether these lesions were proliferative benign breast diseases or whether these were fibroadenomas or cysts, which are less associated with breast cancer than proliferative lesions. In a retrospective study of over 400 women, women with >75% MD had a higher risk of benign hyperplasia on biopsy, compared to 0% MD (crude relative risk: 13.85 [2.65, 72.49]).[185] A prospective study of 676 women reported that women with self-reported BBD had 3.44% (0.40, 6.49) higher percentage MD than women without BBD.[186] The authors adjusted for age at mammogram, race/ethnicity and BMI. A population screening study in South Korea of over 3 million women reported an OR for high versus low MD of 1.59 (1.51, 1.67), where age, age at menarche, parity, breastfeeding, oral contraceptive use, age at menopause, and MHT use was adjusted for.[187] Given that MD is linked to the proliferation within the breast and BBD is also effectively the proliferation of abnormal breast epithelial tissue, a plausible explanation of the association is that the cellular mechanisms promoting MD also promote BBD. [186] An unexpected finding was the positive association between exercise and MD. Previous studies, including trials[188,189] and a systematic

review[190] have either reported an inverse association or no association between exercise or physical activity and MD.[191,192] After adjusting for BMI, however, the association between exercise and MD was attenuated. It is therefore possible that the association between exercise and MD observed in this analysis is due to residual confounding with BMI. Alternatively, it may be that the composition of the breast changes in response to exercise, irrespective of BMI and overall body composition.

There was no evidence of an association between MD and breastfeeding, oral contraceptive use, bilateral oophorectomy, and time since menopause in these data. Only 439 women were reported to have a bilateral oophorectomy before their natural menopause, and therefore, there was likely insufficient power to detect a difference in MD. Given that women are unlikely to have recently used oral contraceptives prior to the screen, if there was an effect on MD, as with past MHT use, it is likely that the impact on MD was no longer present at the time of the screen. Time since menopause has been reported to be associated with a reduction in MD, although there lacked evidence of such in the MWS.[7] Again, this could be due to limited power, as a significant proportion of women had missing time since menopause. Alternatively, this may be because the women had gone through menopause over 21 years before the digital mammogram, and therefore, changes in MD would not have been as evident.

The associations between all risk factors and MD were generally more apparent in those with higher BMI. However, this apparent difference was only statistically significant for BBD. In the subgroup analysis by categories of BMI, the associations seen between BBD and MD were greater in women with higher BMI than in women with lower BMI. This may be because women with higher BMI have relatively little dense tissue, and therefore, the relative changes in MD would be much more apparent than in women with lower BMI. Titus-Ernstoff et al. (2006), in a population of over 100,000 women, examined the association of several risk factors on MD by BMI.[5] In contrast to what I saw, they reported that the influence of some factors was attenuated with increasing BMI. In particular, the positive effect of age at menarche and age at first birth, and the inverse effect of parity on MD,

attenuated as BMI increased. The association between age at menarche and MD was no longer significant (p=0.25) in the highest BMI group (≥30kg/m^2). The authors suggested that hormonal or reproductive factors may have less of an effect in women with higher BMI, whose circulating hormone levels may already be relatively high due to the peripheral conversion in adipose tissue to oestrogens. However this is unlikely to be the case as higher BMI is associated with lower MD, whereas higher hormone levels are associated with higher MD.

In a population of less than 10,000 women, there were roughly 600 breast cancer cases, and therefore, I could not examine the associations between MD and breast cancer in the MWS. Although I could not formally examine the role of MD as a mediator, I examined whether this could be plausible. This was done by checking whether the effect of a given risk factor (MHT, alcohol, parity, age at first birth) on breast cancer risk might be fully explained through its effect on MD. This crude analysis suggested that MD could not fully explain associations between risk factors and breast cancer, and perhaps just partially mediated these associations. Therefore, it is likely that additional mechanisms may be involved in the relationship between these risk factors and the development of breast cancer. As I did not have actual percentage MD estimates and roughly calculated an average difference in MD based on the BIRADS, the approach taken here was very crude. Therefore, the results presented here cannot be taken as a proper examination of mediation. Interestingly though, a study of 3392 breast cancer cases and 8882 controls found that MD partially mediated the association between nulliparity, age at first birth, and current MHT use, and breast cancer risk in postmenopausal women, which is similar to the crude findings of the MWS analysis.[69]

The study has several strengths and limitations. One of the main strengths of this study was that it was able to address some of the limitations of the systematic review in Chapter 3. The inclusion of roughly 8000 women provided enough power to detect the effects of several breast cancer risk factors in postmenopausal women, although this was not possible for type of MHT, bilateral oophorectomy, breastfeeding, oral contraceptive use and time since menopause. By examining postmenopausal

women who were not current MHT users, I was able to avoid any masking or compounding effects due to the presence of exogenous hormones. Given the comprehensive questionnaires administered to the women in the MWS, I was able to examine a variety of lifestyle and reproductive risk factors. This also allowed me to control for important and relevant confounders. For example, it was decided to control for BMI in finer categories to minimise residual confounding. The large population meant that I did not lose power by doing so.

This study does, however, have several limitations. A main limitation was the generalisability of the study. Whilst I aimed to characterise the distribution of MD in the available dataset, it is likely that the observations are not representative of all postmenopausal women who attend routine breast screening in Oxford, the rest of the UK or populations with different demographics or ethnicities. This is because the current study population are older than the population of all postmenopausal women in England. As women in the MWS were recruited in the 1990's and digital mammography was not available until 2010s, all women were postmenopausal and nearing their 70's and hence represent an older demographic compared to other postmenopausal populations. Therefore, I was not able to examine the effects of the risk factors in premenopausal women or women who had recently gone through menopause, where the variation in MD may have been greater to detect significant differences in MD. It would have been of interest to study the associations in premenopausal women who attend screening and compare the results to postmenopausal women. The systematic review suggested that for some of the risk factors, the associations were greater in premenopausal women than postmenopausal women.

Misclassification of risk factors was potentially an important limitation. As information on risk factors, for most women, was collected several years before their screen, risk factor information may have changed by the time of their first digital mammogram. The average time between the last available questionnaire and the first screen was 2.9 years. This is unlikely to have been an issue for reproductive factors, but a possible limitation for BMI, MHT and lifestyle factors, where the effects of the risk factors

may be more transient. Current MHT users were set to “Missing” if their questionnaire was more than 4 years prior to their digital screen to avoid some degree of misclassification. I did not examine the associations in women with “Missing” MHT in a sensitivity analysis as it would not have represented a meaningful classification of MHT use. As current MHT users and women with missing status were removed in subsequent analyses, looking at “Missing” separately would not have affected the overall interpretation of the results.

Furthermore, there was a large proportion of missing data on time since menopause. This was crudely managed by including hysterectomy before natural menopause as a category. This is a general limitation of the MWS and other studies with women who had hysterectomies, as true age at menopause cannot be determined. The sensitivity analysis though, suggested that this did not change the overall interpretation of the results. 16% of the women in the MWS had a hysterectomy without a bilateral oophorectomy. 6% of women had a bilateral oophorectomy. The proportion of hysterectomies have been decreasing since the MWS first began, from around 20% around the recruitment of the MWS to 6% between 2011 and 2015.[193–195]

The BIRADS score represents four categories of increasing density, which I dichotomised into high and low MD, in order to have enough power for each category. The analysis would not have captured differences within a category because there is a large range of densities within the two categories. Therefore, although a risk factor may have influenced MD, it may not have been picked up in the current classification. A limitation of the data available was that I was not able to use continuous or more absolute measures of MD, such as percentage or dense area or dense volume. These would have allowed me to see smaller effects of the risk factors on MD. Furthermore, as the method of obtaining MD was used based on machine learning algorithms reading processed mammograms, that had not been used much previously, there may have been measurement error. Although the methods were validated.[126] Therefore, it would be useful to see how the models perform in a population of women with more varying ages.

The initial aim was to have data from 500,000 mammograms in the MWS. However, as the NHSBSP was paused in 2020 due to the COVID-19 pandemic, the collection of MD data was also delayed. As more MD data becomes available within the MWS, these associations can be investigated further with more power. The associations could also be examined by MHT use, to determine any effect modification by MHT use. There would additionally be more cases to investigate breast cancer as an outcome or conduct more formal mediation analyses. Understanding these associations can help explain the pathophysiology of breast cancer or how these risk factors, such as smoking, might be associated with the risk of breast cancer solely by affecting screening sensitivity. This could perhaps be explored by simulation and may be grounds for future work.

4.6 Conclusions

The results of the chapter shed light on the relationships between various risk factors on breast tissue composition. As MD is an important predictor for breast cancer risk, the present study has perhaps provided insight into identifying individuals with a higher risk of having their cancers missed at screening due to an increase in MD and, therefore, can facilitate targeted screening and preventative interventions. Given that some of the risk factors are modifiable, interventions to modify these or public awareness may improve breast cancer detection. Whilst the preliminary analysis suggested MD may only partially mediate the associations between some risk factors and breast cancer, larger studies with more breast cancer cases would be needed to investigate this more thoroughly.

Chapter 5. Menopausal hormone therapy use and measures of mammographic screening sensitivity in the Million Women Study

5.1 Introduction

The purpose of this chapter is to explore the effects of MHT use on the sensitivity of breast cancer screening. This will effectively explore the first three parts of the pathway, as shown in Figure 1 of Chapter 1. Rather than directly examining sensitivity, I shall use a proxy measure comparing the risk of breast cancer detected in the interval between screens to the risk of breast cancer detected at screening. This shall be explained in more detail in this chapter.

I shall begin the chapter with a background on population-based screening in different countries, discuss routes to breast cancer diagnosis in the UK, and then give an overview of MHT and its relationship with breast cancer risk, mammographic density, and screening sensitivity.

I will then present a background literature review and meta-analysis of the available evidence on MHT and its associations with screen-detected and interval cancers from screening centres around the world. In the meta-analysis, I will estimate the risk of interval versus screen-detected cancer in relation to MHT recency (status and time since last use), duration, and type (combined or oestrogen-only).

Finally, I shall present an analysis in the MWS on the associations of MHT use (recency, duration, and type) with risk of interval relative to screen-detected breast cancer in postmenopausal women attending routine breast screening in England.

5.2 Background

5.2.1 Population-based screening

Population-based breast cancer screening programmes have been introduced in several countries during the last four decades, including the UK.[71,72] During the 2022/2023 screening year (April 2022 – March 2023), 1.93 million women were screened in England, which resulted in 18,942 women being diagnosed with breast cancer.[196]

The aim of breast cancer screening programmes is to facilitate breast cancer diagnosis at an earlier stage when treatment offers a better prognosis.[73] A Cochrane review of eight randomised clinical

trials and a review of 11 randomised trials by the UK Independent Panel on Breast Cancer Screening reported that screening reduced breast cancer mortality by 10% to 20%, compared to no screening.[197,198]

Broadly speaking, women aged 45-70 are invited at two- to three-yearly intervals. Furthermore, whilst all programmes offer x-ray mammography, some countries, such as Austria and Belgium, additionally offer ultrasound scans. A summary of population-based screening programmes is shown in Table 19. The United Kingdom is the only country which screens women every three years. The remaining countries screen every two years. The USA has not been included in this table as it does not have a population-based breast screening programme, instead it has widespread insurance-based screening. Most health insurance plans allow for breast screens every 1 to 2 years for women from the age of 40.[199]

The UK NHS Breast Screening Programme (NHSBSP) was implemented in 1988.[73] Further details have been provided in Chapter 1. Every three years, women aged 50-70 are invited to attend for a screening mammogram. Screening at an earlier age is offered to the relatively small number of women who are at a higher risk of breast cancer because of their family history, personal history of breast cancer, gene-mutation status, or have a history of chest radiotherapy.

Table 19. Summary of some population-based screening programmes, ordered by screening frequency[71,72,200]

Country	Start	Test	Number of people reading mammogram	Age range for invitation of screening (years)	Screening frequency (years)
Sweden[201]	1986-1989	FM, DM	2	Varies by region	1.5–2
Australia[202,203]	1991	DM	2	50-74	2
Austria[204]	2014	DM, US	2	45–69	2
Belgium[205]	2000-2001	DM, US	2	50–69	2
Canada[206–208]	1988	FM, DM*	1	50-74	2
Denmark[209,210]	1991	DM	2	50–69	2
Finland[211,212]	1987	DM, US	2	50–69	2
France[213]	1989	FM, DM, CBE	2	50–74	2
Germany[214,215]	2002	DM	2	50–69	2
Iceland[216,217]	1987	DM	2	40–69	2
Ireland[218]	2000	DM	2	50–64	2
Italy[219,220]	1991	DM	2	50–69	2
Luxemburg[221]	1992	DM	2	50–69	2
Norway[222,223]	1995	DM	2	50–69	2
Portugal[224]	1990	DM	2	45–69	2
Slovenia[225]	2008	DM	2	50–69	2
Spain[226,227]	1990	DM	2	50–69	2
Switzerland[228,229]	1999	FM, DM	2	50–70	2
Holland[230–232]	1989	FM, DM	2	50–75	2
United Kingdom[73,233]	1988	DM	2	50–70	3

**Province dependent*
FM, Film mammography; DM, digital mammography; US, ultrasound; CBE, clinical breast examination

5.2.2 Routes to breast cancer diagnosis in the UK

Amongst women who attend breast cancer screening, breast cancer can either be diagnosed through screening or following the development of symptoms in the interval between screens, such as a breast lump. The women in the latter group usually present to a GP and are then referred to secondary care breast services.

Cancers detected through screening are termed “screen-detected”, and cancers diagnosed between screens, when symptoms may have developed, are termed “interval cancers”. The “interval period” is the time period between two consecutive screens. Interval cancers are sometimes referred to as “symptom-detected cancers”.

As discussed in Chapter 1, the effectiveness of a population-based breast cancer screening programme depends on the ability to correctly detect cancer from the mammogram, namely its sensitivity, and its ability to correctly identify those without cancer, its specificity. For women with dense breasts, the screening sensitivity of mammography is significantly reduced.[2] An increase in MD masks the radiographic evidence of a tumour, as both density and tumours appear white on an x-ray mammogram. Women with increased breast density may be at a greater risk of having breast cancer missed at screening and being diagnosed during the interval period following the development of symptoms.

Interval cancers can be either "missed interval" or "true interval" cancers.[234] "Missed interval" cancers are cancers that are present on the mammogram at screening but are not diagnosed, possibly due to high MD. They are usually classified as "missed" during a retrospective review of the previous screening mammogram after diagnosis of the interval cancer. "True interval" cancers are not detectable at screening but are present during the interval period.[234] These are likely to represent faster growing and more aggressive cancers, which tend to have a higher grade and negative hormone receptor status and, therefore, have a poorer prognosis.[235]

The NHS has adopted a three-point classification system for radiological review of interval cancers:[236,237]

1. Satisfactory: Normal or benign mammographic features
2. Satisfactory with learning points: Seen with hindsight, difficult to perceive, not obviously malignant
3. Unsatisfactory: Appearance is obviously malignant

Under the Duty of Candour regulations, NHS organisations must be open and transparent when mistakes are made and must inform patients soon as practically possible. They must also apologise.[236] In the context of the NHSBSP, women are informed that screening tests are not 100% accurate so are warned about the possibility of false negatives and false positives. At the time of a cancer diagnosis, women are informed that review of any previous screening mammograms will take place, if she has

previously been screened. If on review it is likely that there was a failure at the time of screening, either from the processes involved or interpretation, Duty of Candour regulations would need to be followed.

5.2.3 Menopausal hormone therapy and breast cancer risk

MHT is taken by women to manage the vasomotor and genitourinary symptoms associated with menopause, which are thought to be due to changes in levels of circulating oestrogen.[238] Symptoms include hot flushes, night sweats, sleep disorders, mood changes, and vaginal dryness, amongst others. There are an estimated 12 million current users of MHT in Western countries, with around 2 million users in the UK.[14,239] Approximately 15% of women aged 45-64 years in England are on MHT, with use rapidly rising.[240]

There are two main types of MHT: combined (containing oestrogen and progestogen) or oestrogen-only. Combined MHT can be taken either continuously (oestrogen and progestogen taken daily) or cyclically (oestrogen taken daily and progestogen taken during the last 14 days of the month).[241] Initial MHT preparations contained only oestrogen, as the oestrogen managed the symptoms of menopause. However, unopposed oestrogen therapy in postmenopausal women with a uterus was found to increase the risk of endometrial cancer, and the addition of a progestogen was found to confer protection against this.[242–245] Therefore, current recommendations are that women who have not had a hysterectomy should take combined MHT preparations to mitigate the increased risk of endometrial cancer. MHT can be administered orally, transdermally, vaginally, or sub-dermally.

Tibolone is a synthetic hormone used to treat the symptoms of menopause, such as hot flushes, low mood, and vaginal dryness or irritation.[246] Its mechanism of action is like that of combined MHT and has some androgenic activity.[247]

MHT is an independent risk factor for breast cancer.[14,163,248,249] The most robust evidence from the early 2000s came from the Women's Health Initiative (WHI), a randomised controlled trial, and the MWS, a prospective cohort study.[163,248] The WHI randomised over 16,000 postmenopausal women to

either receive combined MHT or placebo. Results based on a follow-up to a mean of 11 years, comparing the combined MHT group to the placebo group, yielded an HR for breast cancer of 1.25 (1.07, 1.46).[249] In another WHI trial of 10,739 women with a hysterectomy, there was no association between oestrogen-only MHT use and breast cancer risk, compared to the placebo group, after an average of 6.8 years of follow-up.[250] The MWS showed that current MHT users, of any preparation, had a 66% increase in the risk of developing breast cancer compared to never users (RR 1.66; 1.58, 1.75).[163] Past users did not have an increased risk of breast cancer compared to never users (1.01; 0.94, 1.09). The MWS results were included in an individual participant meta-analysis of observational studies by the Collaborative Group on Hormonal Factors in Breast Cancer, which included over 100,000 breast cancer cases. The authors found that ever MHT users had a 26% increased risk for developing breast cancer compared to never users (HR 1.26; 1.24, 1.28).[14] The risk was greater when comparing current combined MHT users to never users (2.01; 1.98, 2.06) than when comparing current oestrogen-only users to never users (1.37; 1.33, 1.41). Amongst current users of 5-14 years duration, those using continuous preparations had a relative risk of 2.30 (2.20, 2.40), and amongst those on cyclical preparations, the relative risk was 1.93 (1.84, 2.01).[14] The authors additionally found that the risk of breast cancer increased with the duration of MHT use and persisted for more than 10 years after cessation of use. For current tibolone users of 5-14 years duration, the authors reported an RR of 1.57 (1.43, 1.72).

MHT use has been shown to be primarily associated with luminal A-like (ER+/PR+) breast cancer, which is slow-growing, better differentiated, and less aggressive and has a more favourable prognosis than ER- breast cancers.[14,165,251–253] The Collaborative Group on Hormonal Factors in Breast Cancer found the relative risks for current MHT use (for both combined and oestrogen-only preparations), during years 5–14 were substantially greater for ER+ than ER- tumours but still significantly greater than 1 for ER- tumours (RR oestrogen-only for ER+: 1.45 [1.38,1.53]; ER-: 1.25 [1.13, 1.38]; RR combined for ER+: 2.44 [2.35, 2.54]; ER-: 1.42 [1.30, 1.55]).[14]

5.2.4 Relationship between MHT, mammographic density and screening sensitivity

There is considerable evidence to show that MHT is associated with MD.[60,166] This is likely because exogenous hormones will stimulate the proliferation of glandular tissue within the breast and therefore increase MD. In the Women's Health Initiative trial, 435 post-menopausal women were randomly assigned to either receiving oestrogen-only MHT or placebo. After one year of follow up, mean percentage MD increased by 1.6% (0.8, 2.4%) among women receiving oestrogen-only MHT and decreased by 1% (-1.7, -0.4%) in the placebo group.[61] The WHI also randomised 413 post-menopausal women to combined MHT or placebo. After one year of follow up, mean percentage MD increased by 6.0% (4.6, 7.5%) among women receiving combined MHT and decreased by 0.9% (-1.5, -0.2%) in the placebo group.[62]

A systematic literature review of twenty-two studies, including six trials and 16 observational studies, published in 2020 reported that MHT use has been consistently found to increase MD.[60] The review reported that combined MHT was associated with higher MD than oestrogen-only MHT. Furthermore, women on continuous combined MHT had higher MD than women on cyclical MHT. There is a lack of data on how long the impact of MHT on MD lasts and whether the duration of MHT use is associated with increasing MD.

In Chapter 4, I examined the associations between MHT use and MD in the MWS. Like the previous evidence, current MHT use was associated with high MD, compared to never users: OR 1.54 (1.11, 2.15). No associations were seen in past users or by MHT type, suggesting the effects on MD are transient. However, given that mean time since menopause was 21.9 years prior to their first digital mammography, past users were likely taking MHT over a decade before digital mammography (as used in Chapter 4). Therefore, I could not reliably assess effects that might have persisted up 10 years after stopping, only those which would have persisted for much longer.

Given the known association of MHT use with increased MD, it might be expected that screening sensitivity would be reduced in those taking MHT. A review of eight studies published in 2001,[254] and

two further studies [255,256] published subsequently, have reported that current MHT use reduces the sensitivity of breast screening, with reductions in sensitivity ranging from approximately 2% to 22% between current and never MHT users. One of these studies included a subset of the MWS.[256] From June 1996 to March 1998, 122,355 women from the MWS were selected for a special study on the effect of MHT on mammographic screening sensitivity and specificity. There were 629 screen-detected breast cancer cases and 97 cases that were screen-negative but had breast cancer diagnoses within 12 months of the screen. Sensitivity was 83.0% (95% CI 77.4, 87.6) in current MHT users, 84.7% (73.9, 91.6) in past users, and 92.1% (87.6, 95.0) in never users.

Whilst it is evident that MHT use reduces mammographic screening sensitivity, less is known about how MHT recency, duration, and type affect sensitivity. Understanding this may help inform screening practices and help guide how to advise women of the risks of MHT use. It may also be beneficial to warn women of the limitations of screening if on MHT so that they can make informed decisions about MHT use.

The MWS only has screening sensitivity data for a small proportion of breast cancers.[256] It does, however, have information on the route to diagnosis (screen-detected or interval) for approximately 27,000 breast cancers, alongside detailed information on MHT recency, type, and timing of use. The extent to which the associations of MHT use with the risk of interval cancer differ from those with risk of screen-detected cancer can be used as a proxy measure of the effect of MHT on screening sensitivity. I examined the risk of interval versus screen-detected breast cancer by MHT recency, duration, and type. As mentioned earlier, as MHT use is associated with slower growing and less aggressive breast cancer, any association between MHT use and interval cancer risk is likely to reflect an association with missed cancers rather than true interval cancers.[14,165,251–253] Therefore, an examination of interval versus screen-detected breast cancer risk by MHT use can act as a surrogate for screening sensitivity.

Before presenting the analysis, I shall summarise the published evidence on interval versus screen-detected breast cancer by MHT use.

5.3 Literature review and meta-analysis

5.3.1 Aims

The aim of this section of the chapter was to summarise the published evidence on interval versus screen-detected breast cancer by patterns of MHT use (status, time since last use, type, and duration of use), in a meta-analysis.

5.3.2 Methods

5.3.2.1 Literature searches

A PubMed search was conducted on 26th May 2022 to identify all published studies examining the effect of MHT on the risk of screen-detected and interval cancers since the inception of the database. Studies which reported on associations of MHT use with the risk of screen-detected and interval cancers were included. Full-text peer-reviewed scientific papers, including cohort, case-case and case-control studies, were included. Scientific papers not reporting an age-adjusted relative risk for the association of MHT and screen and interval cancers were excluded, as were reviews, case reports, and case series. In addition, the references of included studies were checked for further relevant published papers for inclusion.

The following search was entered into PubMed:

"((screen[Title/Abstract]) AND (interval OR symptom[Title/Abstract]) AND (breast cancers OR breast cancer[Title/Abstract]) AND ((hormone replacement therapy[Title/Abstract]) OR (menopausal hormone therapy[Title/Abstract]) OR (hormone therapy[Title/Abstract])))"

5.3.2.2 Harmonisation of results

The types of estimates presented in the included studies differed according to the endpoint used in the analysis (Table 20).

Table 20. Summary of studies by type of analysis and endpoint used

Type of analysis	Endpoints and presentation of results
Logistic regression (case-case)	OR for risk of interval vs screen-detected cancer, comparing MHT users to never users
Cox regression (longitudinal)	HR of interval cancer, comparing MHT users to never users HR for risk of screen-detected cancer, comparing MHT users to never users
Logistic regression (case-control)	OR of interval cancer, comparing MHT users to never users OR for risk of screen-detected cancer, comparing MHT users to never users

OR, odds ratio; HR, hazard ratio; MHT, menopausal hormone therapy

The relative risks, with their associated 95% confidence intervals, were extracted. Where studies presented the results from two types of analysis, the case-case results were used. As the primary aim was to determine the association between MHT use and the risk of screen-detected versus interval cancers, where studies reported separate relative risks (for screen-detected and for interval cancers), a single estimate was generated by dividing the relative risk (RR) for interval cancers by that for screen-detected cancers. A corresponding confidence interval was estimated using the principle that the variance of the difference (on the log scale) is the sum of the variances. Where a study reported results separately by MHT type (combined and oestrogen-only), or by interval cancer type (missed or true), results were combined using the method detailed by Berrington and Cox.[257]

For studies reporting duration of MHT use and time since last MHT use, by ordinal categories of time, a trend through the extracted measures of association was fitted, using generalised least squares, as described by Greenland and Longnecker.[134,135]

5.3.2.3 Meta-analysis

To summarise the evidence for an association of MHT use with the risk of interval as opposed to screen-detected breast cancer, separate meta-analyses were conducted for:

- ever MHT users versus never users;

- current or recent MHT users versus never users;
- past users versus never users;
- time since last MHT use;
- current combined MHT users versus never users and current oestrogen-only users versus never users; and
- duration of use.

For the purposes of this meta-analysis, recent use was defined as MHT use within six months prior to breast cancer diagnosis.

Pooled RRs were generated using a fixed-effects meta-analysis, using the inverse variance method of weighting. Heterogeneity was assessed statistically using the I^2 test. Analyses were conducted in Stata 17.[127] Figures were generated in R-Studio using the Jasper package.[129]

5.3.3 Results

The PubMed literature search yielded 44 studies. After applying the exclusion and inclusion criteria, 15 studies were identified for potential inclusion in the analysis, which examined the risk of screen-detected and interval breast cancers according to MHT use. Two of these studies reported results on the same population of patients.[255,258] The most recent of these two publications was used, leaving 14 studies for inclusion in this systematic review. All studies were conducted outside the UK. Not all studies restricted their populations to postmenopausal women. Five studies did restrict to peri/postmenopausal women. In some studies, menopausal status was not specified (N = 3). 11 studies recorded MHT use around the time of screening, whilst the remaining recorded use around the time of breast cancer diagnosis. Seven studies adjusted for MD. Nine studies had a case-case design,[234,258–265] one study was a screening cohort,[266] one was case-control,[267] and three studies utilised both case-case analyses and longitudinal cohort analyses[268–270] (Table 21). The screening programmes in most studies (N=13) offered biennial mammograms.

5.3.3.1 Status and recency of MHT use

The forest plots for the fixed-effects meta-analyses of the risk of interval as opposed to screen-detected breast cancer, for (i) ever MHT users versus never users, (ii) current or recent MHT users versus never users, and (iii) past users versus never users are shown in Figure 20.

Being an ever user as opposed to a never MHT user was associated with an increased risk of interval versus screen-detected cancer (pooled RR 1.34; 95% CI 1.19, 1.51).[260,267,269]

Being a current or recent user was also associated with an increased risk compared to non-users: 1.69 (1.55, 1.84). Five studies additionally examined the effect of past MHT use.[234,261–263,266] Meta-analysis of these studies showed a statistically non-significant increased risk of interval cancers compared to screen-detected, among past MHT users compared to never users: pooled RR 1.11 (0.97, 1.28).

Chiarelli et al. (2006) and Chiarelli et al. (2015) additionally examined time since last MHT use in women participating in the Ontario Breast Screening Programme, during different time periods (Figure 20).[234,262] In the pooled analysis, there was no evidence of an association between time since last MHT use and the risk of interval versus screened detected cancers: pooled RR for an increment of five years since last MHT use 0.99 (0.63, 1.55).

The I^2 test showed no statistical evidence of heterogeneity across studies in the pooled analysis of ever MHT users, past MHT users, and time since last MHT use. There was evidence of heterogeneity across studies looking at current or recent MHT use (I^2=65.4%, p=0.001).

5.3.3.2 Type of MHT preparation

The different types of MHT preparation in current users were examined in three studies (Figure 21).[234,264,266] The pooled relative risks, were higher for current combined MHT users vs never users than current oestrogen-only vs never users, although were not statistically significant: pooled RRs 1.72 (1.47, 2.02) and 1.60 (1.33, 1.92), respectively. There was evidence of statistical heterogeneity across the studies examining combined MHT, but not amongst those examining oestrogen-only formulations: I^2=62.4%, p=0.070 and I^2=0.0%, p=0.57, respectively.

5.3.3.3 Duration of MHT use

Three studies examined the risk of interval versus screen-detected breast cancer by duration of MHT use.[234,262,269] The meta-analysis of the studies is displayed in Figure 22. A five-year increase in duration of MHT use was associated with a 27% increased risk of developing an interval cancer as opposed to a screen-detected cancer: pooled RR 1.27 (1.13, 1.43). The I^2 test showed no statistical evidence of heterogeneity across studies in the pooled analysis of duration of MHT use (I^2=0.0%, p=0.70).

Table 21. Summary of the literature on MHT use and screen-detected and interval breast cancer

Study	Location & frequency of screen	Analysis method	Number of participants	Menopausal status of participants (pre/post/both/not specified)	Ascertainment of MHT use	Number of breast cancer cases	Adjustments
Blanch 2014[268]	Spain (2-yearly)	Cox regression and case-case analysis	645,764	Both	Questionnaire at screening	Screen: 1428 Interval: 877	Age, menopausal status, FH, BBD, mammographic density, type of mammogram, initial or successive participation, year of last mammogram
Chiarelli 2006[234]	Canada (2-yearly)	Case-case analysis	431,480	Both	Questionnaire after diagnosis asking about MHT use at time of screening	Screen: 450 Interval: 375	Age, region and year of last screen matched Adjusted for time since screening, BBD, BMI, smoking, FH, parity, education, age at menarche, menopausal status, mammographic density
Chiarelli 2015[262]	Canada (2-yearly)	Case-case analysis		Postmenopausal	Questionnaire after diagnosis asking about MHT use at time of screening	Screen: 955 Interval: 286	Age, BMI, smoking, time since last screen, mammographic density
Crane 2002[259]	Australia (2-yearly)	Case-case analysis		Peri/Postmenopausal	Questionnaire at screening	Screen: 628 Interval: 270	Matched by age and year of diagnosis Adjusted for BBD, stage, differentiation

Study	Location & frequency of screen	Analysis method	Number of participants	Menopausal status of participants (pre/post/both/not specified)	Ascertainment of MHT use	Number of breast cancer cases	Adjustments
Domingo 2010[260]	Spain (2-yearly)	Case-case analysis		Both	Questionnaire at screening	Screen: 115 Interval: 115	Age, TNM, first or successive screening
Gaudet 2021[266]	USA (2-yearly)	Cox regression	77,206 14.8 years follow-up	Both	Questionnaire at screening	Screen: 2711 Interval: 1281	Stratified on age Adjusted for race, education, age at menarche, OC use, parity, age at first birth, age at menopause, reason for menopause, BMI, weight change since age 18, height, physical activity, sitting time, FH, BBD, alcohol, smoking, menopausal status
Hofvind 2006[269]	Norway (2-yearly)	Cox regression and case-case analysis	296,651	Both (premenopausal not on MHT)	Questionnaire with screening invitation	Screen: 1512 Interval: 814	Age, calendar year, age at menarche, education, menopausal status, age at menopause, parity, OC use, alcohol, FH
Holm 2015[261]	Sweden (every 18-24 months)	Case-case analysis		Not specified	From questionnaire in 2009 asking about MHT use in year of diagnosis	Screen: 1307 Interval: 550	Age, BMI, mammographic density
Kavanagh 2008[258]	Australia (2-yearly)	Case-case analysis		Not specified	Questionnaire at screening	Screen: 1025 Interval: 369	Age, symptoms, family history, previous mammography, mammographic density

Study	Location & frequency of screen	Analysis method	Number of participants	Menopausal status of participants (pre/post/both/not specified)	Ascertainment of MHT use	Number of breast cancer cases	Adjustments
Njor 2009[263]	Denmark (2-yearly)	Case-case analysis		Postmenopausal	Linkage through drug prescription register at time of screening	Screen: 361 Interval: 68	Age, mammographic density, screen number
Njor 2011[264]	Denmark (2-yearly)	Case-case analysis		Postmenopausal	Linkage through drug prescription register at time of screening	Screen: 1267 Interval: 400	Age, mammographic density, screen number
Seradour 1999[270]	France (3-yearly)	Cox regression and case-case analysis	41,062	Not specified	Questionnaire at screening	Screen: 395 Interval: 47	Stratified on age
Sprague 2015[267]	USA	Case-control		Postmenopausal	Questionnaire after diagnosis asking about MHT use within 1 year of diagnosis	Screen: 8372 Interval: 7276 Controls: 17,602	Age, state of residence, year of diagnosis, education, FH, age at menarche, age at first birth, parity, age at menopause, alcohol, pre and postmenopausal BMI
Strand 2016[265]	Sweden (every 18-24 months)	Case-case analysis		Both	From questionnaire in 2009 asking about MHT use in year of diagnosis	Screen: 958 Interval: 373	Age, BMI, mammographic density

BBD, benign breast disease; BMI, body mass index; FH, family history of breast cancer; OC, oral contraceptive use

Figure 20. Meta-analysis of the relative risk for interval vs screen-detected breast cancer by MHT status and recency of use

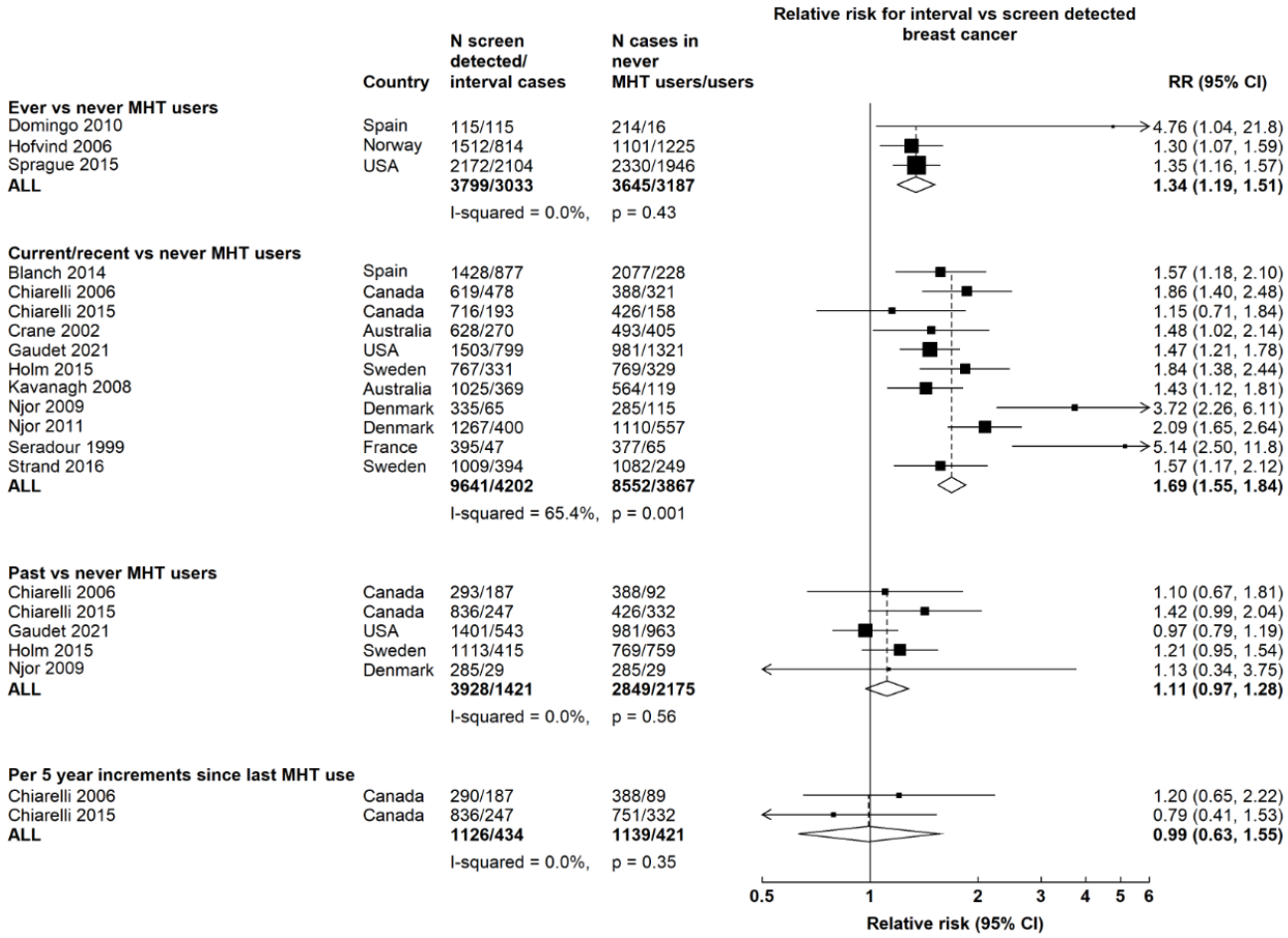


Figure 21. Meta-analysis of the relative risk for interval vs screen-detected breast cancer by current MHT type

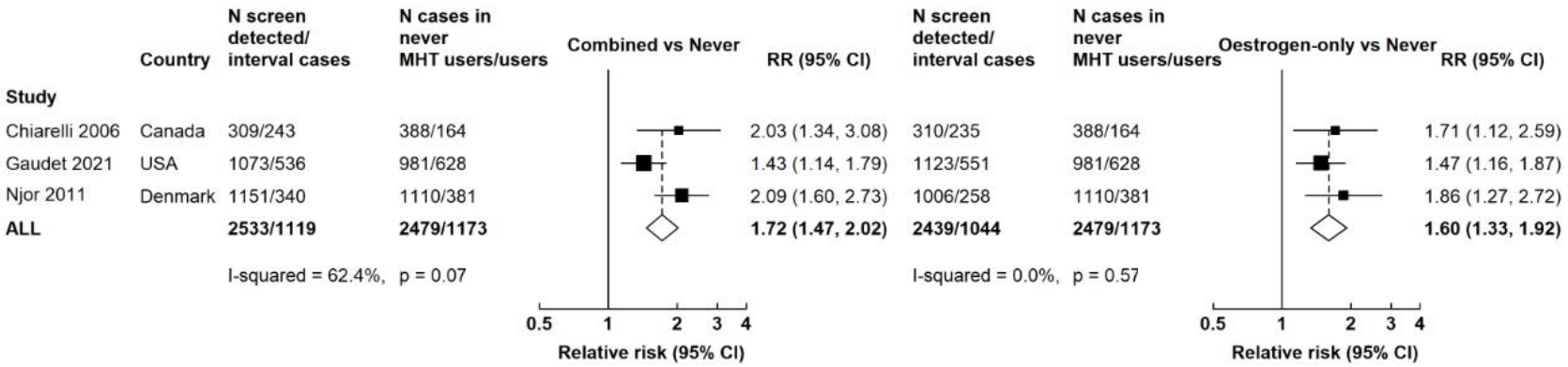
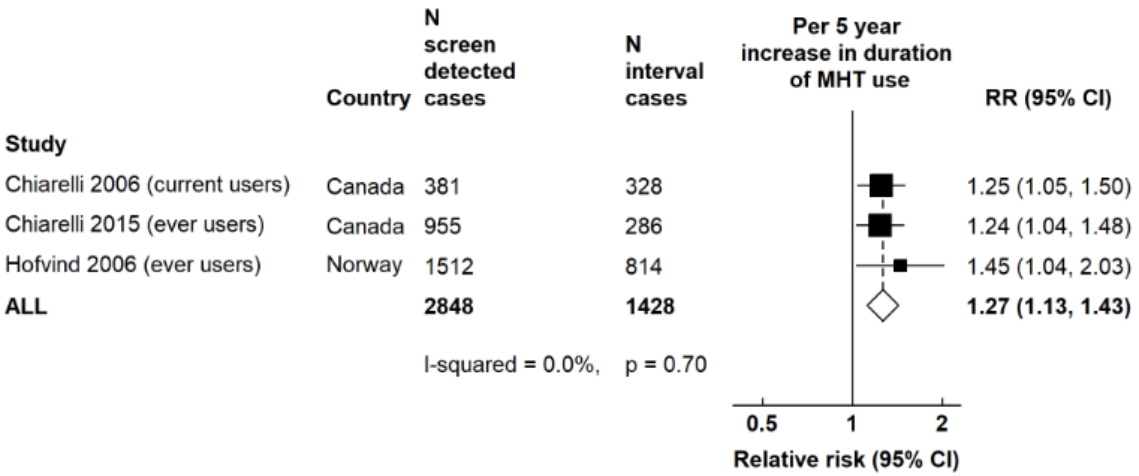


Figure 22. Meta-analysis of the relative risk for interval vs screen-detected breast cancer by duration of MHT (relative risk is for an increment of 5 years of MHT use)



5.3.4 Discussion

The results of this review suggest that current or recent MHT users have around a 69% increased risk of interval versus screen-detected breast cancers compared to never users, however there is no residual effect after women cease MHT use. The relative risk increased with duration of use. There was no clear evidence that the association between MHT use and interval versus screen-detected breast cancer varied by MHT type.

The results on current MHT use in this review agree with current understanding on how MHT affects mammographic density and screening sensitivity.[60,254] One possible mechanism to explain this could be increased breast proliferation of the glandular tissue in women on MHT.[65,66] MHT activates the oestrogen receptors within the breast tissue to stimulate cell growth.[271] This increase in MD, as shown in the analysis in Chapter 4, reduces screening sensitivity and hence, breast cancers are diagnosed outside screening during the interval between two screens when they become symptomatic. Duration of use was also associated with an increased relative risk. This is likely because duration indicates the overall exposure to MHT, and therefore, more exposure to exogenous hormones will increase breast tissue proliferation. As mentioned previously, it is unlikely that MHT causes more aggressive cancers, as it is associated with slower growing cancers, and therefore the results are likely to represent delay in diagnosis rather than a new cancer. [14,165,251–253]

The lack of an increased risk of interval versus screen-detected cancers between past MHT users and never users suggests the effect of MHT on MD, and hence screening sensitivity, may be relatively short-lived. These results are in keeping with the MWS analysis in Chapter 4, in which past MHT use was not associated with high MD. However, it is challenging to interpret these results with certainty, as past use or duration of use may be correlated with recency of use. Hence, when examining the effects of past MHT use or duration of use on screening sensitivity, it is essential to consider the recency of use in these women. The studies also had quite low power to detect a small residual effect in past users.

The review showed some indication that combined MHT users had a higher risk of interval cancers versus screen-detected cancers, compared to oestrogen-only users. However this was not statistically significant, possibly due to lack of power. Previous studies have suggested that combined users have higher MD than oestrogen-only users.[60,163]

A strength of this review is that I was able to summarise the totality of the evidence on interval and screen-detected cancers and MHT use and synthesise the results in a pooled analysis. A strength of many of the studies is that they mostly adjusted for important and relevant potential confounders in their analyses, such as age, BMI, and even MD.

A limitation of the meta-analyses is that there were few studies, and most were based on small populations, and as such may be underpowered to detect a difference in the risk of screen-detected and interval cancers by MHT type, duration of use, and time since last use. Only three studies reported results by current MHT type. Five studies examined past use of MHT, three studies looked at the duration of MHT, and two studies looked at time since last MHT use. The results seen in the studies could be due to small numbers of cases within each time category; therefore, again, the studies lacked sufficient power to reliably assess differences by pattern of use, if they exist. Robust data, with sufficient power, on past use of MHT and time since last use is pertinent to understanding how long the effects of MHT on screening last after cessation of use.

Some of the studies recorded MHT use at diagnosis rather than at screening. As an interval cancer may be missed at screening, it is important to know the MHT use at screening, when the mammogram was taken, and not at diagnosis. This is to ensure the correct classification of exposure status. It is plausible that some women who had an interval cancer may have started MHT use after their screen and before their cancer was diagnosed. If this potential misclassification of the exposure was random, it would have attenuated the results. Recall bias could have caused further misclassification of exposure as MHT information was taken from questionnaire data in all but one of the studies. Whilst classification of never/past/current/ever is unlikely to be mistaken, duration or time since last use may

be forgotten. One study used linkage to prescription records to ascertain information on MHT use.[263] Whilst this may mitigate recall bias, a record of a prescription does not necessarily mean the medication was taken. It would have been interesting to stratify the analysis by method of MHT ascertainment and study type, to determine whether the results would have differed. Given the relatively few studies though, these analyses would likely have been underpowered.

Furthermore, some studies included premenopausal women. This may be an issue for many reasons. If these women do take MHT, they may use it for indications other than menopausal symptoms, which could change their background breast cancer risk, or could confound the associations between MHT and MD or screening sensitivity. Premenopausal women may use MHT for premature ovarian insufficiency, which is where women under 40 years old experience a decline in ovarian function, or for prevention of osteoporosis. It would have been difficult to adjust for these factors in these analyses. To ensure internal validity and avoid confounding by indication, future analyses of MHT use should be restricted to postmenopausal women, who are eligible to take MHT for menopausal symptoms.

Most of the studies in the review had screening every two years. In the NHS Breast Screening Programme women are invited every three years for a mammogram. Given the greater length of time between screens, women in the UK have more opportunity to present with an interval cancer than women living in a country where biennial screening is available. As none of the studies in the literature review were conducted in UK women and only one study included women who were offered screening every three years, the results of the meta-analysis are not generalisable to the UK. Therefore, an analysis of UK data, with three-yearly screens, is necessary.

In the next part of this chapter, I shall present an analysis in the MWS where I can address several of the limitations discussed here. Given the large statistical power, I will be able to explore these associations in finer detail by duration, recency, and type. Furthermore, I will be able to restrict the

analysis to postmenopausal women and MHT use at the time of screen rather than diagnosis to ensure the appropriate classification of my exposure.

5.3.5 Conclusions

The published literature suggests that MHT users have an increased risk of both screen-detected and interval cancers, compared to never users. However, MHT users have an increased of having breast cancer diagnosed during the interval period than at screening, compared to never users. Given the limitations of the present review, a proper assessment of the impact of type, duration and recency of use on this association cannot be made due to the lack of evidence. This is particularly important as the efficacy of the screening programme in MHT users is likely to depend critically on these factors. Additional analyses in the MWS to examine the results in a UK population with greater statistical power, and with the ability to understand in finer detail how the risks differ by duration, recency and type of MHT are warranted. Such results may inform whether the breast screening programme needs to become more personalised for MHT users and whether advice on ceasing MHT use before a screen may be beneficial.

5.4 Analysis in the Million Women Study

5.4.1 Aims

The aim of this MWS analysis was to determine the relative risk of interval breast cancers, compared to screen-detected cancers, as a proxy of screening sensitivity,

- by recency of MHT use
- by duration of MHT use
- by type of MHT

among postmenopausal women attending routine breast cancer screening in England.

5.4.2 Methods

5.4.2.1 Study population and linkage

Details about the MWS design, population, and NHS breast screening programme, are given in Chapter 2. Briefly, the MWS is an open-ended prospective study of 1.3 million women in England and Scotland,

which recruited women from 1996 to 2001. In 2013, the MWS cohort was linked to the breast screening programme records on breast screening episodes for women in England that had started by 31st December 2012. A screening episode (or screening round) is the time interval between consecutive breast cancer screening invitations (usually 3 years). Each MWS participant was linked via participant identification numbers, matched to their name, date of birth, and NHS number on screening records. The database currently only has records from women recruited and living in England. Linkage provides information on the dates women were invited for their routine screen and whether they attended or not.

5.4.2.2 Exposure

The main exposure of interest was MHT use at screening. The classifications of MHT use were recency of MHT use (never, current, <5, 5-9, 10+ years since last use), duration of use (never, <5, 5-9, 10+ years), and type of MHT (oestrogen-only, combined, tibolone, other).

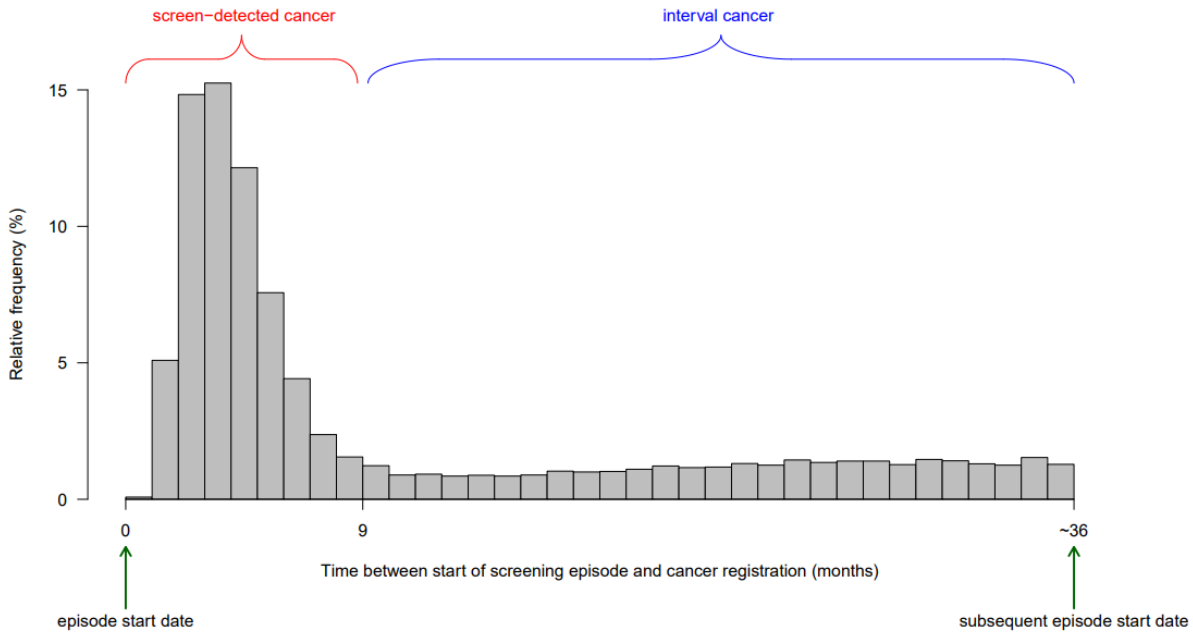
As individual MHT use changes over time, MHT status provided at the recruitment questionnaire was updated based on information provided at the resurveys in median year 2001 and 2008. For the purposes of this analysis, MHT use at each screen was taken from a woman's MHT status as reported at the last available survey prior to that screen. The exact rules used to update MHT use are given in Appendix 7. The main conventions were that it was assumed that women who were never or past users at the first resurvey did not subsequently start, or recommence, use of MHT, respectively. In addition, the status of women who were current users at the last survey prior to the screen was set to unknown to avoid misclassification of MHT use and dilution of effect.

5.4.2.3 Outcomes

The two main outcomes of interest were invasive screen-detected and interval breast cancers (ICD10: C50) at any time between the start and end of an attended screening episode. Date of diagnosis was taken from cancer registry data. Cancer registry data provide information on mode of detection but is only available for roughly 62% of cases. Therefore, the definitions of screen-detected and interval cancers were determined using the empirical distribution of time from the start of an attended

screening episode to cancer registration. As the distribution tails off around nine months, screen-detected cancers were defined as those registered within nine months of the start of a screening episode. Interval cancers were defined as those registered after the nine months, but prior to the next screening episode (Figure 23). The distribution of screen-detected and interval cancers based on this definition showed good agreement with that based on mode of detection as provided by cancer registry data, where available (percentage agreement = 96%).

Figure 23. Definition of screen-detected and interval breast cancer



5.4.2.4 Statistical analysis

This analysis was restricted to postmenopausal women, defined as those who had a natural menopause or bilateral oophorectomy at baseline, or were aged at least 55. This age cut-off was used because 96% of women with known age at menopause reported being postmenopausal at 55 years.[272] Women were excluded if information on MHT use was missing or if they had had a previous cancer, as this may affect their health-seeking behaviour, or put them at a greater risk of breast cancer if they had a cancer syndrome. Women in Scotland were excluded as screening data was only available for women in England.

Women were followed from the date of recruitment if they reported having had a natural menopause, a bilateral oophorectomy, or were aged 55 years or above at recruitment. Those who did not fall into these categories were followed from their 55th birthday. Follow-up ended on the earliest of the date of first cancer registration, date of death, date of cessation of NHS registration, the latest date that a woman's screening status could be ascertained (three years after the start of the last routine screening episode on record), or 31st December 2012.

Follow-up time was subdivided into that during which women were at risk of a screen-detected cancer (namely up to nine months after receiving the invitation for a routine screen) and that during which women were at risk of an interval cancer (between nine months after screening and the end of that screening episode). Women who did not attend for a routine screen exited analyses at the start of the screening episode and re-entered at the start of the next attended screening episode.

Cox regression models were used to estimate the association of MHT with breast cancer, separately within each risk period (screen-detected and interval). Results are presented as hazard ratios (HRs) and 95% confidence intervals (95% CIs), by time since last MHT use (recency), duration of use, and type last used, at the start of the latest screening episode. The rationale for presenting the HR separately by the risk periods was to understand the underlying patterns of the associations between the interval and screen-detected cases by MHT use. Analyses were stratified by year of birth (≤1930, 1931, …, 1950+) and year of recruitment (1996, 1997, …, 2001+), and adjusted for year of screen, age at menarche (<12, 12-13, 14+), age at menopause (<45, 46-48, 49-51, 52+, hysterectomy before natural menopause), weekly units of alcohol (<3, 3-7, 8+), current smoking (yes/no), oral contraceptive use (ever/never), parity (0, 1-2, 3+), benign breast disease (yes/no), first degree relative with history of breast cancer (yes/no), BMI (<25, 25-30, 30+), region of recruitment, and Townsend deprivation quintiles. Likelihood ratio tests were used to assess heterogeneity in the HRs (for MHT type) or trends in HRs (recency and duration), separately for screen-detected and interval breast cancers.

After understanding the underlying patterns of the associations with interval cancers and with screen-detected cancers, a case-case analysis was used to compare the odds of interval cancers to that of screen-detected cancers. This analysis provided a proxy measure for screening sensitivity by quantifying the relative difference in associations between interval and screen-detected cancers. Multivariable logistic regression was used to estimate odds ratios (ORs) comparing interval to screen-detected breast cancers, and their corresponding 95% confidence intervals, by time since last MHT use (recency), duration of use, and type last used, at the start of the latest screening episode. Analyses were adjusted for the same potential confounders as in the Cox regression. The analysis of duration MHT use in the case-case analysis was further adjusted by recency to examine the effect of adjusting for recency on duration. Likelihood ratio tests were used to assess heterogeneity in the ORs (for MHT type) or trends in ORs (recency and duration). This was to determine if MHT use was associated with a statistically significant increased risk of interval breast cancer relative to screen-detected breast cancer.

To understand the associations in finer detail, further analyses were conducted examining HRs by duration (<5, 5+ years) of use and by MHT type (oestrogen-only, combined, other), within categories of recency of use (current, past <5 years, past 5-9 years, 10+ years). This was to determine whether duration and type of MHT use were independently associated with risk given recency of MHT use. Furthermore, the case-case analyses were also examined by age at screen, <60 years old and 60+ years old, to determine if the associations seen differed by age at screen, as MD declines with age. Likelihood ratio tests were used to determine if there was an interaction between recency and duration, recency and type and also age with all classifications of MHT (recency, duration and type).

In general, missing values for adjustment variables were included as a separate category. Age at menopause was missing for 517,763 (47.3%) women. This, in part, was because 174,774 (16.0%) women had a hysterectomy prior to their natural menopause. In main analyses, age at menopause was adjusted for using a variable with separate categories for women with hysterectomy prior to

natural menopause and for women who did not provide an appropriate age ("Missing"). This allowed for a crude adjustment of age at menopause. A sensitivity analysis explored the likely impact of this lack of complete adjustment for age at menopause. In the sensitivity analysis, the main analysis was repeated, adjusting for and not adjusting for age at menopause among the women with known age at menopause. Women with a hysterectomy were excluded from the sensitivity analysis. This was to determine the effect of properly adjusting for age at menopause on the overall interpretation of the results.

Analyses were conducted in Stata 17.[127] Figures were generated in R-Studio using the Jasper package.[129]

5.4.3 Results

1,095,762 women were included in the analysis. Women were followed up for a mean of 7.4 years (median 8.0), during which 27,564 invasive breast cancer cases were diagnosed. A total of 15,902 breast cancers were detected at routine screening and 11,662 were diagnosed during the interval between screens. Table 22 displays the baseline characteristics of the women by recruitment MHT use. Age at menopause was, on average, 2.7 years earlier among current users than never users (45.9 vs 48.6 years). Current MHT users were more likely to have ever used oral contraception than never users (66.2% vs 52.2%).

Table 22. Baseline characteristics of women in the Million Women Study by recruitment MHT use

Characteristics at baseline	Never	Past	Current	All women
Number of women	538,156	184,686	367,679	N=1,090,521
Age at recruitment	56.3 (4.8)	56.4 (4.3)	55.2 (4.0)	55.9 (4.5)
Age at menopause	48.6 (5.2)	47.1 (6.0)	45.9 (6.2)	47.5 (5.8)
Body mass index (kg/m^2)	26.4 (4.9)	26.6 (4.6)	25.8 (4.3)	26.2 (4.7)
Current smokers	96,557 (17.9%)	37,900 (20.5%)	76,908 (20.9%)	211,365 (19.4%)
Alcohol drinks per week	3.8 (5.1)	4.1 (5.3)	4.5 (5.5)	4.1 (5.3)
Parous	475,956 (88.4%)	167,997 (91.0%)	330,619 (89.9%)	974,572 (89.4%)
Number of full-term pregnancies	2.2 (1.3)	2.2 (1.2)	2.2 (1.2)	2.2 (1.2)
First-degree relative with breast cancer	53,257 (9.9%)	17,179 (9.3%)	28,890 (7.9%)	99,326 (9.1%)
Ever oral contraceptive user	281,025 (52.2%)	116,110 (62.9%)	245,379 (66.7%)	642,514 (58.9%)
Ever breastfed among parous women	254,009 (47.2%)	91,561 (49.6%)	177,418 (48.3%)	522,988 (48.0%)
Benign breast disease	58,183 (10.8%)	24,849 (13.5%)	47,552 (12.9%)	130,584 (12.0%)

Results shown are N (%) or mean (SD)

The risks of both interval cancer and screen-detected cancers were elevated in all MHT users, but in general, HRs were greater for interval cancers than screen-detected cancers (Figure 24).

The HRs of screen-detected cancers were not elevated among any categories of past users, compared to never users. However, the HRs of interval cancers were elevated among past users, compared to never users, but decreased as recency of use increased ($p<0.001$).

The ORs for interval vs screen-detected cancers allowed for a comparison between the HRs, which acted as a proxy for screening sensitivity. Current users had a 26% increased risk of an interval versus screen-detected cancer, compared to never users (OR 1.26, 95% CI 1.18, 1.35). The OR of interval

versus screen-detected breast cancer appeared to decrease as time since last use increased. Nevertheless, the OR remained elevated, even after 10+ years of cessation (p=0.003).

The trend in the case-case analysis, comparing the odds of interval to screen-detected cancer, for duration of use, compared to never users, was of borderline significance (p=0.063). After adjusting for recency of use, the p-value was 0.54. This suggested that the possible effect of duration on screening sensitivity can be explained by recency of use. In the case-case analysis MHT type, overall, did not affect the risk of interval versus screen-detected breast cancers (p=0.72).

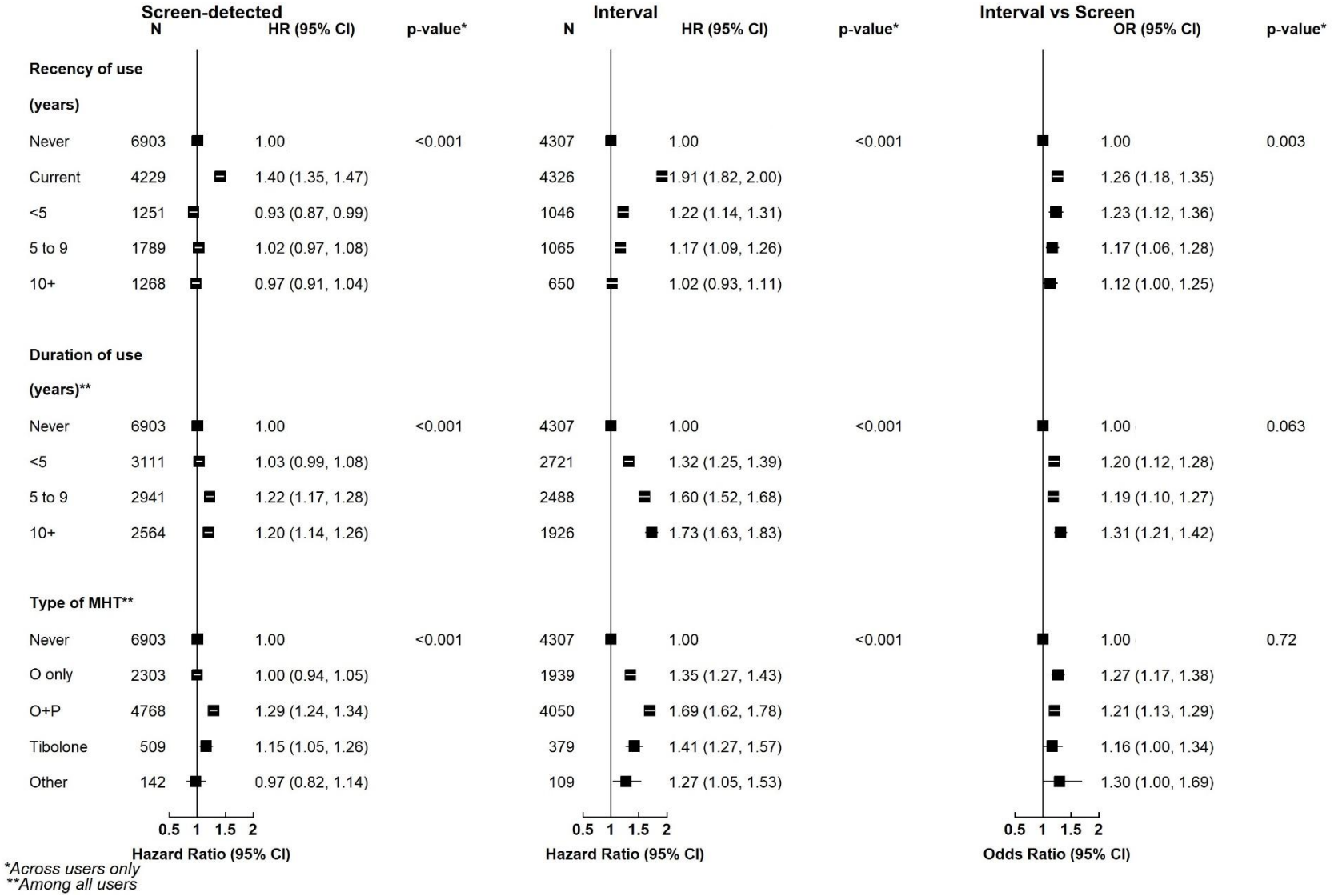
To further explore the association between MHT use with interval and screen-detected breast cancer, recency was cross-classified with duration of use (Figure 25) and type of MHT (Figure 26). This allowed me to determine the conditional effect of duration and type given recency. There was some evidence that recency explained part of the associations seen for duration in women who used MHT within 5 years. There was no effect of duration on interval vs screen-detected cancers among those who had ceased use 5+ years ago. However, there was no evidence of an interaction between duration and recency (p=0.092) nor an interaction between MHT type and recency (p=0.64).

The associations of MHT use by age at screen (<60 years and 60+ years) are displayed in Figure 27. There was evidence of interaction by age at screen for recency of MHT use (p=0.001). The effect of MHT recency on interval versus screen-detected cancers appeared greater in the 60+ years group than the <60 years group. This was not seen for duration and type of MHT.

5.4.3.1 Sensitivity analysis

The sensitivity analysis adjusting for and not adjusting for age at menopause, among women with known age at menopause, revealed that not adjusting for age at menopause in the current population made no material difference to the interpretation of the case-case analysis (Figure 28). The results shown in the figure were largely identical to each other and to the main analysis, which crudely adjusted for age at menopause.

Figure 24. Relative risk of breast cancer among postmenopausal women in England by recency, duration and type of MHT use



*Across users only
**Among all users

Figure 25. Relative risk for breast cancer among postmenopausal women in England in relation to duration and recency of MHT use

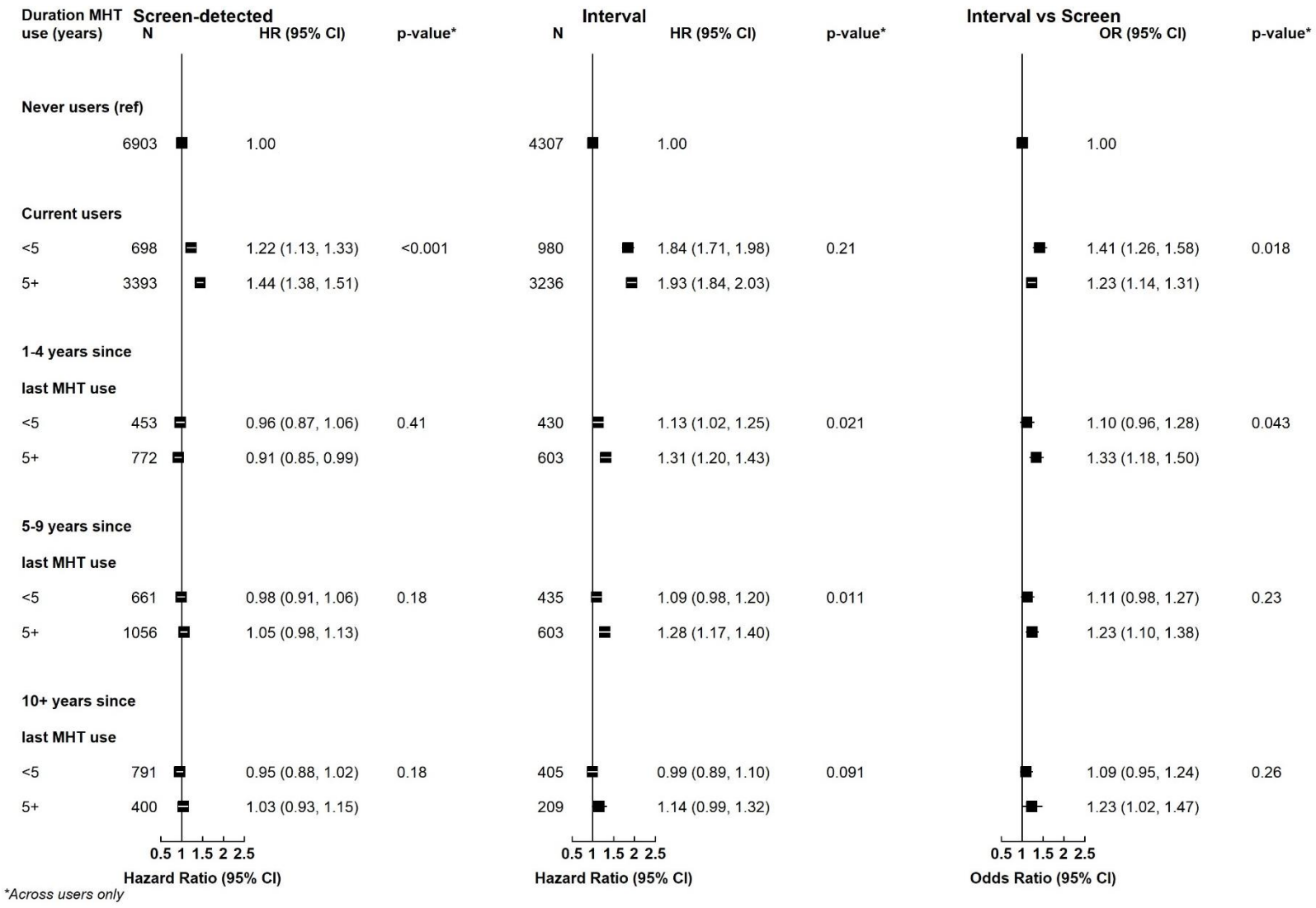


Figure 26. Relative risk for breast cancer among postmenopausal women in England in relation to type and recency of MHT use

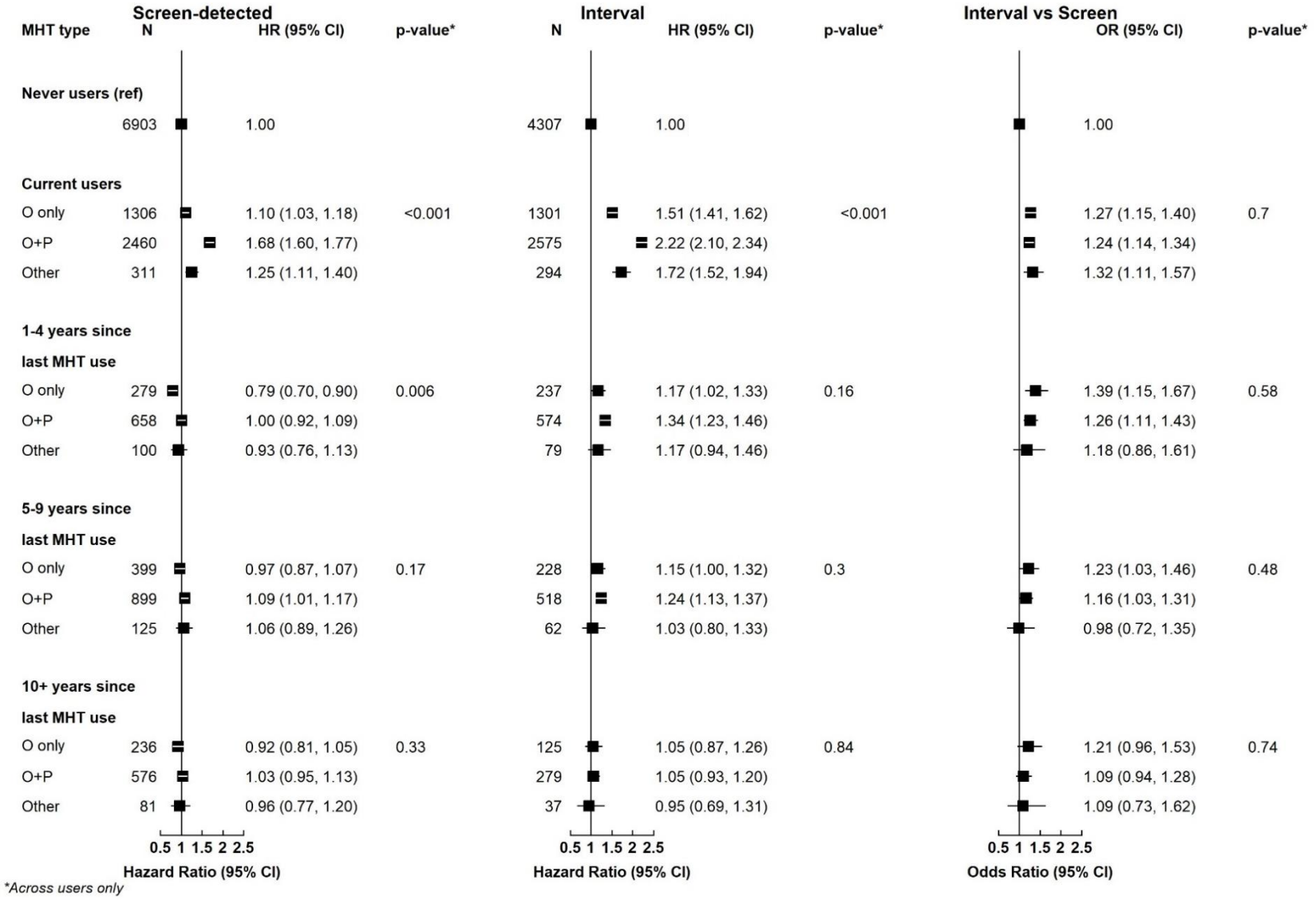
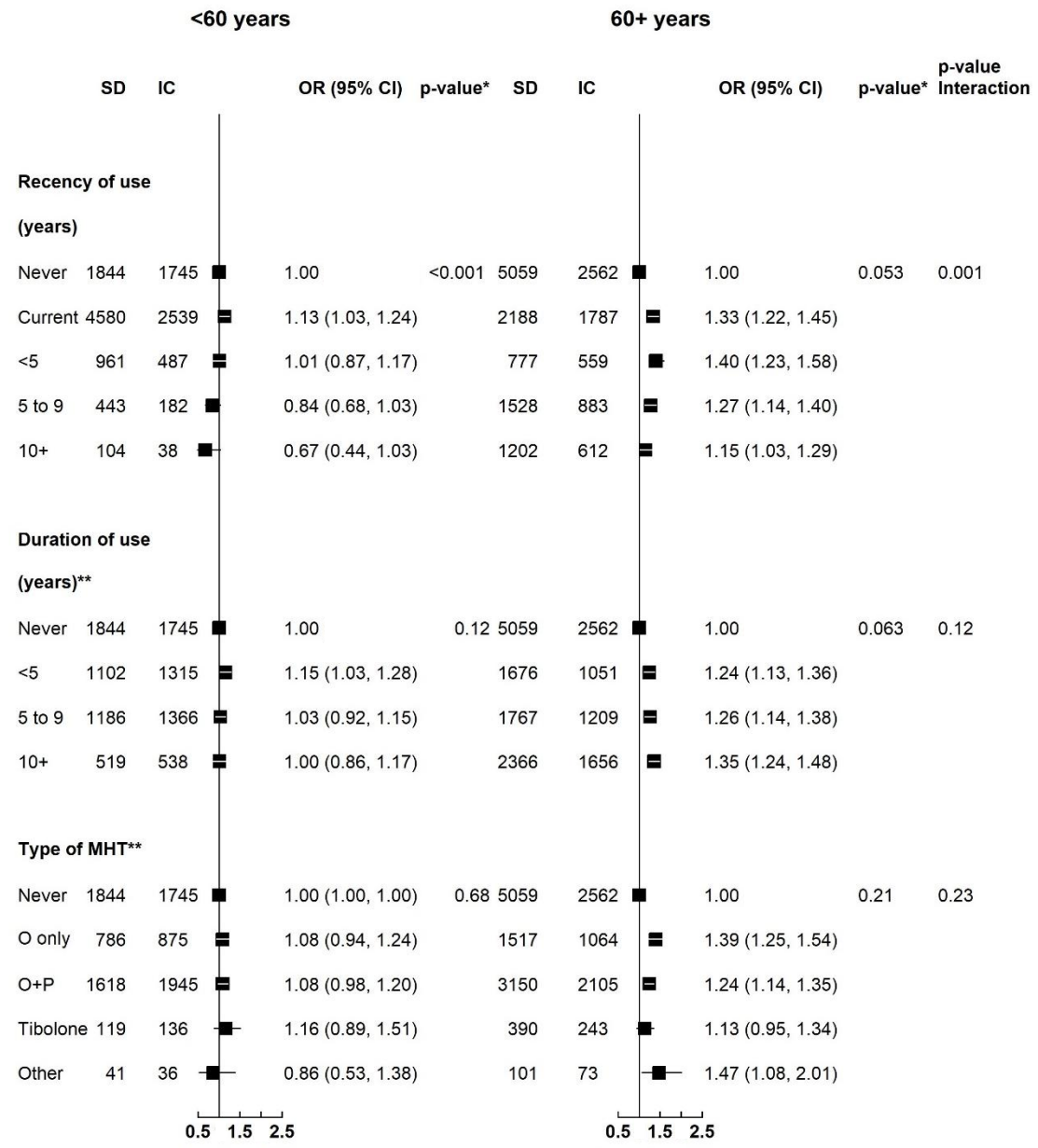
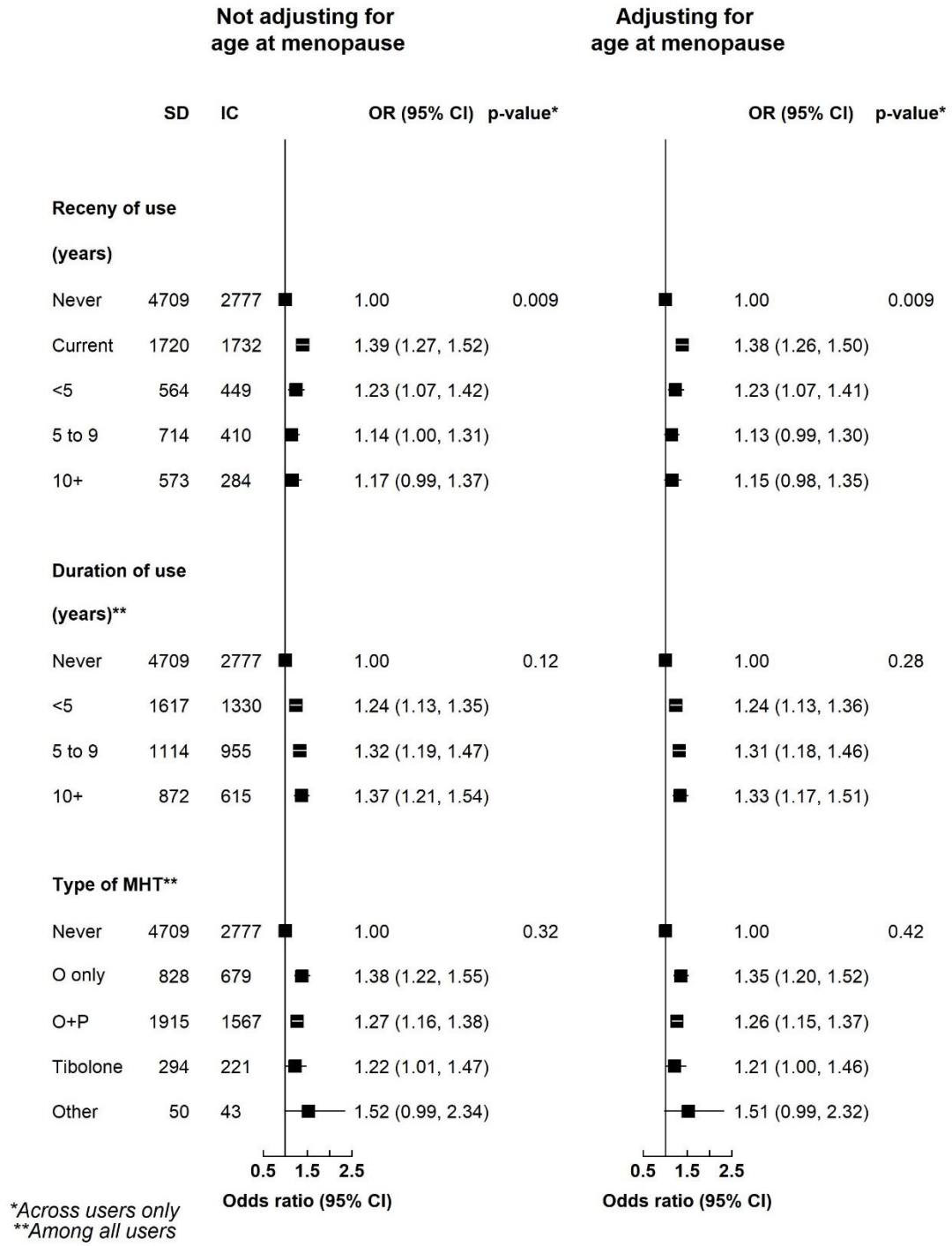


Figure 27. Relative risk for breast cancer among postmenopausal women in England in relation to recency and duration of MHT use by age at screen



*Across users only
**Among all users
SD, number screen-detected cancers;
IC, number interval cancers

Figure 28. Sensitivity analysis for interval vs screen-detected breast cancer adjusting for age at menopause



**Across users only*
***Among all users*

SD, number screen-detected cancers;
IC, number interval cancers

5.4.4 Discussion

With over one million women who attended breast screening, and over 27,000 breast cancer cases, this is one of the largest studies investigating the associations of MHT use with interval and screen-detected breast cancer risk and is the largest in the UK breast screening population. This analysis provides evidence that current MHT users have a 26% greater risk of interval versus screen-detected cancers, compared to never users. This excess risk appears to decline gradually after stopping use, but remains slightly elevated, even after 10+ years of ceasing use.

The UK NHSBSP screens women every three years for breast cancer. In most of the studies in the meta-analysis presented at the start of this chapter, screening occurred every two years. This meant an analysis in the UK population was necessary. Women in the UK are subjected to a greater time period during which an interval cancer could develop. This makes comparisons between the present study and the meta-analysis challenging.

Like the MWS analysis, the meta-analysis revealed that current MHT use is associated with a greater risk of interval versus screen-detected cancers, compared to never users: RR 1.69 (1.55, 1.84). This, however, is greater than what was found in the MWS: OR 1.26 (1.18, 1.35). The meta-analysis did not show any significant increase in risk among past users. In the present study, the odds ratio remained elevated amongst all categories of past users. This is likely because past use is related to recency of use, and therefore, studying recency within past use is more useful than past use alone.

In the meta-analysis, three studies provided some evidence that higher duration of MHT use was associated with an increased risk of interval cancers compared to screen-detected cancers.[234,262,269] The evidence in the MWS was less clear, and any slight association could be explained by the recency of MHT use.

The observed increased risk of interval versus screen-detected cancers among current MHT users could be explained by increased MD among current users, as seen in Chapter 4, reducing screening sensitivity. If this were the case, cancers among these women would be more likely to be detected

during the interval period when they became symptomatic. The results on current MHT use are consistent with the current understanding of how MHT affects MD and screening sensitivity.[60,254] A possible mechanism to explain this is the effect of MHT on breast tissue proliferation.[273,274] MHT activates oestrogen receptors within breast tissue to stimulate proliferation, corresponding to increased radiological density.[271] The results would suggest that the effects of MHT on breast tissue remain, even after cessation of use. This does seem to contradict the finding in Chapter 4, in which past MHT use was not associated with higher MD. This, however, may be because more recent past use was captured in the present analysis. In the MD analysis, past use likely represented MHT use over a decade before MD measurements were available. Therefore, the effects of past MHT on MD, beyond 10 years since cessation, were not substantial enough to be detected.

In the MWS I was able to look at multiple preparations of MHT. There was no clear evidence that type of MHT influenced the association of interval versus screen-detected breast cancer risk. Of note, the HRs for oestrogen-only and combined preparations were comparable. This is also what was suggested in the meta-analysis. [234,264,266] The finding is perhaps surprising given the previously published evidence that combined MHT has a greater effect on mammographic density than oestrogen-only preparations.[60] Therefore, one would expect combined MHT to have a more significant effect on screening sensitivity than oestrogen-only MHT. This is also consistent with Chapter 4, in which there was no evidence of an association between MHT type and MD.

I was able to examine the effect of the associations by age at screen. The influence of MHT recency on screening sensitivity appeared to be greater in the 60+ year group than the <60-year group. An explanation for this could be that as younger women have comparatively denser breasts, the effects of MHT on screening were not as apparent. Whereas in older women with fattier breasts, the effects of MHT were comparatively greater and therefore, a difference in risk was detected.

With almost 2 million current users, MHT use has risen from 11% to 15% of women aged 45 to 64 in England, between 2021 and 2023. This is likely to contribute to a rise in breast cancer cases which

needs to be addressed in the context of current screening practices and communication of patient information.[240] Understanding the effects of MHT on screening can aid in tailoring targeted screening programmes. To do this however, cost implications would need to be considered and therefore health economic evaluations would be necessary to understand the cost/benefit of tailored screening. Furthermore, when being counselled on the risks and benefits of MHT use, women should be informed of the impact MHT use may have on screening. Whilst it is well accepted that MHT use increases breast cancer risk, women may feel this risk is mitigated by attending screening. The current study provides evidence that not only does MHT use limit the effectiveness of screening whilst presently using MHT, but that this effect continues to a lesser degree for several years after cessation. This would need to be communicated to potential and current MHT users so they can make informed choices regarding MHT use.

The strengths of this study include the large sample size, providing statistical power to detect the association between MHT and interval versus screen-detected cancers. Furthermore, the repeat questionnaires allow the use of the most up-to-date MHT status, to avoid misclassification of exposure and dilution of effect. The MWS questionnaires collected data on several important potential confounders, which could be adjusted for in the analysis.

A disadvantage of this study is that the risk of screen-detected versus interval cancer serves as a proxy for screening sensitivity, not a direct measure. It is possible that MHT has a greater effect on more aggressive, fast-growing, tumours, which are more likely to become clinically apparent during the interval period. This, however, is unlikely because MHT is more strongly associated with luminal A-like (ER+/PR+) breast cancer, which is slow growing, better differentiated, and less aggressive, and less so associated with luminal B cancers which more aggressive.140,222–224 The Collaborative Group on Hormonal Factors in Breast Cancer found the relative risks for current MHT use (for both combined and oestrogen-only preparations), during years 5–14 were substantially greater for ER+ than ER- tumours but still significantly greater than 1 for ER- tumours (RR oestrogen-only for ER+: 1.45

[1.38,1.53]; ER-: 1.25 [1.13, 1.38]; RR combined for ER+: 2.44 [2.35, 2.54]; ER-: 1.42 [1.30, 1.55]). [14] Furthermore, MHT is also more associated with lobular breast cancers, which are harder to detect at screening because they are more diffuse than ductal breast cancers. A study with 880 ductal cases, 1,027 lobular cases, and 856 controls, reported that oestrogen-only and combined MHT use was associated with an increased risk of lobular breast cancer: OR 1.6 (1.1–2.2) and 2.3 (1.7–3.2), respectively, but neither were associated with risk of ductal breast cancer.[275] This would further contribute to the increased risk of interval vs screen-detected cancers in MHT users.

Unfortunately, data on MD was not generally available. It would be interesting to examine whether adjusting for MD affects the results. This would help understand whether the increased risk, among MHT users, is likely due to increased MD and reduced screening sensitivity. At present, the MWS has MD data for roughly 10,000 women, which includes 600 breast cancer cases. This would not provide enough cases for the present analysis. However, as more MD data become available for the women in the MWS, the analysis could be repeated, adjusting for MD.

A further limitation was that data collection relied on personal recall. Women may unknowingly recall information about their MHT use or other confounders incorrectly. This would lead to misclassification of exposure and would likely attenuate the results. However, the self-reported information in the MWS has been validated in subsets of the participating women against objective measures or routinely collected health records. Good agreement has been found between the reported measures and validated measures.[118] Despite the repeat questionnaires, it was still possible that current/recent MHT use may have been four years out of date as women may have stopped after the questionnaire and before their next screen.

Approximately half of the women had an unknown age at menopause. This is potentially an important confounder as it affects breast cancer risk and MHT use. The sensitivity analyses suggested that adjusting for age at menopause had little effect on the results. I was able to crudely adjust for age at menopause by adding those who had a hysterectomy, without a bilateral oophorectomy, as a separate

category. Furthermore, the average age at menopause was 47.5 years which is below the current average of 51 years. It is not clear why there is such a discrepancy, and may bring the generalisability of the results into question as age at menopause would determine MHT use and breast cancer risk.

5.4.5 Conclusions

The MWS results suggest that MHT use affects mammographic sensitivity, but that this effect declines after stopping. Importantly though, some reduction in sensitivity continues even after 10 years of cessation. Recognising the effect of MHT on screening and how long this lasts can help direct public health education on the limitations of screening for MHT users and on advice about the risks of MHT, even after cessation. The results need to be communicated to women on MHT and those considering using MHT in the future so that they can make informed decisions. The understanding can also assist in tailoring the screening programme to individuals for whom current screening practices may not be as effective. Health economic evaluations will be necessary to understand the financial burden of the excess risk of interval cancers and whether new screening practices can mitigate this.

Chapter 6. Predicting risk of breast cancer subtypes in the Million Women Study

6.1 Introduction

With the incidence of breast cancer increasing, efforts to improve its prevention and early detection are essential. This includes breast cancer risk prediction models. These have been developed to quantify the effect of multiple risk factors on risk and potentially guide targeted screening practices. Identifying women at higher risk of breast cancer and tailoring screening programmes to them may reduce over-diagnosis and missed diagnoses and improve the cost-effectiveness of population-based screening programmes. These models can also highlight women who may benefit from preventative measures or lifestyle modifications, ultimately reducing the overall burden of breast cancer.

There are many breast cancer risk prediction models.[116] These models were designed to determine the general risk of breast cancer or assess the probability of carrying high penetrance variants such as mutations in the BRCA1 or BRCA2 genes. Examples of these models include Gail/BCRAT, Tyrer-Cuzick, Claus, BRCAPRO, and Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA).[117]

There are, however, few published models aimed at predicting specific subtypes of breast cancer. This is important as various risk factors are associated with certain subtypes of breast cancer. Some risk factors may only be related to specific subtypes or may have qualitatively different associations across various subtypes, and so a generic model may not be predictive of the risk of each individual subtype. Therefore, in this chapter, I will assess how well a risk prediction model for any breast cancer – developed from risk factor information in questionnaires – predicts different subtypes of breast cancer based on the data from the MWS. The subtypes which will be considered here are based on mode of detection (interval and screen-detected breast cancer), and grade (high and low grade breast cancer). There is only limited hormone receptor data available in the MWS and therefore this could not be examined.

As discussed in Chapters 1 and 5, interval cancers are associated with poorer prognosis than screen-detected cancers because of a combination of later stage and more aggressive tumour biology.[88–90] Therefore, identifying women who are more likely to develop them can help guide preventative measures or better-tailored screening programmes. The grade refers to the degree to which the cancer cells histologically resemble the cells of the surrounding breast tissue. A lower grade indicates the cells are similar and are, therefore, termed highly differentiated. A higher grade indicates that the cancer cells look significantly different from normal breast tissue and are, therefore, termed poorly differentiated. High grade cancers usually grow faster and spread more rapidly than low grade cancers, and therefore, are not usually detected at screening as they likely develop in the interval between screens. Hence the former are associated with a poorer prognosis.[85] As with interval cancers, predicting women who may develop high grade cancer can help guide clinical pathways.

The MWS data provides risk factor data from questionnaires and does not possess data from genetics, blood markers, or MD data for all women. Reproductive and hormonal risk factors, which are readily available from the MWS, have stronger associations with ER+ breast cancer than with ER- breast cancer. ER+ breast cancer is less aggressive, slower growing and more indolent than ER- and, therefore, is more likely to be detected at screening and more likely to be of lower grade. Hence, it is expected that a risk prediction model based on reproductive and hormonal risk factors would be better at predicting screen-detected cancers and low grade cancers.[276]

In the next section, I shall provide a brief background to the literature on breast cancer risk prediction models and subtype-specific models, before going onto the main methods and results of the chapter.

6.2 Background review of the literature

To examine earlier work on prediction models, I reviewed the literature on general breast cancer prediction models and subtype-specific prediction models. Several systematic reviews on breast cancer prediction models exist. A recent systematic review from 2022 included 40 studies published between 1989 and 2021.[116] This review included population-based studies and used data from cross-

sectional studies, cohort studies, case-control studies, and randomised controlled trials. The review used models based on regression analyses, rather than studies that used models from machine learning algorithms. Therefore, the studies in this review are more comparable with the analysis that will be presented in this chapter. Age, family history, and reproductive factors were the most used factors in the models. The period of time over which the model was assessed varied by study.

A systematic review of breast cancer risk prediction using machine learning was published in 2024.[277] The review included models based purely on imaging data from mammograms, and models that also used clinical information, gene expression profiles, and biochemical data. The models were developed over varying periods of time. The models in this review performed better than those that did not use machine learning methods, as reported in the 2022 systematic review. This is likely because they had machine learning analysed information from mammograms which is a strong risk factor for breast cancer. The AUC of the studies using only imaging data ranged from 0.70 to 0.86, for non-imaging data, from 0.61 to 0.95, and for those using a combination of both, from 0.65 to 0.89.

The more commonly known models include Gail/BCRAT, Tyrer-Cuzick (IBIS), BRCAPRO, and Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA).[117,278] The Gail model uses non-genetic risk factors such as reproductive factors, benign breast disease, atypical hyperplasia, age, race, and first-degree relatives with breast cancer.[279] The IBIS model includes similar parameters but also requires information on second and third-degree relatives, Ashkenazi Jewish ancestry, a family history of ovarian cancer, and mammographic density.[280] The BRCAPRO and BOADICEA models also use detailed family history information and principles of genetic inheritance to determine the probability of carrying high penetrance variants such as the BRCA1 or BRCA2 mutations. The BRCAPRO model does not consider non-genetic risk factor information but uses detailed information on familial breast and ovarian cancer, including male breast cancer, genetic testing, and biomarker testing.[279] The BOADICEA model also includes the effects of many genetic variants and

polygenic risk scores (PRS), in addition to demographic, reproductive and other information available.[281]

The IBIS and BOADICEA tools are used in clinical practice to assess risk, particularly in those with a strong family history of breast cancer. Based on the BOADICEA, the predicted lifetime risks for women in the UK population range from 2.8% for the 1st percentile to 30.6% for the 99th percentile.[282] Breast cancer risk is assessed clinically through a detailed review of family of breast cancer and ovarian cancer, as well as lifestyle and reproductive factors. The NHS has endorsed models like BOADICEA and IBIS help stratify risk in order to guide enhanced screening, as well as determine whether a woman is eligible for testing of a proven germline pathogenic variant like BRCA1/2.[283,284] The tools are used to assess the 10-year risk of breast cancer. Screening will then depend on age and the 10-year breast cancer risk.[285]

PRS scores can be included into breast cancer prediction models to improve their predictive accuracy. Genome-wide association studies have identified common, low-penetrance genetic variants associated with breast cancer risk, which when combined together may confer significant breast cancer risk. Prediction models have been extended to include PRS scores.[286] A study with 750 breast cancer cases examined the extent to which breast cancer risk prediction models can be improved by adding PRS scores. Using the BOADICEA tool, they reported an AUC of 0.66 for the model with clinical risk factors only and of 0.70 with PRS was combined, for 5-year risk. The same study also used the BRCAPRO tool and reported an AUC of 0.65 and 0.69, respectively. For the IBIS, the AUC went from 0.57 to 0.63, after adding PRS score.[287] In another study with 619 cases women, the AUC went from 0.57 to 0.65 after adding PRS score to the BOADICEA.[288]

There are limited models on breast cancer subtypes and fewer models for breast cancers defined by mode of detection. Table 23 summarises the literature on models that predict screen-detected or interval breast cancers. The AUCs or C-statistics for interval breast cancer ranged from 0.63 to 0.86, and for screen-detected cancer, from 0.53 to 0.64. While most studies used logistic regression to

generate the model, a few used neural networks and deep learning models to incorporate information from screening mammograms. All studies used MD in their model, which is not generally available in the MWS. The process of validation was not clear from the studies. All models presented in Table 23 were developed specifically for each subtype. This will differ from the initial part of my analysis in which I will determine how well an overall breast cancer model predicts subtype-specific breast cancer. This will be useful as it will give insight into whether subtype specific models would be better used in clinical practice or risk stratification rather than generic models.

My review of the literature yielded no risk prediction models for high and low grade breast cancer. Other classifications of breast cancer subtypes are determined through immunohistochemistry based on the expression of hormone receptors: oestrogen (ER), progesterone (PR), and human epidermal growth factor (HER2).[86] Cancers which are ER-, PR- and HER2- (triple negative breast cancers) are usually considered the most aggressive. The Nurses' Health Study developed separate models based on questionnaire data on reproductive factors, benign breast disease, MHT use, BMI, height, alcohol use, and family history for ER+/PR+ and ER-/PR- breast cancer. The AUCs were 0.64 and 0.61, respectively.[289]

The performance of general breast cancer models has not yet been assessed by subtype. Analysis from the MWS will allow me to assess how well a risk prediction model for overall breast cancer predicts different subtypes of breast cancer. This will help determine which risk factors are better at predicting the breast cancer subtypes more likely to be not detected at screening. Furthermore, many of the above models require complex machine learning processes, MD data, genetic data, or blood biomarkers. These tools may not be available in general clinic settings. Therefore, exploring general and subtype-specific breast cancer risk prediction models using easily accessible questionnaire data could be of interest.

Table 23. Summary of studies on interval/screen-detected breast cancer risk prediction models

Study	Location	Study design	N	Model	Available data	AUC/C statistic (95% CI)
Celik 2023[290]	Turkey	Case-control	IC: 323 SC: 441	Neural network model	Digital mammograms	AUC: IC: 0.86
Vachon 2023[291]	USA	Nested case-control	IC:286 Matched controls: 599 SD: 4670 Matched controls: 3567	Logistic regression	MD (different measures) and digital mammograms	AUC IC: 0.68 (0.65, 0.72) SD: 0.65 (0.64, 0.67)
Hacek 2022[292]	Croatia	Case-control	IC: 789	Logistic regression	MD, OC use, MHT, age first birth, FH of breast/ovarian cancer	AUC: 0.66 (0.60, 0.71)
Wanders 2022[293]	Netherlands	Nested case-control	IC: 2222 Matched control: 4661	Neural network model combining AI and MD	MD and digital mammograms	AUC 0.79 (0.77, 0.81)
Burnside 2021[294]	UK	Case-control	IC: 297 SD: 302 Matched controls: 605	Logistic regression	Different measures of MD: Fibroglandular volume (FGV), volumetric breast density (VBD), visual analogue scale of MD (VAS)	AUC: **FGV:** IC: 0.65 SD: 0.61 **VBD:** IC: 0.63 SD: 0.53

Study	Location	Study design	N	Model	Available data	AUC/C statistic (95% CI)
Zhu 2021[295]	USA	Case-control	IC: 351 SD: 1609 Control: 6369	Deep learning model and logistic regression	MD, age, BMI, FH, BBD, race, deep learning predictors from digital mammograms	C statistic: IC: 0.72 (0.66, 0.78) SD: 0.66 (0.63, 0.69) Clinical risk factors only: IC: 0.71 (0.65, 0.77) SD: 0.62 (0.59, 0.65)
Nguyen 2020[296]	Australia	Nested case-control	IC: 168 Matched controls (no cancer): 498	Logistic regression	MD, FH, breast tissue aging, interactions: breast tissue aging x MD; breast tissue aging X FH	AUC 0.73 (0.69, 0.77)
Hinton 2019[297]	USA	Nested case-control	IC:173 SD: 182	Logistic regression	MD and deep learning predictors from digital mammograms	AUC: 0.82
Kerlikowske 2018[298]	USA	Nested case control	IC: 351 SD: 1609 Matched controls: 4409	Logistic regression	MD, age, race, FH, BBD, BMI	C statistic: IC: 0.72 (0.69, 0.75) SD: 0.62 (0.61, 0.64)
Nguyen 2018[299]	Australia	Nested case-control	IC: 168 SD: 422 Matched controls: 498 for IC and 1197 for SD	Logistic regression	MD, BMI, age, mode of detection	AUC IC: 0.75 (0.70, 0.79) SD: 0.64 (0.61, 0.67)

IC, interval breast cancer; SD, screen-detected breast cancer; AUC, area under the curve; MD, mammographic density; FH, family history of breast cancer; BBD, benign breast disease; FBV, fibroglandular volume; VBD, volumetric breast density

6.3 Aims

The aim of this chapter was to determine how accurately a general breast cancer risk prediction model, based on risk factor questionnaire data, could predict the risk of specific breast cancer subtypes:

- Interval breast cancers
- Screen-detected breast cancers
- High grade breast cancers
- Low grade breast cancers

From this, I aimed to determine which risk factors were better at predicting certain breast cancer subtypes, particularly those missed at screening. Another aim of this chapter was to develop separate models to predict the risk of specific subtypes and compare their performance to that of the general breast cancer model. Finally, a further aim was to determine how the performance varied according to follow-up time periods of 5 years and 10 years.

6.4 Methods

6.4.1 Participants and data

Details about the MWS and the NHS Breast Screening Programme are given in Chapter 2. In 2013, the MWS cohort was linked to NHS Breast Screening Programme records on screening episodes which had started by 31st December 2012. Linkage provided information on the dates women were invited for their routine screen and whether they attended or not. The MWS study also has linked information on breast cancer tumour characteristics, including tumour grade. Information on both screening episodes and tumour characteristics were only available for women living in England (not Scotland). At recruitment, participants provided information on anthropometric, lifestyle, reproductive, medical, familial, exogenous hormone use (including MHT use), and sociodemographic risk factors. Participants were resurveyed approximately every three to five years after recruitment.

6.4.2 Candidate predictors

Information on candidate predictors for the risk prediction models was obtained from the recruitment questionnaires. The candidate predictors were chosen based on their association with breast cancer risk, as documented in the literature, and their use in other risk prediction models. They included: MHT use (never, past, current); duration of MHT use (never, <1, 1-4, 5-9, 10-14, 15+ years); MHT type (never, oestrogen-only, combined, other); alcohol (<3, 3-7, 8+ units/week); smoking status (never, past, current); BMI (<25, 25-29, 30+ kg/m^2); age at menarche (<12,12-13,14+); oral contraceptive use (yes/no); parity (nulliparous, 1-2, 3+ births); age at first birth (nulliparous, <20,20-24, 25-29, 30+ years); first degree relative with breast cancer (yes/no); benign breast disease (yes/no); Townsend index of deprivation (quintiles); and menopausal status/age at menopause (premenopausal, hysterectomy, <46, 46-48, 49-51, 52+ years). Hysterectomy was included as a separate category in age at menopause. This was because a woman with a hysterectomy would not know when she reached natural menopause as she would immediately stop menstruating after a hysterectomy but would still retain ovarian function if she had not undergone a bilateral oophorectomy.

6.4.3 Outcome response for the risk prediction model

The main outcomes of interest were invasive breast cancer (ICD10: C50) and its subtypes based on mode of detection and grade. The definitions and explanations of screen-detected and interval breast cancers are provided in detail in Chapter 5. Briefly, screen-detected cancers were defined as those registered within nine months of the start of a screening episode. Interval cancers were defined as those registered after nine months but prior to the next screening episode. Low grade was defined as cancers with grade 1 or 2, and high grade was defined as cancers with grade 3.

6.4.4 Statistical analysis

6.4.4.1 Study design

The presented study design was a matched case-control study nested within a prospective cohort study. The main rationale for this design was to manage the computational power required to carry out the analysis. A nested case-control study method was less computer intensive and would give similar results provided enough controls were selected per case.

Women in this case-control study had to be at risk of a screen-detected or interval cancer. Cases were matched to four controls on attained age, age at the most recent screen (prior to the age at which the case was diagnosed), and the region of recruitment. Controls were matched with replacement; therefore, a control could be a control for more than one case. Furthermore, a woman could be both a case and control if she was cancer-free when another woman had cancer and then developed cancer later. Women who returned their recruitment questionnaire after receiving the results of their screen were excluded from my analyses. This was because those who were diagnosed with breast cancer may have been more likely to recall certain risk factors, especially those thought to be associated with breast cancer, than women who were not diagnosed with breast cancer. Women with a previous cancer diagnosis were also excluded. Given that there was no screening data for women in Scotland, these women were also excluded.

Missing values were imputed using multiple imputation. The numbers of women with missing values for each specified variable were: MHT status (3356 [1.2%]); MHT type (23,768 [8.3%]); MHT duration (5909 [2.1%]); smoking (16,530 [5.8%]), alcohol (2080[0.7%]); age at menarche (5776 [2.0%]); age at first birth (7515 [2.6%]); parity (1084[0.4%]); oral contraceptive use (3130 [1.1%]); family history of breast cancer (17,553 [6.1%]); benign breast disease (2379 [0.8%]); BMI (14,202 [5.0%]); deprivation (1821 [0.6%]); and age at menopause (73,018[25.5%]). All variables in the model were used to impute the missing values, as well as data on age at screen and outcome. Five imputations were completed.

6.4.4.2 Development of model for all breast cancer

The model was developed using all invasive breast cancer cases and their matched controls within the MWS, where screening information was available. This was done using conditional logistic regression. As described earlier, candidate predictors were chosen based on their association with breast cancer risk and use in the other common breast cancer prediction models. Therefore, no formal model selection process was employed, and all candidate predictors were included in the risk prediction model. The large number of breast cancer cases within the MWS further justified this.[300] Linear

predictors of the conditional logistic regression model were then generated. The model was assessed overall and by subtype based on the mode of detection and grade of breast cancer.

6.4.4.3 Indicators of model discrimination

Discrimination was assessed by estimating the AUC and its standard error (SE), producing the receiver operating characteristic (ROC) curve, plotting the histograms of the linear predictors of cases and controls, and calculating the proportion of cases within each quintile of the linear predictor. Traditional prediction analysis assumes independent observations, but matching in conditional logistic regression creates dependencies which had to be taken into account as the nested case-control design over samples cases. Inverse probability weighting adjusts for this by giving lower weights to cases and higher weights to controls to reflect their original prevalence in the cohort.[301]

The overall breast cancer model was used in the data restricted to specific subtypes (interval cancer, screen-detected cancer, high grade cancer, low grade cancer) separately. The performance of the overall model on the subtypes was then assessed using the AUC. This was to determine how well a model generated in all breast cancers could predict breast cancer subtypes.

The model was also assessed on cases and matched controls in the period 5 years or 10 years after reporting exposures at recruitment. This was to observe how well the model predicted risk in the short- and longer-term.

To determine the impact of the inclusion of each variable in the model on the model, each variable was added separately to the basic model with deprivation, and the AUC was calculated and compared to that of the overall model. A separate model with all MHT variables (status, duration and type) was evaluated along with a model with all reproductive risk factors (age at menopause, age at menarche, parity, and age at first birth). This was to assess the collective role of MHT and reproductive factors. The single risk factor models were then applied to the subtypes separately, and their performance was assessed using AUC to determine how well each risk factor was able to predict any of the subtypes.

6.4.4.4 Development of subtype-specific and time period specific models

To develop subtype-specific models, cases of a specific subtype and their matched controls were taken from the original nested case-control dataset. Subtype-specific models were then generated as described for the overall model and their performance compared to the overall breast cancer model. Similar approaches to the assessment of performance and validation, as outlined for the overall breast cancer model, were applied to the subtype-specific models. Time period specific models (5- and 10-year follow-up) were also generated for all cancers, and their performance was compared to the cancer model based on the entire follow-up period, using AUC. This was to examine whether the accuracy of the model varied over time.

6.4.4.5 Validation

As no suitable dataset linked to screening records was available for external validation, internal validation was necessary. Internal validation was completed using K-fold cross-validation with 10 folds.[300] The unimputed matched dataset was divided into 10 equal-size folds. For each of the 10 iterations, I imputed the data in the training set (consisting of 9 groups) and then imputed the test set (consisting of one group). This was to avoid information from the testing group influencing the model development datasets and vice versa. The model was developed in 9 groups from the training set, and the performance was evaluated in the remaining group (test set). This process was completed 10 times, so each time, a different set of 9 groups was used to generate the model, and a different group was used to evaluate model performance. The average performance over the 10 iterations was used to estimate the overall model performance. Validation was carried out for the overall model, as well as the subtype and time period specific models.

Analyses were conducted, and figures were generated using Stata 18.[127]

6.5 Results

There were 57,189 invasive breast cancer cases matched to four controls on attained age, age at most recent screen (prior to the age at which the case was diagnosed), and on region of recruitment. Mean

age at diagnosis was 63.1 years (SD 5.2). Mean time between recruitment and breast cancer diagnosis was 8.9 years (SD 4.9).

6.5.1 All breast cancer cases

The coefficients and odds ratios of the full conditional logistic regression model are shown in Table 24.

The AUC for the above model was 0.59 (SE 0.0013) (Figure 29). The histogram of the linear predictors of cases and controls, and proportion of cases within each quintile of linear predictor are shown in Figure 30 and in Figure 31, respectively. Whilst the histogram peak of the cases was to the right of the controls, the histograms did overlap considerably, indicating some, albeit relatively poor, discrimination in the model. The proportion of cases increased with each quintile of linear predictor: 14.3% and 27.7% of cases in the lowest and highest quintiles, respectively.

Table 24. Conditional logistic regression model of breast cancer risk factors and breast cancer

Variable	Coefficient (95% CI)	Odds ratio (95% CI)
MHT status		
Never	0.00 (ref)	1.00 (ref)
Past	-0.19 (-0.23, -0.15)	0.83 (0.79, 0.86)
Current	0.06 (0.01, 0.1)	1.06 (1.01, 1.11)
MHT duration (years)		
<1	0.00 (ref)	1.00 (ref)
1-4	0.05 (0.01, 0.1)	1.06 (1.01, 1.1)
5-9	0.08 (0.04, 0.12)	1.08 (1.04, 1.13)
10-14	0.12 (0.07, 0.17)	1.13 (1.07, 1.19)
≥15	0.16 (0.08, 0.24)	1.17 (1.08, 1.27)
MHT type		
O-only	0.00 (ref)	1.00 (ref)
O+P	0.27 (0.24, 0.3)	1.31 (1.27, 1.35)
Other	0.10 (0.04, 0.16)	1.11 (1.05, 1.17)
Age at menarche (years)		
<12	0.00 (ref)	1.00 (ref)
12-13	-0.02 (-0.05, 0.0006)	0.98 (0.95, 1.00)
14+	-0.05 (-0.08, -0.03)	0.95 (0.92, 0.97)
Parity		
Nulliparous	0.00 (ref)	1.00 (ref)
1-2	-0.27 (-0.31, -0.23)	0.76 (0.73, 0.8)
3+	-0.32 (-0.36, -0.27)	0.73 (0.70, 0.76)
Age at first birth (years)		
20	0.00 (ref)	1.00 (ref)
20-24	0.04 (0.004, 0.07)	1.04 (1.00, 1.07)
25-29	0.14 (0.10, 0.17)	1.15 (1.10, 1.19)
30+	0.27 (0.23, 0.32)	1.31 (1.26, 1.38)
Age at menopause (years)		
<46	0.00 (ref)	1.00 (ref)
46-48	0.11 (0.06, 0.16)	1.12 (1.07, 1.17)
49-51	0.15 (0.11, 0.19)	1.16 (1.11, 1.2)
≥52	0.18 (0.13, 0.22)	1.2 (1.14, 1.25)
Hysterectomy	0.11 (0.07, 0.16)	1.12 (1.07, 1.17)
Premenopausal	0.32 (0.26, 0.39)	1.38 (1.30, 1.47)
Alcohol units/week		
<3	0.00 (ref)	1.00 (ref)
3-7	0.04 (0.02, 0.06)	1.04 (1.02, 1.07)
>8	0.15 (0.12, 0.17)	1.16 (1.13, 1.19)

Variable	Coefficient (95% CI)	Odds ratio (95% CI)
Smoking status		
Never	0.00 (ref)	1.00 (ref)
Past	0.07 (0.04, 0.09)	1.07 (1.04, 1.09)
Current	0.20 (0.18, 0.23)	1.22 (1.19, 1.26)
Oral contraceptive use		
No	0.00 (ref)	1.00 (ref)
Yes	0.001 (-0.02, 0.02)	1.00 (0.98, 1.02)
Family history breast cancer		
No	0.00 (ref)	1.00 (ref)
Yes	0.44 (0.42, 0.47)	1.56 (1.52, 1.6)
Benign breast disease		
No	0.00 (ref)	1.00 (ref)
Yes	0.35 (0.32, 0.37)	1.42 (1.38, 1.45)
BMI		
<25	0.00 (ref)	1.00 (ref)
25-29	0.15 (0.13, 0.17)	1.16 (1.13, 1.18)
30+	0.28 (0.26, 0.31)	1.33 (1.29, 1.36)
Deprivation quintile		
Q1, least deprived	0.00 (ref)	1.00 (ref)
Q2	-0.01 (-0.04, 0.02)	0.99 (0.97, 1.02)
Q3	0.01 (-0.02, 0.04)	1.01 (0.98, 1.04)
Q4	0.01 (-0.02, 0.04)	1.01 (0.98, 1.04)
Q5, most deprived	0.03 (0.004, 0.07)	1.04 (1.00, 1.07)

O-only, oestrogen-only MHT; O+P, combined MHT

Figure 29. ROC curve of risk prediction model of all breast cancer cases

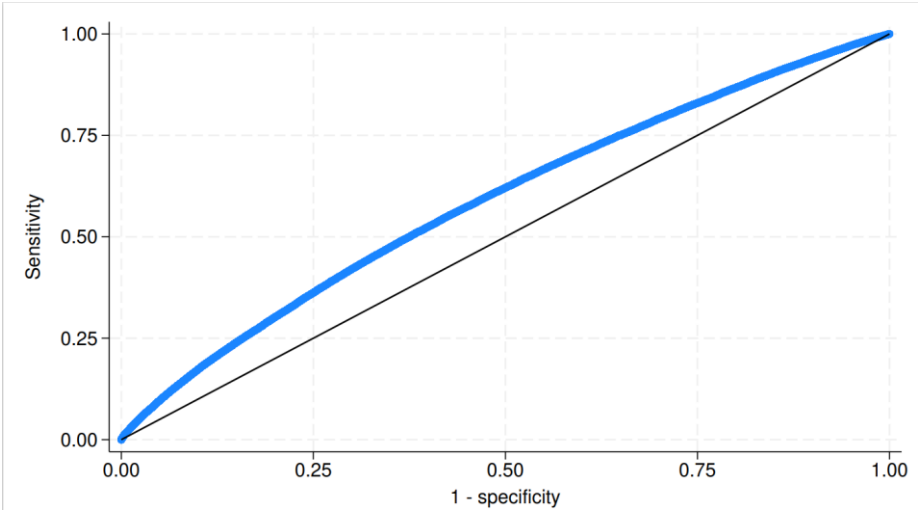


Figure 30. Histogram of linear predictor by cases and controls

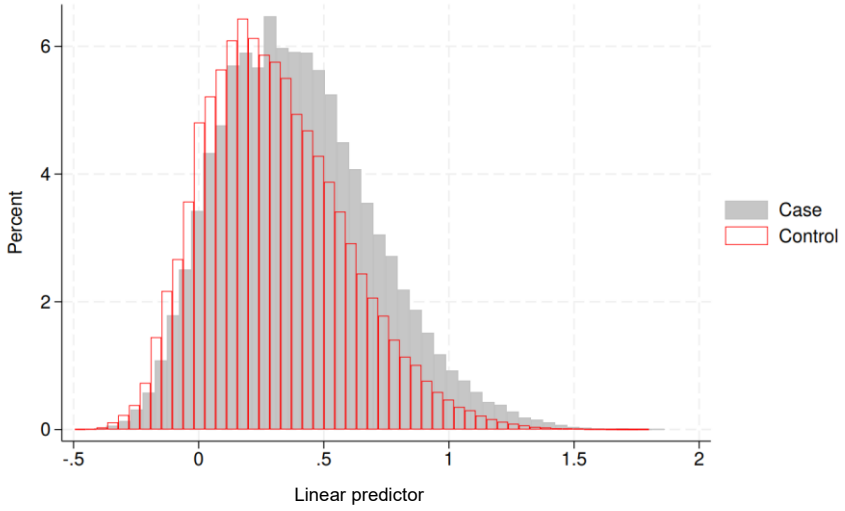
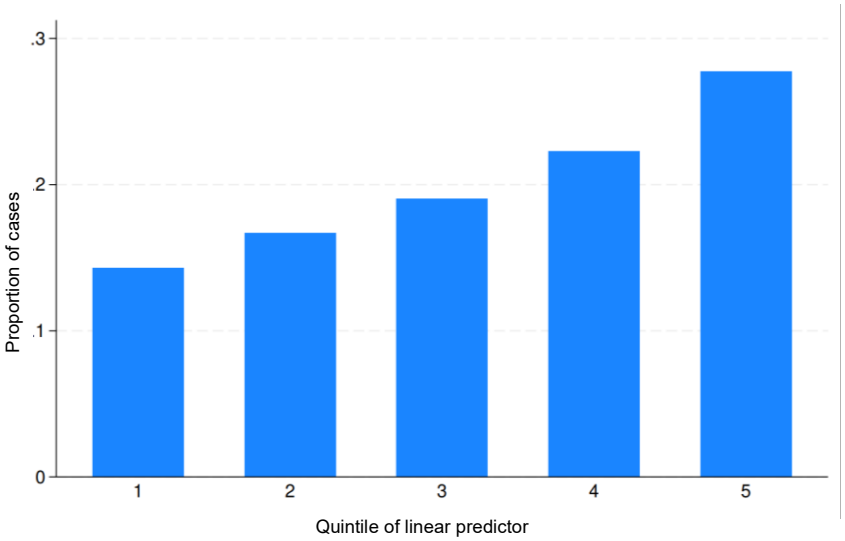


Figure 31. Proportion of cases within each quintile of linear predictor



6.5.2 Interval and screen-detected breast cancers

There were 15,902 screen-detected cases and 11,662 interval cases. The above model was, in turn, applied to screen-detected and interval cancers, and the performance was determined. The model performed slightly better for interval cancers than screen-detected cancers (AUC = 0.60 [0.0030] vs 0.58 [0.0026]).

6.5.3 Low and high grade breast cancers

As with the screen-detected and interval breast cancers, the overall breast cancer model was applied to the low and high grade breast cancers separately. There were 37,656 low grade cancers and 15,078 high grade cancers. The AUC was marginally higher in low grade breast cancers than in high grade breast cancers (AUC = 0.59 [0.0016] vs 0.57 [0.0026]).

6.5.4 Performance of each candidate predictor

The impact of the inclusion of each candidate predictor on discrimination was assessed to determine their relative predictive value in the model. Separate models of each risk factor were trained on all breast cancers and then applied to the subtypes (Table 25). MHT variables had a greater impact on discrimination for interval cancers than for screen-detected cancers (AUC = 0.56 [0.0031] vs 0.52 [0.0026]). Conversely, BMI had a greater impact on discrimination for screen-detected cancers than for interval cancers (0.53 [0.0026] vs 0.48 [0.0030]). The performance of the variables was comparable for both low and high grade breast cancer.

Table 25. Performance of risk prediction model by each breast cancer risk factor

Variables		AUC				
		All breast cancer	**Screen-detected**	**Interval**	**Low grade**	**High grade**
All		0.59	0.58	0.60	0.59	0.57
Deprivation		0.50	0.49	0.49	0.50	0.50
+	**Age menopause**	0.52	0.52	0.50	0.52	0.52
+	**Age menarche**	0.51	0.51	0.49	0.51	0.51
+	**Alcohol**	0.52	0.51	0.51	0.52	0.51
+	**Smoking**	0.52	0.51	0.50	0.52	0.51
+	**Oral contraceptive use**	0.51	0.49	0.49	0.50	0.51
+	**Parity**	0.52	0.52	0.52	0.52	0.51
+	**Family history**	0.53	0.52	0.52	0.53	0.53
+	**Breast lump**	0.52	0.51	0.53	0.52	0.52
+	**BMI**	0.52	0.53	0.48	0.52	0.53
+	**Age first birth**	0.52	0.53	0.53	0.53	0.52
+	**MHT status**	0.53	0.51	0.54	0.53	0.52
+	**MHT duration**	0.53	0.51	0.55	0.53	0.51
+	**MHT type**	0.53	0.51	0.54	0.54	0.52
+	**All MHT variables**	0.54	0.52	0.56	0.54	0.52
+	**All reproductive variables**	0.53	0.54	0.53	0.53	0.53

6.5.5 Restricting follow-up time

The model for all breast cancer was applied to cases diagnosed within 5 years of recruitment and within 10 years of recruitment. Within 5 years, there were 13,845 breast cancer cases, 6432 screen-detected breast cancers, 5127 interval breast cancers, 8676 low grade breast cancers, and 3411 high grade breast cancers diagnosed. Within 10 years, there were 32,995 breast cancer cases, 12,212 screen-detected breast cancers, 8640 interval breast cancers, 21,191 low grade breast cancers, and 8571 high grade breast cancers diagnosed. The performance of the overall breast cancer model for the follow-up periods and for the breast cancer subtypes is shown in Table 26. The performance of

the model did not improve but rather remained consistent across the time periods for all breast cancer and the breast cancer subtypes.

Table 26. Summary of performance of model based on all breast cancers applied to different subtypes and different time periods

	All cases		Cases up to 5 years post recruitment		Cases up to 10 years post recruitment	
	N	**AUC**	**N**	**AUC**	**N**	**AUC**
All breast cancer	57,189	0.59	13,845	0.58	32,995	0.58
Screen-detected	15,902	0.58	6432	0.57	12,212	0.57
Interval	11,662	0.60	5127	0.60	8640	0.60
Low grade	37,656	0.59	8676	0.59	21,191	0.59
High grade	15,078	0.57	3411	0.56	8,571	0.56

6.5.6 Time period specific models

Separate models were then developed for the risk of any breast cancer within 5 years of recruitment and any breast cancer within 10 years of recruitment, using the same method as described for the entire time period. These models were then tested in the breast cancer subtypes. The coefficients from the conditional logistic regression model for each time period are given in Appendix 8.

The performance of the 5-year model was markedly better when compared to the performance of the model for the whole MWS time period: AUC = 0.67 (0.0025) vs 0.59 (0.0013). The model performed comparably across both screen-detected and interval cancer populations (AUC = 0.66 [0.0037] vs 0.67 [0.0042]) and performed marginally better in the low grade population than in the high grade population (AUC = 0.67 [0.0032] vs 0.65 [0.0051]).

The performance of the 10-year model was slightly improved compared to that of the model for the whole MWS time period: AUC = 0.61 (0.0017) vs 0.59 (0.0013). The model performed slightly better for interval cancers than for the screen-detected cancers (AUC = 0.63 [0.0034] vs 0.61 [0.0028]). The

model also performed marginally better for low grade cancer than for high grade cancer (AUC = 0.61 [0.0022] vs 0.59 [0.0035]).

A summary of the time period specific model performances is given in Table 27. Models for each follow-up period were generated in all breast cancer cases and applied to subtypes.

Table 27. Summary of performance of models based on all breast cancers for different time periods applied to different subtypes

	All cases		**Cases up to 5 years post recruitment**		**Cases up to 10 years post recruitment**	
	N	**AUC**	**N**	**AUC**	**N**	**AUC**
All breast cancer	57,189	0.59	13,845	0.67	32,995	0.61
Screen-detected	15,902	0.58	6432	0.66	12,212	0.61
Interval	11,662	0.60	5127	0.67	8,640	0.63
Low grade	37,656	0.59	8,676	0.67	21,191	0.61
High grade	15,078	0.57	3,411	0.65	8,571	0.59

6.5.7 Subtype-specific models

Separate models were then developed for the subtypes using the same method described for all breast cancers. The coefficients from the conditional logistic regression model for each subtype are given in Appendix 9.

The ROC curves, histogram of the linear predictors of the cases and control, and proportion of cases within each quintile of linear predictor are displayed for all models. The subtype-specific models, in general, performed slightly better than the overall breast cancer model. The AUC for the screen-detected model was 0.58 (0.0025) (Figure 32) and for the interval cancer was 0.62 (0.0030) (Figure 35). The AUC for low grade breast cancer was 0.59 (0.0016) (Figure 38) and for high grade breast cancer was 0.57 (0.0026) (Figure 41). For all models, the histograms of the linear predictors of the cases and controls considerably overlapped.

Figure 32. ROC curve of risk prediction model of screen-detected breast cancer

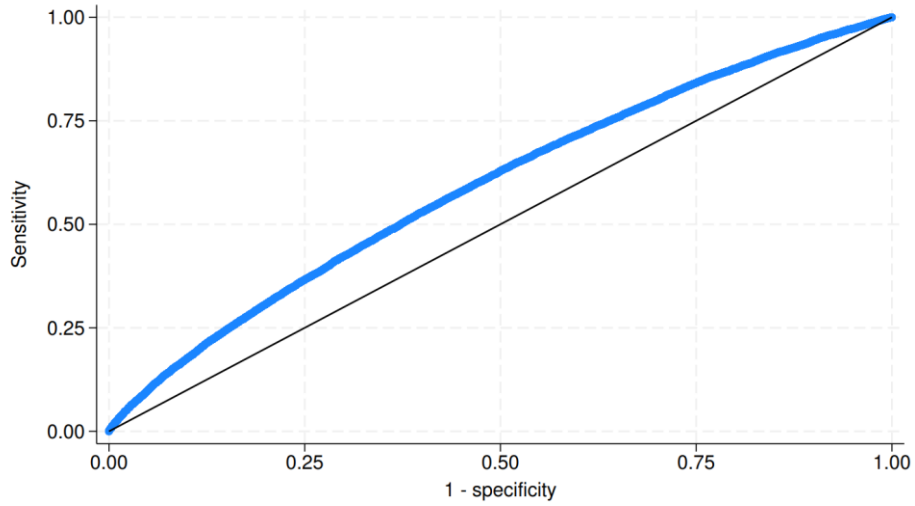


Figure 33. Histogram of linear predictor by screen-detected cases and controls

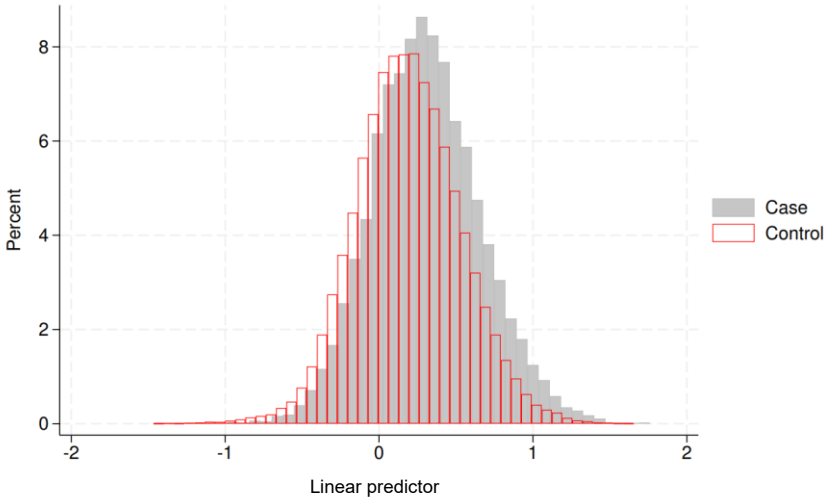


Figure 34. Proportion of screen-detected cases within each quintile of linear predictor

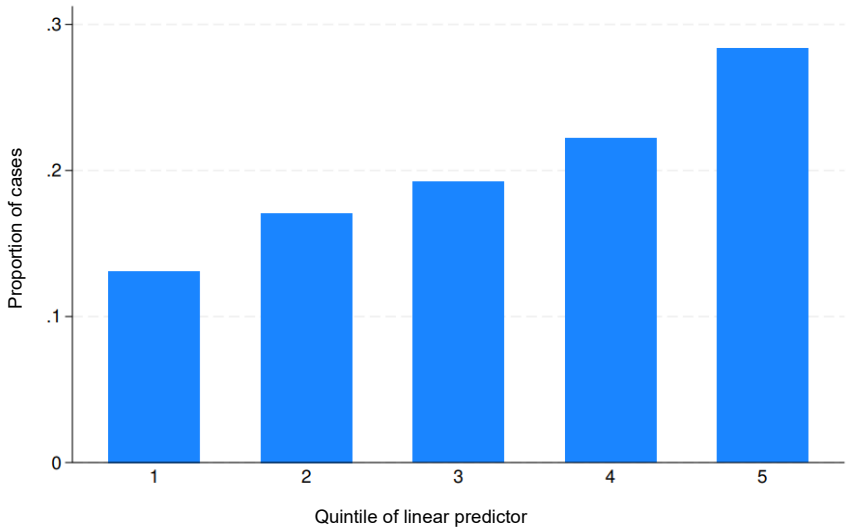


Figure 35. ROC curve of risk prediction model of interval breast cancer

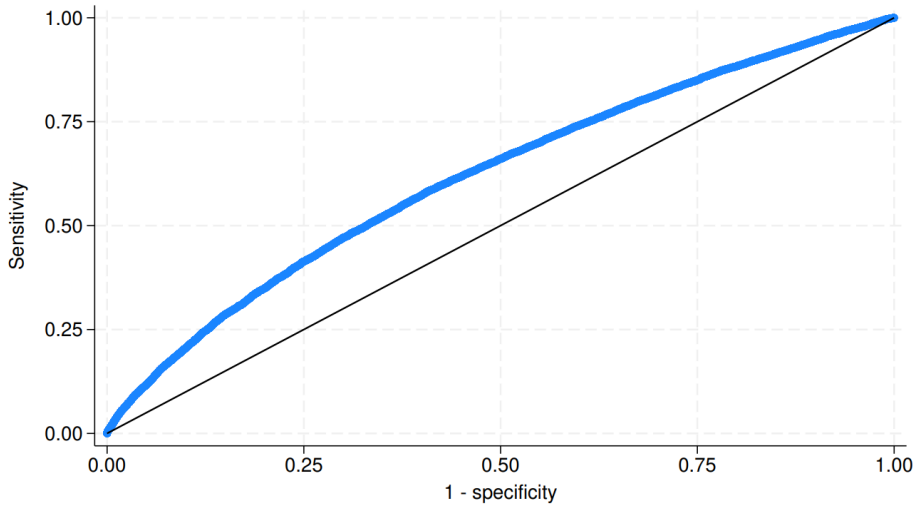


Figure 36. Histogram of linear predictor by interval cases and controls

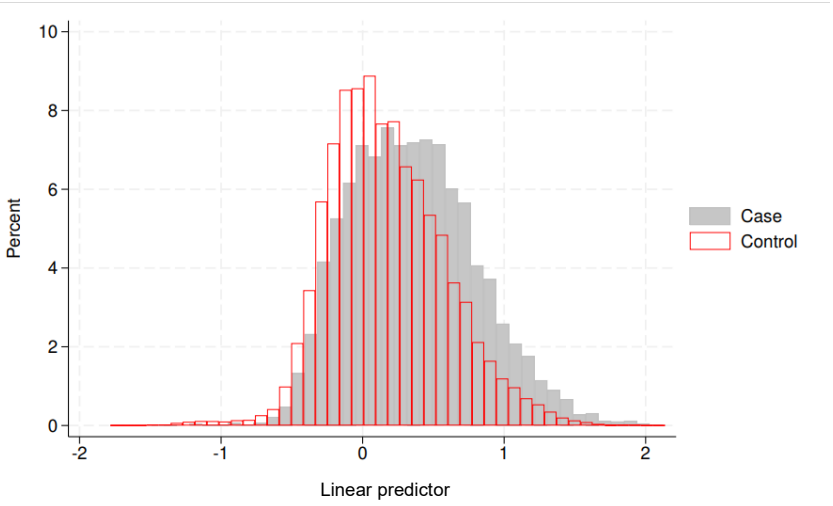


Figure 37. Proportion of interval cases within each quintile of linear predictor

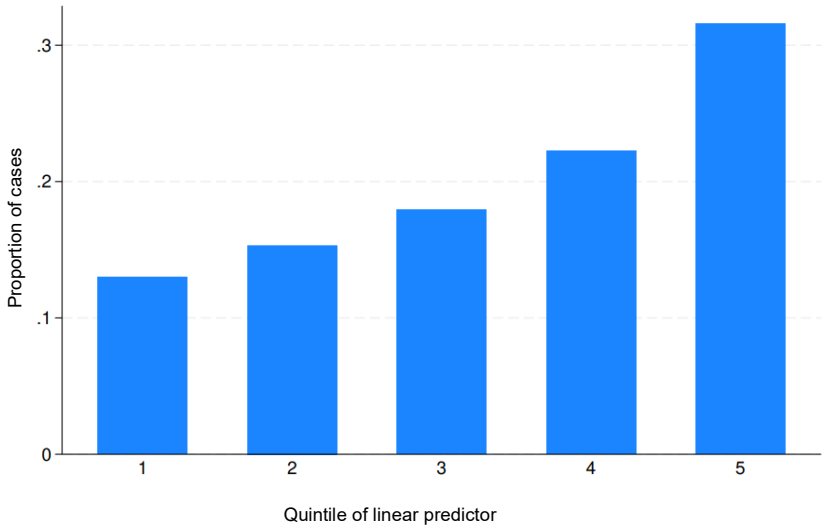


Figure 38. ROC curve of risk prediction model of low grade breast cancer

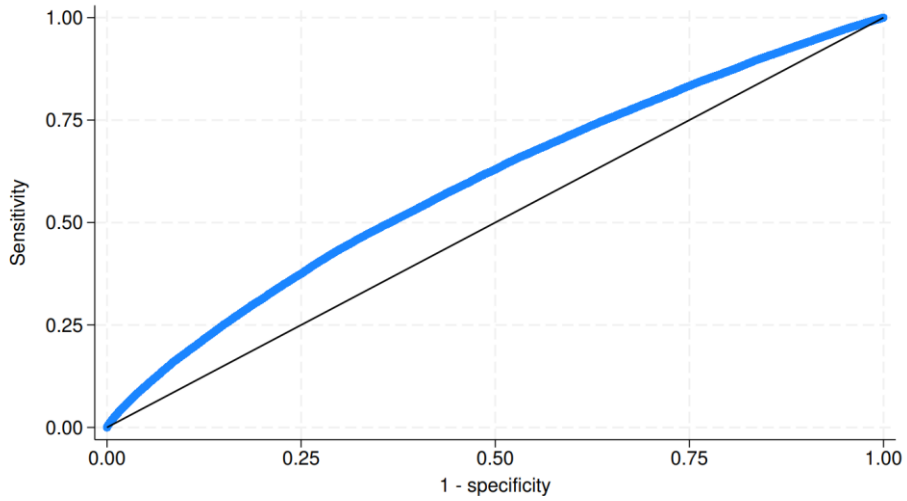


Figure 39. Histogram of linear predictor by low grade breast cancer cases and controls

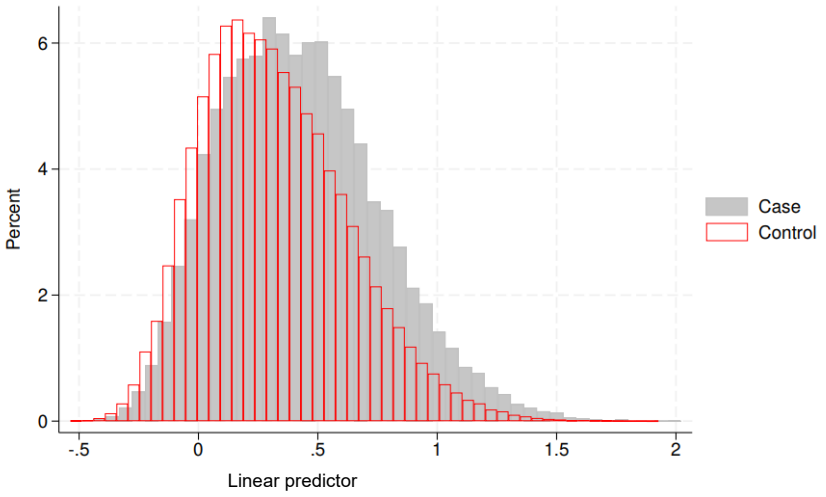


Figure 40. Proportion of low grade breast cancer cases within each quintile of linear predictor

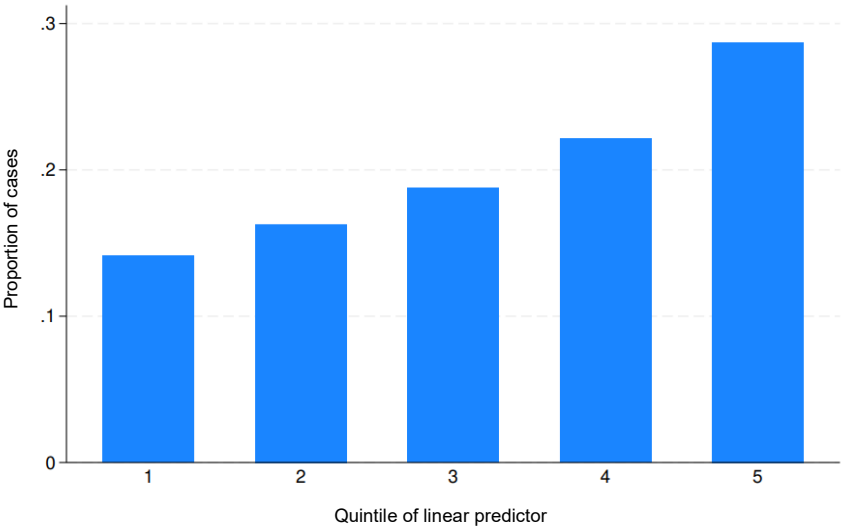


Figure 41. ROC curve of risk prediction model of high grade breast cancer

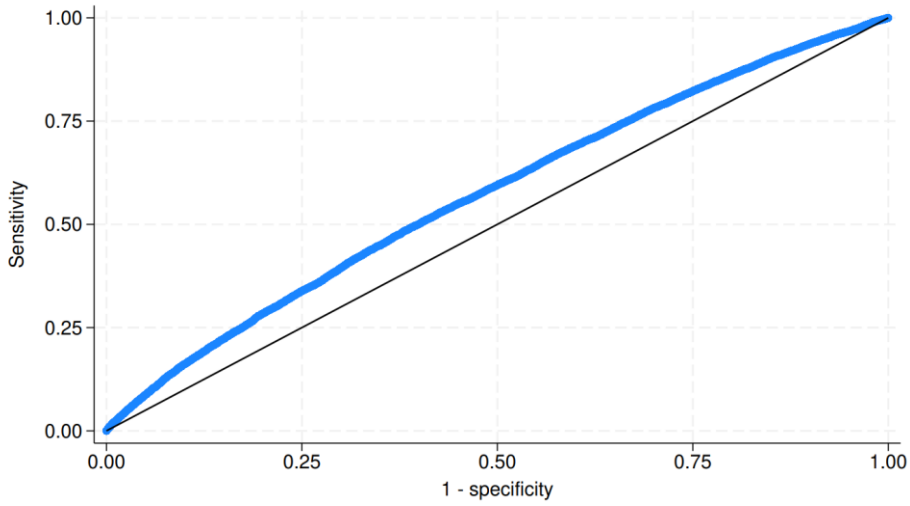


Figure 42. Histogram of linear predictor by high grade breast cancer cases and controls

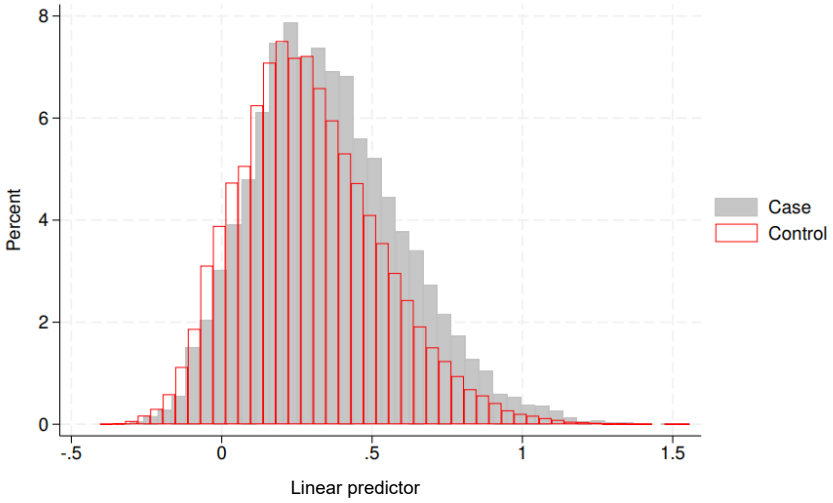
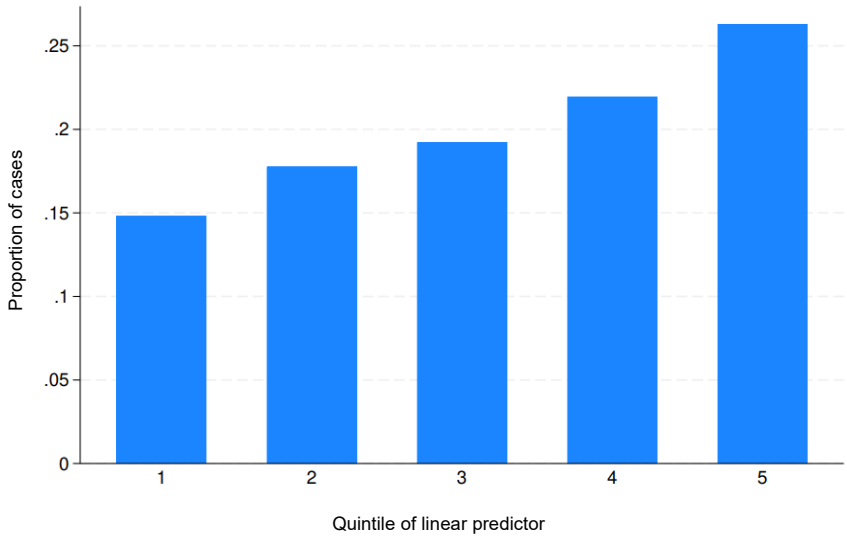


Figure 43. Proportion of high grade breast cancer cases within each quintile of linear predictor



6.5.8 Internal validation

Internal validation of all models generated in this analysis was conducted. Results from each of the 10 folds from the cross-validation are shown in Table 28. The AUCs from the 10 folds for all breast cancers ranged from 0.58 to 0.59, with a mean of 0.59. This is comparable with the AUC of 0.59 in the overall model generated earlier in this chapter. The little variation in AUCs across folds indicated no obvious suggestion of overfitting. The mean AUCs from the internal validation of the subtype and from the time period specific models were also comparable with the AUCs from the models generated earlier in the chapter.

Table 28. K-fold cross-validation of breast cancer risk prediction model and subtype-specific models

Fold	AUC						
	All breast cancer	**Screen-detected**	**Interval**	**Low grade**	**High grade**	**5y**	**10y**
1	0.59	0.59	0.62	0.59	0.58	0.66	0.60
2	0.58	0.59	0.61	0.59	0.56	0.66	0.60
3	0.58	0.57	0.63	0.60	0.57	0.67	0.60
4	0.59	0.60	0.61	0.59	0.57	0.67	0.61
5	0.59	0.60	0.62	0.59	0.56	0.67	0.62
6	0.58	0.59	0.60	0.59	0.57	0.68	0.60
7	0.58	0.59	0.61	0.59	0.58	0.66	0.61
8	0.59	0.59	0.61	0.59	0.57	0.68	0.61
9	0.58	0.59	0.62	0.60	0.56	0.66	0.60
10	0.59	0.59	0.62	0.59	0.58	0.68	0.60
Mean	**0.59**	**0.59**	**0.61**	**0.59**	**0.57**	**0.67**	**0.61**

6.6 Discussion

There are many models predicting breast cancer risk. However, there are fewer predicting specific breast cancer subtypes. The aim of the present chapter was to determine how well a breast cancer model for all breast cancers in a very large general screening population – with questionnaire data which could be easily collected in a clinical setting – could predict breast cancer subtypes based on mode of detection and grade of cancer, and whether any risk factors were more predictive of certain subtypes. The main findings suggest that models developed to predict any breast cancer tend to fare marginally better at predicting interval compared with screen-detected cancers, and at predicting low

grade compared with high grade tumours. However, the discrimination in all cases was moderate at best, as the AUCs typically ranged from 0.57 to 0.60. The models developed to predict specific subtypes did not appear to materially improve on discrimination of the overall model for any particular subtype. Perhaps not surprisingly, the models aimed at determining risk in the short-term (5 years) and moderate-term (10 years) had better discrimination than those aimed at determining risk over the entire follow-up period.

The overall model had an AUC of 0.59, indicating fairly poor discrimination. When examining the impact of each risk factor on the overall discrimination in turn, no single variable had much of an effect. AUCs ranged from 0.51 to 0.53. The effect of combining the MHT variables only very slightly increased the AUC to 0.54. Yet, given the relatively poor performance of all variables individually, it would suggest that there was no single driver of breast cancer risk from easily collectable questionnaire data. It is only when the variables are combined does the discrimination improve somewhat.

Based on the systematic reviews discussed earlier in the chapter, predictive performance is better with the use of MD data, mammographic images, and genetic and biomarker data and can be further improved with more complex machine learning tools analysing data from mammograms.[116,277] However, these are not always available in clinical settings, unlike questionnaire or risk factor data which are easily accessible. A recent systematic review from 2022 included 40 breast cancer risk prediction models published between 1989 and 2021.[116] Like the present study, most of the studies used in the 2022 systematic review used logistic regression to fit the prediction models. The most common risk factors used were reproductive factors, family history of breast cancer, and age. The performance of the MWS model was slightly worse than comparable studies, likely due to lack of MD data and limited family history data in the MWS. Many studies included age in the model, which likely contributed considerably to model discrimination. As I matched on attained age and age at most recent screen, I could not include this in the model; therefore, my model performance would have

been poorer than other models. Crispo et al. used age, age at menarche, number of breast biopsies, age at first birth, and family history in 558 cases and in 127 controls.[302] They reported an AUC of 0.57. Tice et al. used age, ethnicity, first-degree breast cancer, previous biopsies, and breast density in 14,766 cases, with 5.3 years of follow-up.[303] They reported an AUC of 0.66 and in their internal validation. This is comparable with the MWS 5-year model with an AUC of 0.67. Anothaisintawee et al. used age, menopausal status, BMI, and oral contraceptive use in 107 cases.[304] They had an AUC of 0.65 and, using bootstrapping in their internal validation, reported an AUC of 0.65. Maas et al. used age at menarche, menopause, age at first birth, parity, alcohol consumption, height, smoking status, BMI, family history, hormone therapy, and polygenic risk score in 17,171 cases.[305] The model with the polygenic risk score had an AUC of 0.64, and the model without the polygenic risk score (using risk factor information only) had an AUC of 0.59, comparable to the MWS. A model developed in the China Kadoorie Biobank used age, residence area, education, BMI, height, family history of cancer, parity, and age at menarche in 2287 cases over 10.2 years.[306] They reported an AUC of 0.64. In their external validation the AUC was 0.59. This model was developed in an Asian population, whereas the MWS is based around a predominantly Caucasian population. While the MWS model performed comparably or slightly worse than similar studies, it was one of the largest studies with 57,189 cases.

At the start of the chapter, I introduced more widely known breast cancer risk prediction models, some of which are used in clinical practice. The BOADICEA model is used clinically in the NHS for risk stratification. A recent validation of the BOADICEA model, with over 66,000 women, reported an AUC of 0.70, when including PRS and pathogenic variants.[307] This evidently performed better than the MWS model. Another study validated other breast cancer models in 35,921 women for the prediction of breast cancer over 6 years. They reported AUCs of: BRCAPRO, 0.61; Gail, 0.64; IBIS, 0.62.[308] These results are more comparable to the 5- and 10-year MWS models. It is clear that there is no perfect model. However, it is clear that more information than what was provided in the MWS is clearly needed, namely MD data and genetic data, or as a bare minimum, more detailed information on family history, as used by many of the widely known models.

The model for overall breast cancer was applied separately to the breast cancer subtypes. As the performance for each subtype was poor, I developed subtype-specific models. These models also performed relatively poorly. Interestingly, though, all performances were comparable. This may be because the risk factors available from the questionnaire data in the MWS are weak risk factors for the most common types of breast cancer. A summary of interval and screen-detected breast cancer risk prediction models in the literature is given in Section 6.2. The AUC for interval breast cancer ranged from 0.63 to 0.86, and for screen-detected cancer, from 0.53 to 0.64. Whilst most of the studies used logistic regression to generate the model, a few used neural networks and deep learning models to incorporate information from digital screening mammograms. All studies used MD in their model which was not available for the entire MWS cohort. Therefore, it is difficult to compare the performance of the models, as MD is a major predictor of breast cancer risk.

One might have expected that the risk factor data would better predict screen-detected cancers as these are usually the slower growing, more indolent cancers, which tend to have greatest association with the kinds of risk factors (e.g. reproductive factors and hormonal factors) which the MWS collected at recruitment. [309] The opposite was observed, although the difference was marginal. It is not clear why this may be although it may reflect the relatively greater effect of MHT use on interval as compared with screen-detected cancers which was seen in Chapter 5.

I additionally looked at predicting low and high grade breast cancers. The literature did not yield prediction models on grade. In fact, few breast cancer subtype models exist. Although the performance was still poor, the model was better at predicting low grade cancers when compared to higher grade cancers. This is likely because the MWS data primarily captured risk factors associated with hormone receptor-positive cancers, which tend to be lower-grade and have better prognoses. [309] MHT, for example, is associated with luminal type A and ER positive cancers which have a better prognosis than triple negative breast cancers.[310] Reproductive factors and MHT variables performed comparably for both grades. Unfortunately, due to incomplete receptor status data in the MWS,

predicting ER/PR/HER2 or triple-negative breast cancers was not feasible. A study with 35,921 women evaluated the performance of breast cancer models by tumour subtype. Interestingly, all models performed better for ER/PR positive tumours than for HER2+ or triple negative breast cancers. This is likely because the factors in the models are influence hormone levels and therefore are more predictive of hormone receptor positive tumours.[308] For example, the AUCs for the IBIS model were 0.63, 0.57, and 0.52, respectively. For BRCAPRO, they were 0.64, 0.51, and 0.55, respectively. In the WHI cohort of 147,916 women, the Gail model formed better for ER positive cancers than ER negative cancers (0.60 vs 0.58).[311]

For both grade and mode of detection, the individual variables in isolation had low predictive power. When breaking down the model and looking at risk factors separately, MHT better predicted interval cancers (AUC = 0.56 vs 0.52). As shown in Chapter 5, MHT is associated with an increased risk of interval cancers relative to screen-detected breast cancers. MHT increases MD and, therefore, reduces mammographic screening sensitivity, delaying the diagnosis of breast cancer. MHT is, thus, an important determinant of interval cancers. BMI was a poor predictor of interval cancers (AUC = 0.48). As shown in Chapter 4, BMI is inversely associated with MD. Therefore, screening is likely to be more sensitive in women with higher BMI, and so cancers in these women are less likely to be delayed. The remaining variables were otherwise not predictive of any of the subtypes. Only when the variables were combined could the model predict risk to some extent, albeit inadequately. As with the overall model, this suggests that there was no single driver of breast cancer subtype risk within these available data.

When the overall model was applied to shorter follow-up periods, the AUCs remained unchanged. However, after developing time period specific models, the performance of the model improved. This suggests that the recency of risk factor data collection is important in breast cancer risk prediction models. The improved performance is likely because the MHT status was more accurate and predictive of risk at the time of diagnosis during the 5-year model. This would be because the risk associated

with MHT use depends heavily on recency of use. Kerlikowske et al. used age, ethnicity, first degree breast cancer, previous biopsies, and changes in breast density in 13,715 cases.[312] They developed a 5- and 10-year model with AUCs of 0.64 and 0.63, respectively. The AUCs in the 5- and 10-year models in the MWS were 0.67 and 0.61, respectively.

The main strength of this chapter was the size of the study from which the model was developed. With almost 60,000 cases, it is one of the largest populations used to develop a breast cancer risk prediction model. The MWS was also larger than any previous study, which has generated risk prediction models for breast cancer subtypes defined by the mode of detection, with 15,902 screen-detected cases and 11,662 interval cases.

However, there are several limitations. The poor performance overall was likely driven by the lack of data on MD and digital mammographic images, genetics, and other biomarkers, as well as the limited information on family history. It is clear from the latest literature that these are required to produce a model with better predictive performance. Models which do not include genetic variances have AUCs in the range of 0.55-0.67, those with MD have a range of 0.62-0.68 and those that include both genetic information and MD range from 0.62-0.69.[313] As discussed at length in this thesis, MD is an important predictor of breast cancer, and it reduces screening sensitivity. Therefore, this would likely improve the predictive performance models for interval cancers. As MD is also associated with higher grade breast cancer, it would also be needed in grade-specific models.[52] Given that interval cancers (by definition) and high grade cancers tend not to be picked up at screening, models predictive of these subtypes are essential, to detect these cancers earlier. Furthermore, there was a degree of missing data which had to be imputed. However, the coefficients from the regression model were comparable with the literature and other MWS analyses. Using a complete case dataset would have reduced the cases to almost less than 30,000 and, therefore, would limit the power of the analysis. Another limitation was the lack of external validation. While internal K-fold cross-validation showed

no evidence of overfitting, external validation using an independent population is necessary to confirm the model's generalisability.

The decision to carry out a nested case-control study, as opposed to a cohort study, was largely because a smaller dataset made the analysis more computationally manageable. Had I used the entirety of the MWS, the analysis and model development would not be feasible for the timescale of a DPhil. Furthermore, most previous breast cancer risk prediction models referenced in section 6.2 also utilised this study design.

Given the poor performance of the all the models developed in this chapter, they cannot be used in clinical practice. Ideally, a model would predict interval or high-grade cancers to a degree of certainty to inform risk stratification and identify the women who could benefit from personalised screening pathways or preventative measures. Instead, the model showed that risk factor questionnaire data alone is not sufficient for breast cancer prediction and that other more discriminatory factors are necessary beyond self-reported questionnaire-based variables, especially in long-term risk. The improved performance in the 5-year model suggests that risk factor data needs updating when predicting breast cancer.

As I limited my validation to internal validation, future work could look at external validation in perhaps the UK Biobank, as many of the candidate predictors are available in both studies. However, I could only test the performance of the overall breast cancer model as there is no data on mode of detection or on breast cancer grade within the UK Biobank. Additionally, the classification of MHT is not as detailed in the UK Biobank as in the MWS. The MWS is currently in the process of receiving MD data for more women. Therefore, adding MD to the model should improve its predictive performance for both overall breast cancer and its subtypes.

6.7 Conclusions

Overall, the models based on questionnaire-based risk data performed relatively poorly and could not reliably predict breast cancer subtypes, especially for long-term risk. None of these risk factors were

notably able to predict breast cancer subtypes, with the slight exception of MHT for interval cancers. Further work with MD, genetic, biomarker and detailed family history data is necessary to improve its predictive performance before its use in risk stratification or personalised screening pathways.

Chapter 7. Discussion and conclusions

7.1 Introduction

This thesis aimed to elucidate the role of MD, alongside various other risk factors, in the development of breast cancer and their implications for breast cancer detection. This was done through 1) a systematic review and meta-analysis of breast cancer risk factors and MD; 2) a cross-sectional analysis of breast cancer risk factors in the MWS, a UK-based screening cohort; 3) a prospective analysis of the effects of MHT, an important determinant of MD, on mammographic screening sensitivity in the MWS; and finally 4) developing overall breast cancer and subtype-specific risk prediction models using easily collectable risk factor information from questionnaires.

The findings of this thesis contribute to our further understanding of how MD is associated with established breast cancer risk factors, suggesting that MD does not fully account for the association of most risk factors with breast cancer risk. The thesis has also investigated the effects of MHT on screening sensitivity, likely through its effects on MD, highlighting the need for further counselling or patient education upon prescribing MHT by a GP. The development of the risk prediction model highlights the limitations of questionnaire-based risk factor data in the discriminatory ability of breast cancer and subtype-specific prediction models.

In the following sections of this chapter, I will summarise the main findings of this thesis, reflect on the strengths and weaknesses of the thesis, discuss potential public health or clinical implications of the findings, and finally provide recommendations for future research.

7.2 Summary of main findings

7.2.1 Breast cancer risk factors and mammographic density in the literature

The systematic review and meta-analysis in Chapter 3 aimed to examine the associations between certain known breast cancer risk factors and MD, where these relationships were not well established (alcohol, smoking, parity, age at first birth, and age at menarche).

The review and meta-analysis indicated that these lifestyle and reproductive factors were associated with MD. Alcohol consumption, age at first birth, and age of menarche were positively associated with

MD, and smoking and parity were inversely associated with MD. Nevertheless, these associations were much smaller than those between other breast cancer risk factors with established associations with MD, namely age, menopausal status, and BMI. There was a large degree of heterogeneity across the studies for most of the meta-analyses. I attempted to explore this by investigating associations further by menopausal status, MHT use, levels of adjustment, location of study, and method of MD assessment. The meta-analyses by menopausal status suggested that the associations of some of these risk factors (alcohol, smoking, age at menarche) were greater in premenopausal women than in postmenopausal women. This may be due to the compounding or additive effects of the high levels of endogenous hormones. There was also a suggestion that the effects of alcohol and parity were greater in women on MHT than in never users. However, this was just based on two studies. For other potential sources of heterogeneity, the patterns were rather weak, likely because of the few studies in each category. This, therefore, warranted a further study in a population in which I could restrict menopausal status and MHT use, as well as control for potentially important confounders, such as BMI.

7.2.2 Breast cancer risk factors and mammographic density in the Million Women Study

The novel cross-sectional analysis of the associations of several breast cancer risk factors on MD in Chapter 4 allowed me to expand on the work in Chapter 3. In a population of roughly 8000 postmenopausal women, I was able to add to the existing body of evidence that several breast cancer risk factors are associated with MD. Current/recent MHT use, older age at first birth, alcohol consumption, physical activity, and benign breast disease were positively associated with MD. Age, parity, smoking, and BMI were inversely associated with MD. There was no association between MD and age at menarche, breastfeeding, oral contraceptive use, bilateral oophorectomy, family history of breast cancer, and time since menopause. The examination of risk factors, other than MHT, was in women who were not currently using MHT.

Reassuringly, the results in Chapter 4 were almost identical to the results from the meta-analysis. This suggested that the inclusion of premenopausal women and MHT users in most studies which

contributed to the systematic review did not materially affect the findings. After removing current/recent MHT users, BMI had the largest (inverse) association with MD. This is likely because the deposition of fatty tissue within the breast, as BMI increases, either displaces or reduces the overall proportion of dense tissue within the breast. After investigating the effect modification by BMI, on the association between other risk factors and MD, it was clear that the positive association between BBD and MD was more apparent in women with higher BMI than lower BMI. There was no evidence of effect modification by BMI for other risk factors.

Perhaps the most interesting association in both Chapter 3 and Chapter 4 was that current smoking, as compared with never smoking, was associated with lower MD. As discussed previously, the associations between smoking and breast cancer risk have been inconsistent. It is possible that an association in the screening population may be because cancers are easier to detect in smokers, given their reduced MD, compared to never smokers. A further novel finding was the positive association between physical activity and MD. Whilst this has not been observed in previous studies, the results in this thesis may be attributed to residual confounding by BMI, as women who exercise more will have lower BMI and, therefore, denser breasts. On the other hand, the results could be influenced by inflammation or hormones.

I could not examine the associations between the risk factors and MD by MHT use or menopausal status, given the lack of premenopausal women in the MWS and the few current MHT users. It would be interesting to explore the differential effects of risk factors in these populations to understand the effect modification of higher background sex hormone levels on MD.

It was not possible to perform a formal mediation analysis, given the few breast cancer cases within the population for which MD was available. The crude analysis suggested that MD may have partially mediated the associations for some risk factors: MHT use, alcohol consumption, parity, and age at first birth. Therefore, it is likely that these risk factors act through mechanisms beyond MD in breast

cancer development. However, as this was a very crude analysis, more definitive conclusions on the role of MD as a mediator cannot be made.

7.2.3 Menopausal hormone therapy and mammographic screening sensitivity

The aim of Chapter 5 was to understand how type and timing of MHT use affected mammographic screening sensitivity. This was done by determining the relative risk of interval breast cancers, compared to screen-detected cancers, as a proxy of screening sensitivity. The initial review of the literature and meta-analysis indicated that being an ever MHT or current MHT user, as opposed to a never MHT user, was associated with an increased risk of interval versus screen-detected cancer. Current MHT users had a 69% increased risk of having an interval versus screen-detected breast cancer compared to never users. This effect was largely confined to current users, and there was little evidence that it persisted among past users regardless of how long they had stopped using MHT. This was likely due to a few small studies looking at these associations. Therefore, these associations were further explored in the MWS, which had detailed information on MHT patterns of use (recency, type, and duration) and mode of detection. The analysis found that current users, compared to never users, had a 26% increased risk of having an interval versus screen-detected breast and that this increased risk declined after stopping MHT use but remained elevated after 10+ years post-cessation. This, therefore, suggests that normal x-ray mammography as a screening modality may not be as effective in women who are currently on or were previously on MHT. No clear associations were seen by type of MHT. There was no evidence that duration had an independent effect on risk once recency was considered.

As MHT is usually associated less aggressive cancers, these interval cancers seen in Chapter 5 are unlikely to represent fast-growing, aggressive cancers which developed after the screen, although it is possible that some cancers captured in the analysis were aggressive tumours. Furthermore, MHT is also more associated with lobular breast cancers, which are harder to detect at screening because they are more diffuse than ductal breast cancers.[275] This could conceivably contribute to a more persistent increase in the risk of interval vs screen-detected cancers in MHT users.

Interestingly, in Chapter 4 I found that there lacked evidence of an association between past MHT use and MD. However, in Chapter 5 there was evidence that screening sensitivity remained reduced even after ceasing MHT use. In Chapter 4, past users were largely women who had stopped using MHT over a decade before digital mammography was introduced. Therefore, I would not have been able to detect any short-term persistent effects of MHT on MD. Furthermore, due to the limited sample size in Chapter 4, compared to Chapter 5, it is possible the MD analysis was underpowered to detect a pattern seen in the MHT analysis of Chapter 5.

7.2.4 Breast cancer risk prediction

In the final results chapter of this thesis, Chapter 6, I aimed to develop a general breast cancer risk prediction model using information from easily collectable risk factor data. I then determined how well this model was able to predict breast cancer subtypes based on the mode of detection and breast cancer grade and additionally examined the effect of each risk factor on its ability to predict breast cancer subtypes.

Perhaps as expected, given the lack of MD, the overall breast cancer model performed fairly poorly. The model was marginally better at predicting interval compared with screen-detected breast cancers, and at predicting low grade compared with high grade breast cancers. The discrimination in the subtype-specific models was, in reality, no better than the overall model at predicting subtypes. Unsurprisingly, the model developed for predicting 5-year risk was better than models predicting 10-year risk or over the full follow-up period. This is likely because the risk associated with MHT use depends heavily on the recency of use. Compared to models used in clinical practice, the MWS model performed worse than the BOADICEA, likely due to the lack of genetic data, PRS scores and MD data, with an AUC of 0.70 and other models with an AUC greater than 0.60.

When assessing the impact of each risk factor on model discrimination, MHT (status, type and duration) had the largest impact on the overall model. This may explain why the subtype-specific model of interval cancers was slightly better than that for screen-detected cancers. MHT was a

stronger predictor of interval cancers due to its effect on screening sensitivity, as shown in Chapter 5. No other risk factor appeared to affect model discrimination. It was only once the risk factors were combined in the model that it had some discriminatory ability. This suggests that amongst the other risk factors available, none was a significant driver of breast cancer risk or predictive of subtype-specific risk in isolation.

This chapter has shown that prediction models based on easily available risk factor data are unlikely to be useful in clinical practice to determine breast cancer risk or subtype-specific risk. Ideally, one would want a model that would predict cancers that are missed at screening, namely interval cancers and high grade tumours. Then, more tailored screening practices or preventative measures can be implemented for the women at greater risk of these. The lack of MD in the model likely contributed to the overall poor utility of the models. An alternative view would be to have a generic model that was sensitive at predicting all breast cancers to then determine the optimal screening interval for all women, regardless of the cancer detected.

7.3 Strengths and limitations

As detailed in Chapters 3 to 6, this thesis has several strengths and limitations. An overview of the general strengths and limitations are discussed here.

7.3.1 Strengths

7.3.1.1 Large cohort

The main strength of the thesis was the use of the MWS, a large, robust and well-studied prospective cohort, with detailed and ongoing collection of risk factor data relevant to this thesis and, importantly, virtually complete follow-up since the mid-1990s. The MWS allowed for the study of MD, risk factors, and breast cancer in a real-world setting, making it arguably generalisable. The detailed information collected at recruitment and beyond allowed for the study of many risk factors and the ability to adjust for several potential confounders. The large sample size made it possible to detect subtle associations. The prospective designs limited recall bias associated with retrospective recording of exposures. The repeat questionnaires meant that relevant information on risk factors that may change could be

updated from the recruitment survey. The linkage to cancer registries and other external data sets through NHS England enabled a comprehensive ascertainment of relevant outcome data.

7.3.1.2 Detailed exposure data

The MWS questionnaires asked detailed questions on MHT use, as this was the primary exposure of interest when the cohort was conceived by the late Professor Dame Valerie Beral. The ability to determine status, recency, type, and timing meant that these could be investigated as more detailed exposures rather than simply ever/never use. This allowed for a more novel investigation into MHT patterns of use and breast cancer detection at screening, which had not been thoroughly explored previously. Coupled with the power of the MWS, this thesis yielded new and pertinent findings that could have clinical and public health implications.

7.3.2 Limitations

7.3.2.1 Bias

As with most epidemiological or observational studies, there are methodological factors to consider, particularly sources of bias. Bias could systematically influence the results in a certain direction.

Selection bias is a potential limitation given that the women who agree to attend screening and participate in the MWS may be systematically different from the general population in terms of health-seeking behaviours, lifestyle choices, and potential background morbidity. If these characteristics are also related to breast cancer risk, the results may not be representative of the target population about which I have made conclusions. However, the MWS study included 1 in 4 women in the UK born between 1935 and 1950, and the characteristics of the study population were broadly representative of the target population within the general UK population at the time of recruitment.[118,123]

A further limitation is that the MWS relied on recall to collect exposure data. As risk factor data was obtained through self-administered questionnaires, relying on memory, collected information may have some inaccuracies. If participants incorrectly recorded their historical MHT use, reproductive, or lifestyle factors, there could be misclassification of exposure or residual confounding depending on the use of the variable in the analysis. This potential misclassification, though, was likely random

rather than systematic. It was also limited by the repeat questionnaires during the MWS, and therefore, risk factor information was updated so that the most recent documented observation could be used. Furthermore, several of the variables have been validated. Good agreement has been found between the reported information on hormone use, anthropometric factors, and reproductive factors and the validated objective measures or health records.[118] Further residual confounding may be possible for other health, lifestyle, and familial factors not collected on the surveys.

7.3.2.2 Generalisability

Generalisability or external validity is the extent to which the findings can be applied to the general population beyond the context of the study.274 As mentioned, the NHSBSP invites women every three years for screening. Most other population-based screening programmes invite women every two years. Therefore, women in the UK have a greater time period over which an interval cancer can be diagnosed. Hence, the screening-based results in this thesis may not be fully generalisable to other populations with biennial screening.

Breast screening eligibility and practices have changed since the inception of the MWS. The upper age limit for screening was initially 64 years old but was extended to 70 years old after the recruitment in 2004. At present, there is ongoing research into the effects of extending screening further from 47 years to 73 years old.[315] Similarly, initial mammograms were film-based and digital mammography was introduced some 10-15 years after recruitment. These changes will likely influence screening abilities and overall detection rates. This may impact outcomes related to breast cancer detection and MD measurements, potentially leading to differences from what was observed in this thesis.

Furthermore, the MWS does not represent an ethnically diverse population compared to the present general UK population. Given that ethnicity is a predictor of lifestyle behaviours, risk factors, disease risk, health-seeking behaviours, access to health services, and MD patterns, the results here may not be generalisable. Ethnicity and access to healthcare could modify the relationships between screening update, MD, breast cancer risk, and screening sensitivity. 96% of women in the MWS were of white ethnicity, which is similar to the figure at the time of the 2001 census (91% white) in England, but

higher than what was reported in the 2021 census (81%).[316] Therefore, whilst the results would be valid for the population around the time of recruitment (1996-2001), they may not be generalisable to a more ethnically diverse population, as seen in present-day England.

As the participants were born around World War 2, their lifestyle habits, exposures, and reproductive habits will broadly differ from women born after this period. Women in the MWS were likely to have children earlier, have more children, and have a later age at menarche compared to women who are going through menopause now or women who are now eligible for screening. Whilst the relative risk estimates based on internal comparisons should still be valid in the current population, there may still be differences between the women in the study and women who attend screening now.

7.3.2.3 Limited mammographic density data

One of the main limitations was the limited availability of MD data. At the start of my thesis, the initial aim was to have data from 500,000 mammograms in the MWS. However, as breast screening was paused in 2020 due to the COVID-19 pandemic, the collection of MD data was also delayed. MD data became available in 2023 for a subset of women (~10,000) who attended screening in the Oxford screening centre between 2011 and 2018. Had I access to more MD measurements, I could have determined the associations between MD and breast cancer risk in the MWS, conducted a more formal mediation analysis, adjusted for MD in the MHT and screening sensitivity analysis to determine how much of the results were attributable to MD, and finally used MD in the risk prediction models, which would have almost certainly improved the predictive accuracy and therefore the clinical utility of the models.

7.4 Clinical and public health implications

The findings of this thesis have potentially important clinical and public health implications, particularly concerning MHT prescribing, breast cancer risk assessment, and screening practices.

7.4.1 MHT prescribing and counselling

MHT prescribing in the UK declined after the early publications of MWS on associations of MHT use with breast cancer risk in the early 2000s.[310] However, over the past several years, demand for MHT

has increased dramatically to the extent that supply could not meet demand.[317] With approximately 2 million current users in the UK, MHT use has risen from 11% to 15% of women aged 45 to 64 in England, between 2021 and 2023.[240] The rise in use may lead to a long-term increase in breast cancer incidence and false negatives at breast screening. The increased demand is likely due to celebrity endorsement, advocating for the benefits of MHT, increased awareness of MHT and menopause, prepayment options for MHT making it cheaper to access by hundreds of pounds each year, and finally because the retirement age is increasing, women are working for longer. MHT can relieve symptoms that a woman may otherwise find disruptive at work.

For women experiencing the unpleasant symptoms of menopause, such as hot flushes, night sweats, sleep disorders, mood changes, and vaginal dryness, MHT may offer a cheap and accessible solution. Given that MHT relieves the symptoms of menopause fairly rapidly after initiating treatment, an individual may be more inclined to address and manage current issues versus the possibility of developing cancer in the future.

Whilst GPs may counsel women on breast cancer risks associated with MHT use, women may believe that attending breast screening may mitigate some of these risks. This thesis has shown that current MHT use increases the risk of interval versus screen-detected breast cancer by 26% compared to never users and that this risk remains elevated even 10 years after stopping use. Therefore, women may need to be counselled on the limitations of screening and breast cancer detection if they are currently taking MHT use or have previously taken it. This may be more pertinent if they are at higher risk of breast cancer, such as having a family history of breast cancer. GPs may also need to explain that MHT effectively changes the appearance of an x-ray mammogram, which masks the presence of cancer and, therefore, leads to a greater chance of a false negative screen. This will delay the diagnosis of breast cancer. Such consultations will need to discuss the benefits of MHT against the potential risks of breast cancer and false negative screens. This will ensure that women can make informed decisions

about MHT use. Where women decide to use MHT, education and encouragement of self-examination may be beneficial.

7.4.2 Advising on lifestyle modification

The recognition of lifestyle risk factors that may influence MD may guide lifestyle modification programmes into clinical practice to improve the sensitivity and detection of breast cancer at screening. This thesis highlighted that alcohol and physical activity were associated with higher MD, and smoking and higher BMI were associated with lower MD. Evidently, healthcare professionals cannot advise women to reduce their physical activity, maintain a high BMI, or take up smoking to improve screening sensitivity. However, women who attend screening may need to be advised about the effects of alcohol on MD and, therefore, potentially on screening sensitivity. This may be more important for premenopausal women or women using MHT, given the suggestion from the meta-analysis that the effect of alcohol on MD may be greater in these women. Nevertheless, given that I was not able to explore the associations of alcohol by MHT use or menopausal status, this will have to be studied further. Furthermore, the associations were small and therefore, the absolute effect of alcohol on screening sensitivity would also need to be quantified.

7.4.3 Personalised breast cancer screening

Chapters 3 and 4 highlighted the risk factors associated with higher MD and Chapter 5 evidenced that MHT use was associated with a higher risk of interval cancer compared to screen-detected breast cancers. It, therefore, may be warranted that women with several risk factors that are associated with MD or interval cancer risk may require either more regular breast screening (i.e. every two years instead of three years) or may require more sensitive imaging modalities, such as an MRI. Currently, the UK employs screening every three years, whereas most other population-based screening programmes have biennial screening, as discussed in Chapter 5. [71,72,200]

A recent study within the MWS examined the associations between breast cancer risk factors, independent of MHT use, with interval and screen-detected breast cancer. The authors found that having a first-degree relative with a history of breast cancer or having previous BBD increased the risk

of interval versus screen-detected breast cancer (21% and 38% increased risk, respectively).[173] An inverse association was seen with increasing BMI and current smoking. These associations are essentially in keeping with what was found in my MWS analysis of risk factors and MD.

Moving forward, there may be a case for arguing that women with several risk factors associated with higher MD and increased risk of interval versus screen-detected breast cancers should be offered a more targeted screening programme. This, however, would be a costly enterprise and would require extensive health economic evaluations to determine whether the benefits of detecting cancers earlier would outweigh the cost of more frequent screens or more expensive imaging modalities. This would also require extensive collection of risk factor information prior to screening to determine who would be eligible for more tailored programmes. It is well acknowledged that the NHS and GPs are already overworked and understaffed. Therefore, deciding who would collect the information and where it would be stored would need to be addressed and agreed upon.

In the United States, it is a requirement to inform a woman who has a mammogram if she has "not dense" or "dense" breasts.[114,318] She then will have the opportunity, if she wishes, to discuss this with a healthcare provider to determine further options. Currently, the NHSBSP does not report MD. Given that the UK has a publicly funded health system, such a policy may not be feasible as unless there was government funding, only those with access to private health care would be able to act on the information. This would lead to inequitable care based on socioeconomic status. Furthermore, if women were informed, it would likely induce anxiety due to the fear of potentially having a cancer, which would be even worse if no further supplementary screening was readily available.

7.4.4 Breast cancer and subtype risk assessment

The thesis has shown that questionnaire-based risk factor information is insufficient in terms of predicting overall breast cancer risk and has additionally highlighted that none of these risk factors can predict breast cancer subtypes, with the small exception of MHT for interval cancers. It is also not clear how well better-performing models can predict breast cancer subtypes. Therefore, models will

likely need MD data, genetic information, biomarker data, and machine learning-based analysis of mammographic images to improve model performance. Risk prediction may become increasingly utilised with the increasing use of artificial intelligence and machine learning models.

7.5 Recommendations for future research

In addition to the suggestions made earlier in this chapter, in order to address existing gaps made apparent by this thesis, to improve prevention, screening, and intervention strategies, future research will need to:

- expand on the investigation of MD in relation to breast cancer risk factors and breast cancer risk in large and more diverse populations;
- improve risk prediction;
- and further critically understand new imaging and risk prediction technologies as they are being developed and introduced into clinical practice.

7.5.1 Mammographic density data in the Million Women Study

As mentioned previously, I had initially aimed to have access to MD data on 500,000 mammograms in the MWS. The MWS is in the process of receiving more MD data on women attending other screening centres beyond Oxford. Once more data is available, further MD analyses will be possible, including studying the relationship between MD and breast cancer, conducting a formal mediation analysis, and examining more detailed patterns of MHT use (recency and duration) and its associations with MD. Examination of other risk factors could also be stratified by MHT use. Furthermore, MD can be used as a candidate predictor in breast cancer risk prediction models and subtype-specific models, which will likely improve their discriminatory abilities.

7.5.2 Prospective study of mammographic density

The studies in the systematic review and the MWS used a cross-sectional design to determine the associations between breast cancer risk factors and MD. Future prospective studies should track changes in MD over time. This will help establish how changes in MD correspond to changes in risk factor status. For example, understanding the duration of increased risk and the time course for

changes in MD after cessation of alcohol or other lifestyle modifications could inform guidelines and health advice initiatives surrounding screening. This will help elucidate the dynamic nature of these associations over time and possible causal relationships. Furthermore, future research will need to determine how the effects of risk factors on MD correspond to clinically relevant changes in breast cancer risk and the absolute impact on screening sensitivity.

7.5.3 Modern and diverse populations

As discussed earlier, the MWS represents a cohort of women born around World War 2, lacking diversity in terms of ethnicity and possibly other socioeconomic features in the current UK population. To ensure generalisability, future studies investigating MD and breast cancer should focus on more ethnically diverse populations, representing the demographics of women who currently attend a screening or will be approaching screening age. Furthermore, examining the role of MD and other risk factors in women who are younger than those in the MWS, with different reproductive and lifestyle patterns, would enhance the external validity of the results. This would also enable the examination of breast cancer risk factors by menopausal status. Current and new risk prediction models will need to be validated across diverse populations to ensure their applicability and effectiveness in different demographic groups. This will hopefully help identify specific risks and patterns associated with various ethnicities.

7.5.4 Risk prediction

In addition to incorporating MD into the prediction models, combining the questionnaire data with genetic information, biomarker data and machine learning analysed information from mammograms would also improve the predictive accuracy of the models. Since 2006, blood samples have been collected from about a 5% sample of women in the MWS for genetic and biochemical analyses [118] There is no overlap between the women whose blood samples are available and those whose MD is available. Once more MD measurements are available, the MWS may be able to utilise both MD and biomarker or genetic data in the risk prediction models to improve model performance. These refinements could improve personalised risk assessments and inform breast cancer prevention or

early detection strategies. Ideally, one would want to generate models that can predict cancers that are not detected at screening (interval cancers or high grade cancers). This may allow earlier detection of these cancers to improve prognosis and allow for less invasive management options. Therefore, future research will need to refine risk assessment models for more aggressive subtypes.

7.5.5 Pathophysiological mechanisms

Future studies investigating the pathophysiology of breast cancer, including more formal mediation analyses, will also be necessary to elucidate the biological pathways through which MD contributes to breast cancer risk. This would include the hormonal, genetic, and molecular pathways influencing MD. This may also help identify potential therapeutic targets for women with high MD or identify new biomarkers for early detection or risk assessment.

7.5.6 Developing technologies and artificial intelligence

Developments in imaging tools and programmes for measuring MD will be necessary to improve current screening practices and to enhance MD assessment. This would include evaluating the cost-effectiveness and sensitivity of integrating modalities such as digital breast tomosynthesis, which produces 3-dimensional x-ray images, and MRI.

There has been extensive interest in integrating artificial intelligence (AI) and machine learning into healthcare. Machine learning is a subset of AI focused on the study and development of algorithms that learn and understand patterns from data to make predictions.[319] In breast cancer research, AI has been used for reading screening mammograms to understand whether these methods enhance detection or are at least not inferior to radiologist-read scans,[320–322] measuring and quantifying MD (used in Chapter 4) and developing risk prediction models (discussed in Chapter 6).[277] Further research using large, robust and prospective data will be necessary to understand the utility and safety of such tools. AI and machine learning algorithms can provide helpful predictive information by processing data on risk factors, characteristics, demographics, MD, and other radiological features from digital mammograms. Once validated, these can provide valuable insights into a patient's risk profile and guide decision-making regarding screening and preventative measures. However, factors

such as overfitting must be considered as models developed from machine learning can sometimes perform well for the training dataset but are not generalisable to new datasets.[323] Finally, understanding and managing public perception of the use and acceptability of AI in healthcare will be paramount. The public may feel wary about having less "human" input into their clinical decision-making or may have concerns surrounding trust and privacy regarding their data usage. Ultimately, AI has vast potential in breast cancer clinical practice, provided it is fully researched and understood. This field is evidently beyond the remit of this thesis but may possibly be the future of breast cancer detection and risk assessment.

7.6 Conclusions

This thesis has provided insights into the interplay between mammographic density and other breast cancer risk factors in the development and detection of breast cancer. By systematically reviewing the literature and conducting novel analyses using data from the Million Women Study, I have confirmed and expanded upon existing knowledge regarding the role of MD and other breast cancer risk factors on breast cancer risk and detection at screening. Identifying those at greater risk of breast cancer or of having their cancers missed at screening will ultimately guide future public education decisions, prescribing practices, and, after further economic evaluation, possible personalised screening strategies.

Continued research in this area, particularly in diverse populations and prospective cohorts, is vital to understanding the role of MD further. This will guide screening practices and help refine risk assessment tools with more predictive determinants of overall and subtype risk. This will ultimately help with the earlier detection and possible prevention of breast cancer to reduce the overall burden of the disease.

References

1. McCormack, V. A. & Dos Santos Silva, I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiology Biomarkers and Prevention* **15**, 1159–1169 (2006).

2. Weigel, S., Heindel, W., Heidrich, J., Hense, H. W. & Heidinger, O. Digital mammography screening: sensitivity of the programme dependent on breast density. *Eur Radiol* **27**, 2744–2751 (2017).

3. Arnold, M. *et al.* Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast* **66**, 15–23 (2022).

4. Cancer Research UK. Breast cancer statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer.

5. Torre, L. A., Siegel, R. L., Ward, E. M. & Jemal, A. Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev* **25**, 16–27 (2016).

6. Cancer Research UK. Risk factors for breast cancer. https://www.cancerresearchuk.org/about-cancer/breast-cancer/risks-causes/risk-factors (2023).

7. McPherson, K., Steel, C. M. & Dixon, J. M. Breast cancer - epidemiology, risk factors, and genetics. *Br Med J* **309**, 1003–1007 (1994).

8. Łukasiewicz, S. *et al.* Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review. *Cancers (Basel)* **13**, (2021).

9. Beral, V. *et al.* 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* **358**, 1389–1399 (2001).

10. Roberts, E., Howell, S. & Evans, D. G. Polygenic risk scores and breast cancer risk prediction. *The Breast : Official Journal of the European Society of Mastology* **67**, 71 (2023).

11. Key, T. J. *et al.* Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst* **94**, 606–616 (2002).

12. Endogenous Hormones and Breast Cancer Collaborative Group. Sex hormones and risk of breast cancer in premenopausal women: A collaborative reanalysis of individual participant data from seven prospective studies. *Lancet Oncol* **14**, 1009–1019 (2013).

13. Hamajima, N. *et al.* Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol* **13**, 1141–1151 (2012).

14. Collaborative Group on Hormonal Factors in Breast Cancer. Type and timing of menopausal hormone therapy and breast cancer risk: individual participant meta-analysis of the worldwide epidemiological evidence. *Lancet* **394**, 1159–1168 (2019).

15. Calle, E. E. *et al.* 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. *Lancet* **347**, 1713–1727 (1996).

16. 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* **21**, (2019).

17. Trichopoulos, D. *et al.* Age at any birth and breast cancer risk. *Int J Cancer* **31**, 701–704 (1983).

18. Beral, V., Bull, D., Doll, R., Peto, R. & Reeves, G. Breast cancer and breastfeeding: Collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302

women with breast cancer and 96 973 women without the disease. *Lancet* **360**, 187–195 (2002).

19. Guo, W., Key, T. J. & Reeves, G. K. Adiposity and breast cancer risk in postmenopausal women: Results from the UK Biobank prospective cohort. *Int J Cancer* **143**, 1037 (2018).

20. Bhardwaj, P. *et al.* Estrogens and breast cancer: mechanisms involved in obesity-related development, growth and progression. *J Steroid Biochem Mol Biol* **189**, 161 (2019).

21. García-Estévez, L. *et al.* Obesity and Breast Cancer: A Paradoxical and Controversial Relationship Influenced by Menopausal Status. *Front Oncol* **11**, (2021).

22. Hamajima, N. *et al.* Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Cancer* **87**, 1234–45 (2002).

23. National Cancer Institute. Physical Activity and Cancer Fact Sheet. https://www.cancer.gov/about-cancer/causes-prevention/risk/obesity/physical-activity-fact-sheet#what-is-known-about-the-relationship-between-being-sedentary-and-the-risk-of-cancer (2020).

24. Guo, W., Fensom, G. K., Reeves, G. K. & Key, T. J. Physical activity and breast cancer risk: results from the UK Biobank prospective cohort. *Br J Cancer* **122**, 726–732 (2020).

25. Lynch, B. M., Neilson, H. K. & Friedenreich, C. M. Physical activity and breast cancer prevention. *Recent Results Cancer Res* **186**, 13–42 (2011).

26. Hartmann, L. C. *et al.* Benign Breast Disease and the Risk of Breast Cancer. *https://doi.org/10.1056/NEJMoa044383* **353**, 229–237 (2005).

27. Tice, J. A. *et al.* Benign Breast Disease, Mammographic Breast Density, and the risk of Breast cancer. (2013) doi:10.1093/jnci/djt124.

28. Sheldahl. Breast anatomy. *Wikimedia Commons* https://commons.m.wikimedia.org/wiki/File:Breast_anatomy.png (2016).

29. Garrido-Estepa, M. *et al.* Evaluation of mammographic density patterns: reproducibility and concordance among scales. *BMC Cancer* **10**, 485 (2010).

30. Wolfe, J. N. Breast patterns as an index of risk for developing breast cancer. *AJR Am J Roentgenol* **126**, 1130–1139 (1976).

31. Boyd, N. F. *et al.* Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. *J Natl Cancer Inst* **87**, 670–675 (1995).

32. Gram, I. T., Funkhouser, E. & Tabár, L. The Tabár classification of mammographic parenchymal patterns. *Eur J Radiol* **24**, 131–136 (1997).

33. Copp, T. *et al.* Marketing empowerment: how corporations co-opt feminist narratives to promote non-evidence based health interventions. *BMJ* **384**, e076710 (2024).

34. Melnikow, J. *et al.* Supplemental Screening for Breast Cancer in Women with Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. *Ann Intern Med* **164**, 268 (2016).

35. Goddi, A., Bonardi, M. & Alessi, S. Breast elastography: A literature review. *J Ultrasound* **15**, 192 (2012).

36. Dense Breasts - Diagnosis, Evaluation and Treatment. https://www.radiologyinfo.org/en/info/dense-breasts.

37. National Cancer Institute. Dense Breasts: Answers to Commonly Asked Questions. https://www.cancer.gov/types/breast/breast-changes/dense-breasts.

38. Bodewes, F. T. H., van Asselt, A. A., Dorrius, M. D., Greuter, M. J. W. & de Bock, G. H. Mammographic breast density and the risk of breast cancer: A systematic review and meta-analysis. *Breast* vol. 66 62–68 Preprint at https://doi.org/10.1016/j.breast.2022.09.007 (2022).

39. Boyd, N. F. *et al.* Mammographic density and the risk and detection of breast cancer. *N Engl J Med* **356**, 227–36 (2007).

40. Wolfe, J. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer* **37**, (1976).

41. Byrne, C. *et al.* Mammographic Features and Breast Cancer Risk: Effects With Time, Age, and Menopause Status. *JNCI: Journal of the National Cancer Institute* **87**, 1622–1629 (1995).

42. Byrne, C. *et al.* Mammographic features and breast cancer risk: effects with time, age, and menopause status. *J Natl Cancer Inst* **87**, 1622–1629 (1995).

43. Boyd, N. F. *et al.* Mammographic density and the risk and detection of breast cancer. *N Engl J Med* **356**, 227–36 (2007).

44. Mokhtary, A., Karakatsanis, A. & Valachis, A. Mammographic density changes over time and breast cancer risk: A systematic review and meta-analysis. *Cancers (Basel)* **13**, 4805 (2021).

45. Farrar, M. C. & Jacobs, T. F. *Tamoxifen*. (2025).

46. Cuzick, J., Warwick, J., Pinney, E., Warren, R. M. L. & Duffy, S. W. Tamoxifen and Breast Density in Women at Increased Risk of Breast Cancer. *JNCI: Journal of the National Cancer Institute* **96**, 621–628 (2004).

47. Eriksson, M., Czene, K., Conant, E. F. & Hall, P. Use of Low-Dose Tamoxifen to Increase Mammographic Screening Sensitivity in Premenopausal Women. *Cancers (Basel)* **13**, 302 (2021).

48. Cuzick, J. *et al.* Tamoxifen-Induced Reduction in Mammographic Density and Breast Cancer Risk Reduction: A Nested Case–Control Study. *JNCI: Journal of the National Cancer Institute* **103**, 744–752 (2011).

49. Jung, S. *et al.* Endogenous Sex Hormones and Breast Density in Young Women. *Cancer Epidemiology Biomarkers and Prevention* **24**, 369–378 (2015).

50. Jung, S. *et al.* Adolescent endogenous sex hormones and breast density in early adulthood. *Breast Cancer Research* **17**, 1–9 (2015).

51. Martin, L. J. & Boyd, N. F. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. *Breast Cancer Research 2008 10:1* **10**, 1–14 (2008).

52. Boyd, N. F., Martin, L. J., Yaffe, M. J. & Minkin, S. Mammographic density and breast cancer risk: Current understanding and future prospects. *Breast Cancer Research* **13**, 1–12 (2011).

53. Wanders, J. O. P. *et al.* The effect of volumetric breast density on the risk of screen-detected and interval breast cancers: A cohort study. *Breast Cancer Research* **19**, 1–13 (2017).

54. Pettersson, A. *et al.* Mammographic density phenotypes and risk of breast cancer: a meta-analysis. *J Natl Cancer Inst* **106**, (2014).

55. Shepherd, J. A. *et al.* Volume of Mammographic Density and Risk of Breast Cancer. *Cancer Epidemiol Biomarkers Prev* **20**, 1473 (2011).

56. Burton, A. *et al.* Mammographic density and ageing: A collaborative pooled analysis of cross-sectional data from 22 countries worldwide. *PLoS Med* **14**, 1–20 (2017).

57. Alexeeff, S. E. *et al.* Reproductive Factors and Mammographic Density: Associations among 24,840 Women and Comparison of Studies Using Digitized Film-Screen Mammography and Full-Field Digital Mammography. *Am J Epidemiol* **188**, 1144–1154 (2019).

58. Hart, V. *et al.* The effect of change in body mass index on volumetric measures of mammographic density. *Cancer Epidemiol Biomarkers Prev* **24**, 1724–30 (2015).

59. Rice, M. S. *et al.* Reproductive and lifestyle risk factors and mammographic density in Mexican women. *Ann Epidemiol* **25**, 868–873 (2015).

60. Azam, S., Jacobsen, K. K., Aro, A. R., Lynge, E. & Andersen, Z. J. Hormone replacement therapy and mammographic density: a systematic literature review. *Breast Cancer Res Treat* **182**, 555–579 (2020).

61. McTiernan, A. *et al.* Conjugated Equine Estrogen Influence on Mammographic Density in Postmenopausal Women in a Substudy of the Women's Health Initiative Randomized Trial. *Journal of Clinical Oncology* **27**, 6135 (2009).

62. McTiernan, A. *et al.* Estrogen-plus-progestin use and mammographic density in postmenopausal women: Women's Health Initiative randomized trial. *J Natl Cancer Inst* **97**, 1366–1376 (2005).

63. Stone, J., Warren, R. M. L., Pinney, E., Warwick, J. & Cuzick, J. Determinants of Percentage and Area Measures of Mammographic Density. *Am J Epidemiol* **170**, 1571–1578 (2009).

64. Engmann, N. J. *et al.* Combined Effect of Volumetric Breast Density and Body Mass Index on Breast Cancer Risk. *Breast Cancer Res Treat* **177**, 165 (2019).

65. Boyd, N. F. *et al.* Heritability of mammographic density, a risk factor for breast cancer. *N Engl J Med* **347**, 886–894 (2002).

66. Han, Y., Moore, J. X., Colditz, G. A. & Toriola, A. T. Family History of Breast Cancer and Mammographic Breast Density in Premenopausal Women. *JAMA Netw Open* **5**, e2148983–e2148983 (2022).

67. Lindström, S. *et al.* Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. *Nat Commun* **5**, (2014).

68. Rice, M. S. *et al.* Mammographic density and breast cancer risk: A mediation analysis. *Breast Cancer Research* **18**, 1–13 (2016).

69. Rice, M. S. *et al.* Does mammographic density mediate risk factor associations with breast cancer? An analysis by tumor characteristics. *Breast Cancer Res Treat* **170**, 129–141 (2018).

70. Boyd, N. F., Martin, L. J., Yaffe, M. J. & Minkin, S. Mammographic density and breast cancer risk: current understanding and future prospects. *Breast Cancer Res* **13**, 223 (2011).

71. Altobelli, E. *et al.* Breast Cancer Screening Programmes across the WHO European Region: Differences among Countries Based on National Income Level. *Int J Environ Res Public Health* **14**, (2017).

72. Peintinger, F. National Breast Screening Programs across Europe. *Breast Care (Basel)* **14**, 354–358 (2019).

73. Advisory Committee on Breast Cancer Screening. Screening for breast cancer in England: past and future. *J Med Screen* **13**, 59–61 (2006).

74. Guidance on screening and symptomatic breast imaging, fourth edition. (2019).

75. Faridah, Y. Digital versus screen film mammography: a clinical comparison. *Biomed Imaging Interv J* **4**, (2008).

76. Pisano, E. D. *et al.* Diagnostic Performance of Digital versus Film Mammography for Breast-Cancer Screening. *New England Journal of Medicine* **353**, 1773–1783 (2005).

77. Public Health England. Breast screening care pathway. https://www.gov.uk/government/publications/breast-screening-care-pathway/breast-screening-care-pathway (2021).

78. Public Health England. 71 or over? You can still have breast screening if you choose. https://www.gov.uk/government/publications/breast-screening-for-women-aged-71-or-over/71-or-over-you-can-still-have-breast-screening-if-you-choose (2021).

79. Routes to diagnosis. http://www.ncin.org.uk/publications/routes_to_diagnosis.

80. National Cancer Institute. Prognosis. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/prognosis.

81. Breast Cancer Now. Prognosis. https://breastcancernow.org/information-support/facing-breast-cancer/diagnosed-breast-cancer/prognosis.

82. Senkus, E. *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up † incidence and epidemiology. *ESMO Updated Clinical Practice Guidelines* **26**, v8–v30 (2015).

83. Office for National Statistics. Cancer survival in England. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalinengland/latest (2019).

84. Cserni, G., Chmielik, E., Cserni, B. & Tot, T. The new TNM-based staging of breast cancer. *Virchows Arch* **472**, 697–703 (2018).

85. Cancer Research UK. About breast cancer staging and grades. https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/about-breast-cancer-staging-grades.

86. Orrantia-Borunda, E., Anchondo-Nuñez, P., Acuña-Aguilar, L. E., Gómez-Valles, F. O. & Ramírez-Valdespino, C. A. Subtypes of Breast Cancer. *Breast Cancer* 31–42 (2022) doi:10.36255/EXON-PUBLICATIONS-BREAST-CANCER-SUBTYPES.

87. Dunnwald, L. K., Rossing, M. A. & Li, C. I. Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. *Breast Cancer Res* **9**, (2007).

88. Porter, P. L. *et al.* Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* **91**, 2020–8 (1999).

89. Gilliland, F. D. *et al.* Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* **92**, 743–9 (2000).

90. Burrell, H. C. *et al.* Screening interval breast cancers: mammographic features and prognosis factors. *Radiology* **199**, 811–7 (1996).

91. Shen, Y. *et al.* Role of detection method in predicting breast cancer survival: Analysis of randomized screening trials. *J Natl Cancer Inst* **97**, 1195–1203 (2005).

92. Irvin, V. L. *et al.* Comparison of Mortality Among Participants of Women's Health Initiative Trials With Screening-Detected Breast Cancers vs Interval Breast Cancers. *JAMA Netw Open* **3**, e207227–e207227 (2020).

93. Niraula, S., Biswanger, N., Hu, P. Z., Lambert, P. & Decker, K. Incidence, Characteristics, and Outcomes of Interval Breast Cancers Compared with Screening-Detected Breast Cancers. *JAMA Netw Open* **3**, 3–9 (2020).

94. Allgood, P. C. *et al.* Explaining the difference in prognosis between screen-detected and symptomatic breast cancers. *Br J Cancer* **104**, 1680–1685 (2011).

95. Lehtimäki, T. *et al.* Long-term prognosis of breast cancer detected by mammography screening or other methods. *Breast Cancer Research* **13**, 1–11 (2011).

96. Domingo, L. *et al.* Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers. *Eur J Cancer Prev* **22**, 21–28 (2013).

97. Cabioğlu, N. *et al.* Poor Biological Factors and Prognosis of Interval Breast Cancers: Long-Term Results of Bahçeşehir (Istanbul) Breast Cancer Screening Project in Turkey. *JCO Glob Oncol* 1103–1113 (2020) doi:10.1200/go.20.00145.

98. Bellio, G. *et al.* Interval Breast Cancer Versus Screen-Detected Cancer: Comparison of Clinicopathologic Characteristics in a Single-Center Analysis. *Clin Breast Cancer* **17**, 564–571 (2017).

99. Rayson, D. *et al.* Comparison of clinical-pathologic characteristics and outcomes of true interval and screen-detected invasive breast cancer among participants of a canadian breast screening program: A nested case-control study. *Clin Breast Cancer* **11**, 27–32 (2011).

100. van der Waal, D., Verbeek, A. L. M. & Broeders, M. J. M. Breast density and breast cancer-specific survival by detection mode. *BMC Cancer* **18**, (2018).

101. Eriksson, L. *et al.* Mammographic density and survival in interval breast cancers. *Breast Cancer Res* **15**, (2013).

102. Lee, J. S. *et al.* Five-Year Overall Survival of Interval Breast Cancers is Better than Non- Interval Cancers from Korean Breast Cancer Registry. *Asian Pac J Cancer Prev* **20**, 1717 (2019).

103. Collins, S., Woodman, C. B. J., Threlfall, A. & Prior, P. Survival rates from interval cancer in NHS breast screening programme. *BMJ : British Medical Journal* **316**, 832 (1998).

104. Kalager, M., Tamimi, R. M., Bretthauer, M. & Adami, H. O. Prognosis in women with interval breast cancer: population based observational cohort study. *BMJ* **345**, (2012).

105. Chen, H. L., Zhou, J. Q., Chen, Q. & Deng, Y. C. Comparison of the sensitivity of mammography, ultrasound, magnetic resonance imaging and combinations of these imaging modalities for the detection of small (≤2 cm) breast cancer. *Medicine* **100**, e26531 (2021).

106. Chae, E. Y., Kim, H. H., Cha, J. H., Shin, H. J. & Choi, W. J. Detection and characterization of breast lesions in a selective diagnostic population: diagnostic accuracy study for comparison between one-view digital breast tomosynthesis and two-view full-field digital mammography. *Br J Radiol* **89**, 20150743 (2016).

107. Whelehan, P. *et al.* Digital breast tomosynthesis: sensitivity for cancer in younger symptomatic women. *Br J Radiol* **94**, 20201105 (2021).

108. Public Health England. Breast screening: digital breast tomosynthesis. https://www.gov.uk/government/publications/breast-screening-digital-breast-tomosynthesis/breast-screening-digital-breast-tomosynthesis (2023).

109. Contrast-Enhanced Mammography (CEM) | DenseBreast-info, Inc. https://densebreast-info.org/screening-technologies/contrast-enhanced-mammography/.

110. Watt, G. P. *et al.* Association of breast cancer with quantitative mammographic density measures for women receiving contrast-enhanced mammography. *JNCI Cancer Spectr* **8**, 26 (2024).

111. van Breest Smallenburg, V. *et al.* Lower sensitivity of screening mammography after previous benign breast surgery. *Int J Cancer* **130**, 122–128 (2012).

112. Miglioretti, D. L. *et al.* Effect of breast augmentation on the accuracy of mammography and cancer characteristics. *JAMA* **291**, 442–450 (2004).

113. Isautier, J. M. J. *et al.* Clinical guidelines for the management of mammographic density: a systematic review of breast screening guidelines worldwide. *JNCI Cancer Spectr* **8**, (2024).

114. Food and Drug Administration. FDA Updates Mammography Regulations to Require Reporting of Breast Density Information and Enhance Facility Oversight. https://www.fda.gov/news-

events/press-announcements/fda-updates-mammography-regulations-require-reporting-breast-density-information-and-enhance (2023).

115. NHS Business Services Authority. Hormone Replacement Therapy - England - April 2015 to June 2023. https://www.nhsbsa.nhs.uk/statistical-collections/hormone-replacement-therapy-england/hormone-replacement-therapy-england-april-2015-june-2023 (2023).

116. Zheng, Y. *et al.* Risk prediction models for breast cancer: a systematic review. *BMJ Open* **12**, e055398 (2022).

117. Thompson, P., Stopeck, A., Alberts, D. & Hess, L. Breast Cancer Prevention. in *Fundamentals of Cancer Prevention* 347–385 (2008). doi:10.1007/978-3-540-68986-7_14.

118. Green, J. *et al.* Cohort Profile: the Million Women Study. *Int J Epidemiol* **48**, 28–29e (2019).

119. The Million Women Study. Publications. http://www.millionwomenstudy.org/publications/.

120. Banks, E. *et al.* Effect on attendance at breast cancer screening of adding a self administered questionnaire to the usual invitation to breast screening in southern England. *J Epidemiol Community Health (1978)* **52**, 116 (1998).

121. Beral, V. The Million Women Study: design and characteristics of the study population. *Breast Cancer Research* **1**, 73 (1999).

122. Banks, E. *et al.* Effect on attendance at breast cancer screening of adding a self administered questionnaire to the usual invitation to breast screening in southern England. *J Epidemiol Community Health (1978)* **52**, 116 (1998).

123. Banks, E. *et al.* Comparison of various characteristics of women who do and do not attend for breast cancer screening. *Breast Cancer Research* **4**, 1–6 (2002).

124. Townsend P, Phillimore P & Beattie A. Health and Deprivation: inequality and the North. *Nurs Stand* **2**, 34 (1988).

125. Gubern-Mérida, A. *et al.* Volumetric Breast Density Estimation from Full-Field Digital Mammograms: A Validation Study. *PLoS One* **9**, e85952 (2014).

126. Warren, L. M. *et al.* Deep learning to calculate breast density from processed mammography images. *https://doi.org/10.1117/12.2561278* **11513**, 352–358 (2020).

127. StataCorp. Stata Statistical Software: Release 17. *College Station, TX: StataCorp LLC.* (2021).

128. StataCorp. Stata Statistical Software: Release 18. *College Station, TX: StataCorp LLC.* (2023).

129. Arnold, M. Jasper makes plots. https://github.com/arnhew99/Jasper.

130. Kirkwood, B. & Sterne, J. *Essential Medical Statistics*. (Wiley, 2003).

131. Checka, C. M., Chun, J. E., Schnabel, F. R., Lee, J. & Toth, H. The relationship of mammographic density and age: Implications for breast cancer screening. *American Journal of Roentgenology* **198**, 292–295 (2012).

132. Kellermeyer, L., Harnke, B. & Knight, S. Covidence and Rayyan. *Journal of the Medical Library Association* **106**, (2018).

133. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).

134. Orsini, N., Bellocco, R. & Greenland, S. Generalized least squares for trend estimation of summarized dose-response data. *Stata Journal* **6**, 40–57 (2006).

135. Greenland, S. & Longnecker, M. P. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* **135**, 1301–9 (1992).

136. Barnes. I. Personal communication. (2022).

137. Cox, D. R. & Snell, E. J. Analysis of Binary Data, Second Edition (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). 240 (1989).

138. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, W. V. *Cochrane Handbook for Systematic Reviews of Interventions*.

139. Sterne, J. A. C. *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* **343**, (2011).

140. Joanna Briggs Institute. Critical Appraisal Tools. (2021).

141. Maskarinec, G., Meng, L. & Ursin, G. Ethnic differences in mammographic densities. *Int J Epidemiol* **30**, 959–965 (2001).

142. McDonald, J. A. *et al.* Alcohol intake from early adulthood to midlife and mammographic density. *Cancer Causes and Control* **27**, 493–502 (2016).

143. Brand, J. S. *et al.* Influence of lifestyle factors on mammographic density in postmenopausal women. *PLoS One* **8**, 1–8 (2013).

144. Jacobsen, K. K. *et al.* Alcohol consumption and mammographic density in the Danish Diet, Cancer and Health cohort. *Cancer Causes and Control* **28**, 1429–1439 (2017).

145. Yaghjyan, L. *et al.* Relationship between breast cancer risk factors and mammographic breast density in the Fernald Community Cohort. *Br J Cancer* **106**, 996–1003 (2012).

146. Jacobsen, K. K. emp *et al.* Cigarette smoking and mammographic density in the Danish Diet, Cancer and Health cohort. *Cancer Causes Control* **27**, 271–280 (2016).

147. Singletary, K. W. & Gapstur, S. M. Alcohol and breast cancer: review of epidemiologic and experimental evidence and potential mechanisms. *JAMA* **286**, 2143–51 (2001).

148. Ziembicki, S. *et al.* The association between alcohol consumption and breast density: A systematic review and meta-analysis. *Cancer Epidemiology Biomarkers and Prevention* **26**, 170–178 (2017).

149. Baron, J. A., la Vecchia, C. & Levi, F. The antiestrogenic effect of cigarette smoking in women. *Am J Obstet Gynecol* **162**, 502–14 (1990).

150. Tankó, L. B. & Christiansen, C. An update on the antiestrogenic effect of smoking: a literature review with implications for researchers and practitioners. *Menopause* **11**, 104–109 (2004).

151. Jones, M. E., Schoemaker, M. J., Wright, L. B., Ashworth, A. & Swerdlow, A. J. Smoking and risk of breast cancer in the Generations Study cohort. *Breast Cancer Res* **19**, 118 (2017).

152. Garnett, C., Oldham, M., Shahab, L., Tattan-Birch, H. & Cox, S. Characterising smoking and smoking cessation attempts by risk of alcohol dependence: A representative, cross-sectional study of adults in England between 2014-2021. *The Lancet Regional Health - Europe* **18**, 100418 (2022).

153. Russo, J., Moral, R., Balogh, G. A., Mailo, D. & Russo, I. H. The protective role of pregnancy in breast cancer. *Breast Cancer Res* **7**, 131–42 (2005).

154. Lope, V. *et al.* Obstetric history and mammographic density: A population-based cross-sectional study in spain (DDM-Spain). *Breast Cancer Res Treat* **132**, 1137–1146 (2012).

155. Momenimovahed, Z. & Salehiniya, H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer: Targets and Therapy* **11**, 151–164 (2019).

156. Ward, S. V. *et al.* The association of age at menarche and adult height with mammographic density in the International Consortium of Mammographic Density. *Breast Cancer Research* **24**, 1–16 (2022).

157. McBride, R. B. *et al.* Alcohol and tobacco use in relation to mammographic density in 23,456 women. *Cancer Epidemiology Biomarkers and Prevention* **29**, 1039–1048 (2020).

158. Cabanes, A. *et al.* Alcohol, tobacco, and mammographic density: A population-based study. *Breast Cancer Res Treat* **129**, 135–147 (2011).

159. Lope, V. *et al.* Childhood factors associated with mammographic density in adult women. *Breast Cancer Res Treat* **130**, 965–974 (2011).

160. Simmonds, M. Quantifying the risk of error when interpreting funnel plots. *Syst Rev* **4**, 1–7 (2015).

161. Spak, D. A., Plaxco, J. S., Santiago, L., Dryden, M. J. & Dogan, B. E. BI-RADS® fifth edition: A summary of changes. *Diagn Interv Imaging* **98**, 179–190 (2017).

162. Yaghjyan, L. *et al.* Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. *J Natl Cancer Inst* **103**, 1179–1189 (2011).

163. Beral, V. & Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* **362**, 419–27 (2003).

164. Chen, W. Y., Rosner, B., Hankinson, S. E., Colditz, G. A. & Willett, W. C. Moderate Alcohol Consumption During Adult Life, Drinking Patterns, and Breast Cancer Risk. *JAMA* **306**, 1884–1890 (2011).

165. Ellingjord-Dale, M. *et al.* Parity, hormones and breast cancer subtypes - results from a large nested case-control study in a national screening program. *Breast Cancer Research* **19**, 1–21 (2017).

166. Titus-Ernstoff, L. *et al.* Breast cancer risk factors in relation to breast density (United States). *Cancer Causes and Control* **17**, 1281–1290 (2006).

167. Riza, E. *et al.* Correlates of high-density mammographic parenchymal patterns by menopausal status in a rural population in Northern Greece. *Eur J Cancer* **41**, 590–600 (2005).

168. Sung, H. *et al.* Breast cancer risk factors and mammographic density among high-risk women in urban China. *NPJ Breast Cancer* **4**, (2018).

169. Azam, S. *et al.* Determinants of mammographic density change. *JNCI Cancer Spectr* **3**, 1–12 (2019).

170. Wronka, I. Association between BMI and age at menarche in girls from different socio-economic groups. *Anthropol Anz* **68**, 43–52 (2010).

171. Zhu, J., Kusa, T. O. & Chan, Y. M. Genetics of Pubertal Timing. *Curr Opin Pediatr* **30**, 532 (2018).

172. Dvornyk, V. & Waqar-ul-Haq. Genetics of age at menarche: a systematic review. *Hum Reprod Update* **18**, 198–210 (2012).

173. Barnes, I. *et al.* A comparative analysis of risk factor associations with interval and screen-detected breast cancers: A large UK prospective study. *Int J Cancer* **155**, 979–987 (2024).

174. Crandall, C. J. *et al.* Breast tenderness after initiation of conjugated equine estrogens and mammographic density change. *Breast Cancer Res Treat* **131**, 969–979 (2012).

175. Marchesoni, D. *et al.* Postmenopausal hormone therapy and mammographic breast density. *Maturitas* **53**, 59–64 (2006).

176. Greendale, G. A. *et al.* Postmenopausal hormone therapy and change in mammographic density. *J Natl Cancer Inst* **95**, 30–37 (2003).

177. Olsson, Å., Sartor, H., Borgquist, S., Zackrisson, S. & Manjer, J. Breast density and mode of detection in relation to breast cancer specific survival: a cohort study. *BMC Cancer* **14**, (2014).

178. Yaghjyan, L. *et al.* Relationship between breast cancer risk factors and mammographic breast density in the Fernald Community Cohort. *Br J Cancer* **106**, 996–1003 (2012).

179. Azam, S. *et al.* Hormone replacement therapy, mammographic density, and breast cancer risk: a cohort study. *Cancer Causes and Control* **29**, 495–505 (2018).

180. Chen, F. P., Cheung, Y. C. & Soong, Y. K. Factors that influence changes in mammographic density with postmenopausal hormone therapy. *Taiwan J Obstet Gynecol* **49**, 413–418 (2010).

181. Aiello, E. J., Buist, D. S. M. & White, E. Do breast cancer risk factors modify the association between hormone therapy and mammographic breast density? (United States). *Cancer Causes Control* **17**, 1227–1235 (2006).

182. Van Duijnhoven, F. J. B. *et al.* Postmenopausal hormone therapy and changes in mammographic density. *J Clin Oncol* **25**, 1323–1328 (2007).

183. Couto, E. *et al.* Hormone therapy use and mammographic density in postmenopausal Norwegian women. *Breast Cancer Res Treat* **132**, 297–305 (2012).

184. El-Bastawissi, A. Y., White, E., Mandelson, M. T. & Taplin, S. H. Reproductive and hormonal factors associated with mammographic breast density by age (United States). *Cancer Causes and Control* **11**, 955–963 (2000).

185. Boyd, N. F., Jensen, H. M., Cooke, G., Lee Han, H. & Lockwood, G. A. Mammographic densities and the prevalence and incidence of histological types of benign breast disease. Reference Pathologists of the Canadian National Breast Screening Study. *Eur J Cancer Prev* **9**, 15–24 (2000).

186. Reimers, L. L. *et al.* Benign breast disease and changes in mammographic breast density. *Breast Cancer Research* **23**, 1–9 (2021).

187. Lee, J. S. & Oh, M. Breast density of mammography is correlated with reproductive risk factors regardless of menopausal status: A cross-sectional study of the Korean national screening program. *Asian Pacific Journal of Cancer Prevention* **21**, 1011–1018 (2020).

188. Masala, G. *et al.* Can dietary and physical activity modifications reduce breast density in postmenopausal women? The DAMA study, a randomized intervention trial in Italy. *Cancer Epidemiology Biomarkers and Prevention* **28**, 41–50 (2019).

189. Woolcott, C. G. *et al.* Mammographic density change with 1 Year of aerobic exercise among postmenopausal women: A randomized controlled trial. *Cancer Epidemiology Biomarkers and Prevention* **19**, 1112–1121 (2010).

190. Yaghjyan, L., Colditz, G. A. & Wolin, K. Physical activity and mammographic breast density: a systematic review. *Breast Cancer Res Treat* **135**, 367–80 (2012).

191. Azam, S. *et al.* Regular physical activity and mammographic density: a cohort study. *Cancer Causes and Control* **29**, 1015–1025 (2018).

192. Hanna, M., Dumas, I., Jacob, S., Têtu, B. & Diorio, C. Physical activity, mammographic density, and age-related lobular involution among premenopausal and postmenopausal women. *Menopause* **22**, 964–975 (2015).

193. Mukhopadhaya, N. & Manyonda, I. T. The hysterectomy story in the United Kingdom. *J Midlife Health* **4**, 40 (2013).

194. Ruiz de Azua Unzurrunzaga, G., Brewster, D. H., Wild, S. H. & Sivalingam, V. N. Declining hysterectomy prevalence and the estimated impact on uterine cancer incidence in Scotland. *Cancer Epidemiol* **59**, 227–231 (2019).

195. Sievert, L. L., Murphy, L., Morrison, L. A., Reza, A. M. & Brown, D. E. Age at menopause and determinants of hysterectomy and menopause in a multi-ethnic community: The Hilo Women's Health Study. *Maturitas* **76**, 10.1016/j.maturitas.2013.08.007 (2013).

196. NHS England. New breast screening figures prompt fresh uptake appeal. https://www.england.nhs.uk/2024/01/new-breast-screening-figures-prompt-fresh-uptake-appeal/ (2024).

197. Gøtzsche, P. C. & Jørgensen, K. J. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews* **2013**, (2013).

198. Marmot, M. *et al.* The benefits and harms of breast cancer screening: An independent review. *The Lancet* vol. 380 1778–1786 Preprint at https://doi.org/10.1016/S0140-6736(12)61611-0 (2012).

199. Centres for Disease Control and Prevention. Screening for Breast Cancer. https://www.cdc.gov/breast-cancer/screening/index.html.

200. Altobelli, E. & Lattanzi, A. Breast cancer in European Union: An update of screening programmes as of March 2014 (Review). *Int J Oncol* **45**, 1785 (2014).

201. Autier, P., Koechlin, A., Smans, M., Vatten, L. & Boniol, M. Mammography screening and breast cancer mortality in Sweden. *J Natl Cancer Inst* **104**, 1080–1093 (2012).

202. Nickson, C., Velentzis, L. S., Brennan, P., Bruce Mann, G. & Houssami, N. Improving breast cancer screening in Australia: a public health perspective | PHRP. *Public Health Res Pract* **29**, (2019).

203. Li, M., Warner-Smith, M., McGill, S., Roder, D. & Currow, D. History of screening by BreastScreen New South Wales of women with invasive breast cancer. *Cancer Epidemiol* **64**, (2020).

204. Federal Ministry of Social Affairs, H. Screening programme: Early detection of breast cancer. https://www.oesterreich.gv.at/en/themen/gesundheit_und_notfaelle/medizinische_versorgung/screening-programm_brustkrebs-frueherkennung.html (2022).

205. De Gauquier, K., Remacle, A., Fabri, V. & Mertens, R. Evaluation of the first round of the national breast cancer screening programme in Flanders, Belgium. *Arch Public Health* **64**, 71–80 (2006).

206. Canadian Partnership Against Cancer. Breast cancer screening in Canada: Monitoring and evaluation of quality indicators. https://www.partnershipagainstcancer.ca/topics/breast-cancer-screening-quality-indicators-2017/ (2017).

207. Canadian Partnership Against Cancer. *Breast Cancer Screening in Canada: Environmental Scan*. https://www.partnershipagainstcancer.ca/topics/breast-cancer-screening-scan-2019-2020/ (2019).

208. Canadian Cancer Society. Screening for breast cancer. https://cancer.ca/en/cancer-information/cancer-types/breast/screening.

209. Danish Cancer Society. Screening for breast cancer (mammography) . https://www.cancer.dk/international/english/screening-breast-cancer-english/.

210. Lynge, E. *et al.* Outcome of breast cancer screening in Denmark. *BMC Cancer* **17**, 1–9 (2017).

211. Cancer Society of Finland. Breast cancer screening . https://cancerregistry.fi/screening/breast-cancer-screening/.

212. Cancer Society of Finland. Cancer screenings . https://www.cancersociety.fi/publications/reports/cancer-in-finland-2016/cancer-screenings/.

213. Regnault, M. Breast cancer: in France, a screening program for women between 50 and 74 years of age | Institut Curie. https://institut-curie.org/dossier-pedagogique/breast-cancer-france-screening-program-women-between-50-and-74-years-age (2017).

214. Informed Health. The breast cancer screening program in Germany. https://www.ncbi.nlm.nih.gov/books/NBK361021/ (2016).

215. Biesheuvel, C., Weige, S. & Heindel, W. Mammography Screening: Evidence, History and Current Practice in Germany and Other European Countries. *Breast Care* **6**, 104 (2011).

216. Gabe, R. *et al.* A case-control study to estimate the impact of the Icelandic population-based mammography screening program on breast cancer death. *Acta Radiol* **48**, 948–955 (2007).

217. Sigurdsson, K. & Ólafsdóttir, E. J. Population-based service mammography screening: the Icelandic experience. *Breast Cancer : Targets and Therapy* **5**, 17 (2013).

218. Health Service Executive. Breast screening information . https://www2.hse.ie/services/breast-screening/breast-screening-overview/breast-screening-information.html.

219. Giordano, L. *et al.* Breast cancer screening in Italy: evaluating key performance indicators for time trends and activity volumes. *Epidemiol Prev* **39**, 30–9 (2015).

220. Foca, F. *et al.* Decreasing incidence of late-stage breast cancer after the introduction of organized mammography screening in Italy. *Cancer* **119**, 2022–2028 (2013).

221. National Cancer Plan – Luxembourg. Current breast cancer screening. https://plancancer.lu/about/depistage/cancer-du-sein/depistage-actuel-cancer-du-sein/.

222. Cancer Registry of Norway. Breast Screen Norway. https://www.kreftregisteret.no/en/screening/BreastScreen_Norway/breastscreen-norway/.

223. Weedon-Fekjræ, H., Romundstad, P. R. & Vatten, L. J. Modern mammography screening and breast cancer mortality: population study. *The BMJ* **348**, (2014).

224. Dourado, F., Carreira, H. & Lunet, N. Mammography use for breast cancer screening in Portugal: results from the 2005/2006 National Health Survey. *The European Journal of Public Health* **23**, 386 (2013).

225. Jarm, K. *et al.* Quality assured implementation of the Slovenian breast cancer screening programme. *PLoS One* **16**, e0258343 (2021).

226. Ascunce, N. *et al.* Cancer screening in Spain. *Ann Oncol* **21 Suppl 3**, (2010).

227. Bosch, G. *et al.* Impact of the COVID-19 pandemic on breast cancer screening indicators in a Spanish population-based program: a cohort study. *Elife* **11**, (2022).

228. Bähler, C. *et al.* Temporal trends and regional disparities in cancer screening utilization: an observational Swiss claims-based study. *BMC Public Health* **21**, 1–12 (2021).

229. Swiss Cancer Screening. *Early Detection through Mammography Early Breast Cancer Detection-Improved Chances of Recovery*.

230. de Glas, N. A. *et al.* Effect of implementation of the mass breast cancer screening programme in older women in the Netherlands: population based study. *BMJ* **349**, g5410 (2014).

231. Rainey, L., Van Der Waal, D. & Broeders, M. J. M. Dutch women's intended participation in a risk-based breast cancer screening and prevention programme: A survey study identifying preferences, facilitators and barriers. *BMC Cancer* **20**, 1–14 (2020).

232. National Evaluation Team for Breast cancer screening. *National Evaluation of Breast Cancer Screening in the Netherlands*. (2014).

233. Richards, M. *Report of THE INDEPENDENT REVIEW OF ADULT SCREENING PROGRAMMES in England*. (2019).

234. Chiarelli, A. M. *et al.* Influence of patterns of hormone replacement therapy use and mammographic density on breast cancer detection. *Cancer Epidemiology Biomarkers and Prevention* **15**, 1856–1862 (2006).

235. Porter, P. L. *et al.* Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* **91**, 2020–2028 (1999).

236. Cornford, E. & Sharma, N. Interval Cancers and Duty of Candour, a UK Perspective. *Curr Breast Cancer Rep* **11**, 89–93 (2019).

237. Fitzpatrick, P. *et al.* Interval cancer audit and disclosure in breast screening programmes: An international survey. *J Med Screen* **30**, 36 (2022).

238. Harper-Harrison G & Shanahan MM. Hormone Replacement Therapy. *StatPearls* (2022).

239. NICE. Menopause: diagnosis and management | Guidance. *NICE guidelines* https://www.nice.org.uk/guidance/ng23/chapter/context (2015).

240. Department of Health and Social Care. Hundreds of thousands of women experiencing menopause symptoms to get cheaper HRT. https://www.gov.uk/government/news/hundreds-of-thousands-of-women-experiencing-menopause-symptoms-to-get-cheaper-hormone-replacement-therapy (2023).

241. NHS. Hormone replacement therapy (HRT) - Types . https://www.nhs.uk/conditions/hormone-replacement-therapy-hrt/types/ (2019).

242. Endometrial cancer and hormone-replacement therapy in the Million Women Study. *Lancet* **365**, 1543–1551 (2005).

243. Woodruff, J. D. & Pickar, J. H. Incidence of endometrial hyperplasia in postmenopausal women taking conjugated estrogens (Premarin) with medroxyprogesterone acetate or conjugated estrogens alone. The Menopause Study Group. *Am J Obstet Gynecol* **170**, 1213–1223 (1994).

244. Smith, D. C., Prentice, R., Thompson, D. J. & Herrmann, W. L. Association of exogenous estrogen and endometrial carcinoma. *N Engl J Med* **293**, 1164–1167 (1975).

245. Ziel, H. K. & Finkle, W. D. Increased risk of endometrial carcinoma among users of conjugated estrogens. *N Engl J Med* **293**, 1167–1170 (1975).

246. NHS. About tibolone. https://www.nhs.uk/medicines/hormone-replacement-therapy-hrt/tibolone/about-tibolone/ (2023).

247. Rymer, J. Why tibolone is different? *Reviews in Gynaecological Practice* **2**, 35–39 (2002).

248. 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* **288**, 321–333 (2002).

249. Chlebowski, R. T. *et al.* Estrogen plus progestin and breast cancer incidence and mortality in postmenopausal women. *JAMA* **304**, 1684–1692 (2010).

250. Anderson, G. L. & Limacher, M. Effects of Conjugated Equine Estrogen in Postmenopausal Women With Hysterectomy: The Women's Health Initiative Randomized Controlled Trial. *JAMA* **291**, 1701–1712 (2004).

251. Chen, W. Y. *et al.* Association of hormone replacement therapy to estrogen and progesterone receptor status in invasive breast carcinoma. *Cancer* **101**, 1490–1500 (2004).

252. Saxena, T. *et al.* Menopausal hormone therapy and subsequent risk of specific invasive breast cancer subtypes in the California Teachers Study. *Cancer Epidemiol Biomarkers Prev* **19**, 2366 (2010).

253. Salagame, U., Banks, E., O'Connell, D. L., Egger, S. & Canfell, K. Menopausal Hormone Therapy use and breast cancer risk by receptor subtypes: Results from the New South Wales Cancer Lifestyle and EvaluAtion of Risk (CLEAR) study. *PLoS One* **13**, (2018).

254. Banks, E. Hormone replacement therapy and the sensitivity and specificity of breast cancer screening: a review. *J Med Screen* **8**, 29–35 (2001).

255. Kavanagh, A. M. *et al. Hormone Replacement Therapy, Percent Mammographic Density, and Sensitivity of Mammography*. *Cancer Epidemiol Biomarkers Prev* http://aacrjournals.org/cebp/article-pdf/14/5/1060/2263369/1060-1064.pdf (2005).

256. Banks, E. *et al.* Influence of personal characteristics of individual women on sensitivity and specificity of mammography in the Million Women Study: Cohort study. *Br Med J* **329**, 477–479 (2004).

257. Berrington, A. & Cox, D. R. *Generalized Least Squares for the Synthesis of Correlated Information*. *Biostatistics* vol. 4 https://academic.oup.com/biostatistics/article/4/3/423/226665 (2003).

258. Kavanagh, A. M. *et al.* Using mammographic density to improve breast cancer screening outcomes. *Cancer Epidemiology Biomarkers and Prevention* **17**, 2818–2824 (2008).

259. Crane, C. E. B., Luke, C. G., Rogers, J. M., Playford, P. E. & Roder, D. M. An analysis of factors associated with interval as opposed to screen-detected breast cancers, including hormone therapy and mammographic density. *Breast* **11**, 131–136 (2002).

260. Domingo, L. *et al.* Phenotypic characterization and risk factors for interval breast cancers in a population-based breast cancer screening program in Barcelona, Spain. *Cancer Causes and Control* **21**, 1155–1164 (2010).

261. Holm, J. *et al.* Risk factors and tumor characteristics of interval cancers by mammographic density. *Journal of Clinical Oncology* **33**, 1030–1037 (2015).

262. Chiarelli, A. M. *et al.* Digital versus screen-film mammography: impact of mammographic density and hormone therapy on breast cancer detection. *Breast Cancer Res Treat* **154**, 377–387 (2015).

263. Njor, S. H., Pedersen, A. T., Schwartz, W., Hallas, J. & Lynge, E. Minimizing misclassification of hormone users at mammography screening. *Int J Cancer* **124**, 2159–2165 (2009).

264. Njor, S. H., Hallas, J., Schwartz, W., Lynge, E. & Pedersen, A. T. Type of hormone therapy and risk of misclassification at mammography screening. *Menopause* **18**, 171–177 (2011).

265. Strand, F. *et al.* Novel mammographic image features differentiate between interval and screendetected breast cancer: A case-case study. *Breast Cancer Research* **18**, (2016).

266. Gaudet, M. M. *et al.* Breast cancer risk factors by mode of detection among screened women in the Cancer Prevention Study-II. *Breast Cancer Res Treat* **186**, 791–805 (2021).

267. Sprague, B. L. *et al.* Variation in Breast Cancer-Risk Factor Associations by Method of Detection: Results from a Series of Case-Control Studies. *Am J Epidemiol* **181**, 956–969 (2015).

268. Blanch, J. *et al.* Impact of risk factors on different interval cancer subtypes in a population-based breast cancer screening programme. *PLoS One* **9**, (2014).

269. Hofvind, S., Møller, B., Thoresen, S. & Ursin, G. Use of hormone therapy and risk of breast cancer detected at screening and between mammographic screens. *Int J Cancer* **118**, 3112–3117 (2006).

270. Séradour, B., Estève, J., Heid, P. & Jacquemier, J. Hormone replacement therapy and screening mammography: Analysis of the results in the Bouches du Rhone programme. *J Med Screen* **6**, 99–102 (1999).

271. National Cancer Institute. Hormone Therapy for Breast Cancer Fact Sheet. https://www.cancer.gov/types/breast/breast-hormone-therapy-fact-sheet (2022).

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

273. Hofseth, L. J. *et al.* Hormone Replacement Therapy with Estrogen or Estrogen plus Medroxyprogesterone Acetate Is Associated with Increased Epithelial Proliferation in the Normal Postmenopausal Breast. *J Clin Endocrinol Metab* **84**, 4559–4565 (1999).

274. Sendag, F. *et al.* Mammographic density changes during different postmenopausal hormone replacement therapies. *Fertil Steril* **76**, 445–450 (2001).

275. Li, C. I. *et al.* Use of menopausal hormone therapy and risk of ductal and lobular breast cancer among women 55–74 years of age. *Breast Cancer Res Treat* **145**, 481 (2014).

276. Anderson, K. N., Schwab, R. B. & Martinez, M. E. Reproductive risk factors and breast cancer subtypes: a review of the literature. *Breast Cancer Res Treat* **144**, 1–10 (2014).

277. Hussain, S. *et al.* Breast cancer risk prediction using machine learning: a systematic review. *Front Oncol* **14**, 1343627 (2024).

278. Evans, D. G. R. & Howell, A. Breast cancer risk-assessment models. *Breast Cancer Res* **9**, 213 (2007).

279. Guan, Z. *et al.* Combining Breast Cancer Risk Prediction Models. *Cancers (Basel)* **15**, (2023).

280. McCarthy, A. M. *et al.* Performance of Breast Cancer Risk-Assessment Models in a Large Mammography Cohort. *J Natl Cancer Inst* **112**, 489–497 (2020).

281. Lee, A. *et al.* BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med* **21**, 1708–1718 (2019).

282. Lee, A. *et al.* BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genetics in Medicine* **21**, 1708 (2019).

283. Carver, T. *et al.* Canrisk tool—A web interface for the prediction of breast and ovarian cancer risk and the likelihood of carrying genetic pathogenic variants. *Cancer Epidemiology Biomarkers and Prevention* **30**, 469–473 (2021).

284. McVeigh, T. CanRisk — Knowledge Hub. https://www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/canrisk/?utm_source=chatgpt.com (2024).

285. GOV.UK. Protocols for the surveillance of women at higher risk of developing breast cancer. https://www.gov.uk/government/publications/breast-screening-higher-risk-women-surveillance-protocols/protocols-for-surveillance-of-women-at-higher-risk-of-developing-breast-cancer (2025).

286. Mbuya-Bienge, C. *et al.* A Systematic Review and Critical Assessment of Breast Cancer Risk Prediction Tools Incorporating a Polygenic Risk Score for the General Population. *Cancers (Basel)* **15**, 5380 (2023).

287. Dite, G. S. *et al.* Breast cancer risk prediction using clinical models and 77 independent risk-associated SNPs for women aged under 50 years: Australian Breast Cancer Family Registry. *Cancer Epidemiol Biomarkers Prev* **25**, 359 (2015).

288. Pal Choudhury, P. *et al.* Comparative validation of the BOADICEA and Tyrer-Cuzick breast cancer risk models incorporating classical risk factors and polygenic risk in a population-based prospective cohort of women of European ancestry. *Breast Cancer Research* **23**, 1–5 (2021).

289. Colditz, G. A., Rosner, B. A., Chen, W. Y., Holmes, M. D. & Hankinson, S. E. Risk factors for breast cancer according to estrogen and progesterone receptor status. *J Natl Cancer Inst* **96**, 218–28 (2004).

290. Çelik, L., Güner, D. C., Özçağlayan, Ö., Çubuk, R. & Arıbal, M. E. Diagnostic performance of two versions of an artificial intelligence system in interval breast cancer detection. *Acta radiol* (2023) doi:10.1177/02841851231200785.

291. Vachon, C. M. *et al. Impact of Artificial Intelligence System and Volumetric Density on Risk Prediction of Interval, Screen-Detected, and Advanced Breast Cancer*. *J Clin Oncol* vol. 41 https://doi.org/10. (2023).

292. Haček, R. T., Antoljak, N. & Erceg, M. Development of a model for predicting risk from breast interval cancer in the female population in the Republic of Croatia. *Neoplasma* **69**, 708–722 (2022).

293. Wanders, A. J. T. *et al.* Interval Cancer Detection Using a Neural Network and Breast Density in Women with Negative Screening Mammograms. *Radiology* **303**, 269–275 (2022).

294. Burnside, E. S. *et al.* Quantitative breast density analysis to predict interval and node-positive cancers in pursuit of improved screening protocols: a case–control study. *Br J Cancer* **125**, 884–892 (2021).

295. Zhu, X. *et al.* Deep learning predicts interval and screening-detected cancer from screening mammograms: A case-case-control study in 6369 women. *Radiology* **301**, 550–558 (2021).

296. Nguyen, T. L. *et al.* Interval breast cancer risk associations with breast density, family history and breast tissue aging. *Int J Cancer* **147**, 375–382 (2020).

297. Hinton, B. *et al.* Deep learning networks find unique mammographic differences in previous negative mammograms between interval and screen-detected cancers: A case-case study. *Cancer Imaging* **19**, (2019).

298. Kerlikowske, K. *et al.* Automated and clinical breast imaging reporting and data system density measures predict risk for screen-detected and interval cancers: A case-control study. *Ann Intern Med* **168**, 757–765 (2018).

299. Nguyen, T. L. *et al.* Predicting interval and screen-detected breast cancers from mammographic density defined by different brightness thresholds. *Breast Cancer Research* **20**, (2018).

300. Collins, G. S. *et al.* Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ* **384**, (2024).

301. Xu, H. *et al.* Estimating the receiver operating characteristic curve in matched case control studies. *Stat Med* **38**, 437 (2018).

302. Crispo, A. *et al.* Gail model risk factors: impact of adding an extended family history for breast cancer. *Breast J* **14**, 221–7 (2008).

303. Tice, J. A. *et al.* Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* **148**, 337–47 (2008).

304. Anothaisintawee, T. *et al.* Development and validation of a breast cancer risk prediction model for Thai women: a cross-sectional study. *Asian Pac J Cancer Prev* **15**, 6811–7 (2014).

305. Maas, P. *et al.* Breast Cancer Risk From Modifiable and Nonmodifiable Risk Factors Among White Women in the United States. *JAMA Oncol* **2**, 1295–1302 (2016).

306. Han, Y. *et al.* Development and external validation of a breast cancer absolute risk prediction model in Chinese population. *Breast Cancer Research* **23**, 1–13 (2021).

307. Yang, X. *et al.* Prospective validation of the BOADICEA multifactorial breast cancer risk prediction model in a large prospective cohort study. *J Med Genet* **59**, 1196–1205 (2022).

308. McCarthy, A. M. *et al.* Performance of Breast Cancer Risk-Assessment Models in a Large Mammography Cohort. *JNCI Journal of the National Cancer Institute* **112**, 489 (2019).

309. Hernández-García, M. *et al.* Differences in breast cancer-risk factors between screen-detected and non-screen-detected cases (MCC-Spain study). *Cancer Causes & Control* **33**, 125 (2022).

310. Group on Hormonal Factors in Breast Cancer, C. Type and timing of menopausal hormone therapy and breast cancer risk: individual participant meta-analysis of the worldwide epidemiological evidence. *The Lancet* **394**, 1159–1168 (2019).

311. Chlebowski, R. T. *et al.* Predicting risk of breast cancer in postmenopausal women by hormone receptor status. *J Natl Cancer Inst* **99**, 1695–1705 (2007).

312. Kerlikowske, K., Gard, C. C., Sprague, B. L., Tice, J. A. & Miglioretti, D. L. One versus two breast density measures to predict 5- and 10-year breast cancer risk. *Cancer Epidemiology Biomarkers and Prevention* **24**, 889–897 (2015).

313. Yamada, M. *et al.* Breast cancer risk assessment based on a predictive model: evaluation of risk factors among Japanese women. *BMC Cancer* **25**, 1–13 (2025).

314. Patino, C. M. & Ferreira, J. C. Internal and external validity: can you apply research study results to your patients? *J Bras Pneumol* **44**, 183 (2018).

315. Cancer Epidemiology Unit. AgeX Trial. https://www.ceu.ox.ac.uk/research/agex-trial.

316. GOV.UK. Population of England and Wales - GOV.UK Ethnicity facts and figures. https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest/ (2022).

317. Wise, J. Why are there shortages of HRT and other drugs in the UK? *BMJ* **377**, (2022).

318. Food and Drug Administration. Important Information: Final Rule to Amend the Mammography Quality Standards Act (MQSA). https://www.fda.gov/radiation-emitting-products/mammography-quality-standards-act-mqsa-and-mqsa-program/important-information-final-rule-amend-mammography-quality-standards-act-mqsa (2024).

319. Rowe, M. An Introduction to Machine Learning for Clinicians. *Acad Med* **94**, 1433–1436 (2019).

320. Ng, A. Y. *et al.* Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. *Nature Medicine 2023 29:12* **29**, 3044–3049 (2023).

321. Lång, K. *et al.* Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety

analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *Lancet Oncol* **24**, 936–944 (2023).

322. Dembrower, K., Crippa, A., Colón, E., Eklund, M. & Strand, F. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *Lancet Digit Health* **5**, e703–e711 (2023).

323. Xu, C., Coen-Pirani, P. & Jiang, X. Empirical Study of Overfitting in Deep Learning for Predicting Breast Cancer Metastasis. *Cancers (Basel)* **15**, 1969 (2023).

324. Ahern, T. P. *et al.* Lifestyle, behavioral, and dietary risk factors in relation to mammographic breast density in women at high risk for breast cancer. *Cancer Epidemiology Biomarkers and Prevention* **30**, 936–944 (2021).

325. Pepłońska, B. & Kałużny, P. CIGARETTE SMOKING and MAMMOGRAPHIC BREAST DENSITY among POLISH WOMEN. *Int J Occup Med Environ Health* **34**, 805–815 (2021).

326. Lee, E. *et al.* Immigration history, lifestyle characteristics, and breast density in the Vietnamese American Women's Health Study: a cross-sectional analysis. *Cancer Causes and Control* **31**, 127–138 (2020).

327. Moore, J. X., Han, Y., Appleton, C., Colditz, G. & Toriola, A. T. Determinants of Mammographic Breast Density by Race Among a Large Screening Population. *JNCI Cancer Spectr* **0**, 1–10 (2020).

328. Albeshan, S. M. *et al.* Preliminary investigation of mammographic density among women in Riyadh: association with breast cancer risk factors and implications for screening practices. *Clin Imaging* **54**, 138–147 (2019).

329. Kaya, C. *et al.* The relation between metabolic syndrome and its components with breast density in postmenopausal women. *Metab Syndr Relat Disord* **17**, 341–345 (2019).

330. Moran, O. *et al.* Predictors of mammographic density among women with a strong family history of breast cancer. *BMC Cancer* **19**, 1–12 (2019).

331. Okamoto, T. & Ito, A. Association between alcohol consumption and mammographic density: a hospital-based cross-sectional study. *Breast Cancer* **26**, 478–484 (2019).

332. Liu, Y., Tamimi, R. M., Colditz, G. A. & Bertrand, K. A. Alcohol consumption across the life course and mammographic density in premenopausal women. *Breast Cancer Res Treat* **167**, 529–535 (2018).

333. Dung Trieu, P., Mello-Thoms, C., Peat, J. K., Doan Do, T. & Brennan, P. C. Associations of Breast Density with Demographic, Reproductive, and Lifestyle Factors in a Developing Southeast Asian Population. *Asia Pac J Public Health* **29**, 377–387 (2017).

334. Pereira, A. *et al.* Determinants of volumetric breast density in Chilean premenopausal women. *Breast Cancer Res Treat* **162**, 343–352 (2017).

335. Yang, Y. *et al.* Influence of factors on mammographic density in premenopausal Chinese women. *European Journal of Cancer Prevention* **25**, 306–311 (2016).

336. Frydenberg, H. *et al.* Alcohol consumption, endogenous estrogen and mammographic density among premenopausal women. *Breast Cancer Research* **17**, 1–12 (2015).

337. Quandt, Z. *et al.* The association of alcohol consumption with mammographic density in a multiethnic urban population. *BMC Cancer* **15**, 1–10 (2015).

338. Trinh, T. *et al.* Background risk of breast cancer influences the association between alcohol consumption and mammographic density. *Br J Cancer* **113**, 159–165 (2015).

339. Borugian, M. J. *et al.* Fasting insulin and endogenous hormones in relation to premenopausal breast density (Canada). *Cancer Causes and Control* **25**, 385–394 (2014).

340. Dai, H. *et al.* Distribution of mammographic density and its influential factors among Chinese women. *Int J Epidemiol* **43**, 1240–1251 (2014).

341. Ahmadinejad, N., Movahedinia, S., Movahedinia, S., Naieni, K. H. & Nedjat, S. Distribution of breast density in Iranian women and its association with breast cancer risk factors. *Iran Red Crescent Med J* **15**, (2013).

342. Obajimi, M. O. *et al.* Mammographic breast pattern in Nigerian women in Ibadan, Nigeria. *Breast Dis* **33**, 9–15 (2011).

343. Qureshi, S. A., Couto, E., Hofvind, S., Wu, A. H. & Ursin, G. Alcohol intake and mammographic density in postmenopausal Norwegian women. *Breast Cancer Res Treat* **131**, 993–1002 (2012).

344. Jeon, J. H. *et al.* Reproductive and hormonal factors associated with fatty or dense breast patterns among Korean women. *Cancer Res Treat* **43**, 42–48 (2011).

345. Sung, J., Song, Y. M., Stone, J., Lee, K. & Lee, D. Reproductive factors associated with mammographic density: A Korean co-twin control study. *Breast Cancer Res Treat* **128**, 567–572 (2011).

346. Tehranifar, P. *et al.* Reproductive and menstrual factors and mammographic density in African American, Caribbean, and white women. *Cancer Causes and Control* **22**, 599–610 (2011).

347. Terry, M. B. *et al.* Prenatal smoke exposure and mammographic density in mid-life. *J Dev Orig Health Dis* **2**, 340–352 (2011).

348. Tseng, M., Olufade, T. O., Evers, K. A. & Byrne, C. Adolescent lifestyle factors and adult breast density in U.S. Chinese immigrant women. *Nutr Cancer* **63**, 342–349 (2011).

349. Butler, L. M. *et al.* Active, but not passive cigarette smoking was inversely associated with mammographic density. *Cancer Causes and Control* **21**, 301–311 (2010).

350. McCormack, V. A., Perry, N. M., Vinnicombe, S. J. & dos Santos Silva, I. Changes and tracking of mammographic density in relation to Pike's model of breast tissue aging: A UK longitudinal study. *Int J Cancer* **127**, 452–461 (2010).

351. Flom, J. D., Ferris, J. S., Tehranifar, P. & Terry, M. B. Alcohol intake over the life course and mammographic density. *Breast Cancer Res Treat* **117**, 643–651 (2009).

352. Stone, J., Warren, R. M. L., Pinney, E., Warwick, J. & Cuzick, J. Determinants of percentage and area measures of mammographic density. *Am J Epidemiol* **170**, 1571–1578 (2009).

353. Butler, L. M. *et al.* Menstrual and reproductive factors in relation to mammographic density: The Study of Women's Health Across the Nation (SWAN). *Breast Cancer Res Treat* **112**, 165–174 (2008).

354. Jeffreys, M., Warren, R., Highnam, R. & Davey Smith, G. Breast cancer risk factors and a novel measure of volumetric breast density: Cross-sectional study. *Br J Cancer* **98**, 210–216 (2008).

355. McCormack, V. A., Perry, N., Vinnicombe, S. J. & Silva, I. D. S. Ethnic variations in mammographic density: A British multiethnic longitudinal study. *Am J Epidemiol* **168**, 412–421 (2008).

356. Tseng, M., Sellers, T. A., Vierkant, R. A., Kushi, L. H. & Vachon, C. M. Mediterranean diet and breast density in the Minnesota breast cancer family study. *Nutr Cancer* **60**, 703–709 (2008).

357. Bremnes, Y., Ursin, G., Bjurstam, N. & Gram, I. T. Different measures of smoking exposure and mammographic density in postmenopausal Norwegian women: a cross-sectional study. *Breast Cancer Res* **9**, 6–9 (2007).

358. Masala, G. *et al.* Dietary and lifestyle determinants of mammographic breast density. A longitudinal study in a Mediterranean population. *Int J Cancer* **118**, 1782–1789 (2006).

359. Modugno, F. *et al.* Breast cancer risk factors and mammographic breast density in women over age 70. *Breast Cancer Res Treat* **97**, 157–166 (2006).

360. Tseng, M., Byrne, C., Evers, K. A., London, W. T. & Daly, M. B. Acculturation and breast density in foreign-born, U.S. Chinese women. *Cancer Epidemiology Biomarkers and Prevention* **15**, 1301–1305 (2006).

361. Haars, G., van Noord, P. A. H., van Gils, C. H., Grobbee, D. E. & Peeters, P. H. M. Measurements of breast density: No ratio for a ratio. *Cancer Epidemiology Biomarkers and Prevention* **14**, 2634–2640 (2005).

362. Heng, D. *et al.* Risk factors for breast cancer associated with mammographic features in Singaporean Chinese women. *Cancer Epidemiology Biomarkers and Prevention* **13**, 1751–1758 (2004).

363. Jeffreys, M., Warren, R., Gunnell, D., McCarron, P. & Smith, G. D. Life course breast cancer risk factors and adult breast density (United Kingdom). *Cancer Causes and Control* **15**, 947–955 (2004).

364. Gapstur, S. M. *et al.* Associations of Breast Cancer Risk Factors with Breast Density in Hispanic Women. *Cancer Epidemiology Biomarkers and Prevention* **12**, 1074–1080 (2003).

365. Roubidoux, M. A. *et al.* Relationship of Mammographic Parenchymal Patterns to Breast Cancer Risk Factors and Smoking in Alaska Native Women. *Cancer Epidemiology Biomarkers and Prevention* **12**, 1081–1086 (2003).

366. Warwick, J. *et al.* Breast density and breast cancer risk factors in a high-risk population. *Breast* **12**, 10–16 (2003).

367. Maskarinec, G., Nagata, C., Shimizu, H. & Kashiki, Y. Comparison of mammographic densities and their determinants in women from Japan and Hawaii. *Int J Cancer* **102**, 29–33 (2002).

368. Soares, D., Reid, M. & James, M. Age as a predictive factor of mammographic breast density in Jamaican women. *Clin Radiol* **57**, 472–476 (2002).

369. Vachon, C. M., Kuni, C. C., Anderson, K., Anderson, V. E. & Sellers, T. A. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). *Cancer Causes and Control* **11**, 653–662 (2000).

370. Jakes, R. W., Duffy, S. W., Ng, F. C., Gao, F. & Ng, E. H. Mammographic parenchymal patterns and risk of breast cancer at and after a prevalence screen in Singaporean women. *Int J Epidemiol* **29**, 11–19 (2000).

Appendices

Chapter 3 appendix

Appendix 1. Searches for systematic review

Database	Risk factor	
PubMed	Alcohol	(breast density[Title/Abstract] OR mammographic density[Title/Abstract]) AND ((alcohol[Title/Abstract]) OR (lifestyle[Title/Abstract]))
	Smoking	(breast density[Title/Abstract] OR mammographic density[Title/Abstract]) AND ((smoking[Title/Abstract]) OR (tobacco[Title/Abstract]) OR (lifestyle[Title/Abstract]))
	All reproductive	(breast density[Title/Abstract] OR mammographic density[Title/Abstract]) AND ((parity [Title/Abstract]) OR (number of births [Title/Abstract]) OR (number of children [Title/Abstract]) OR (menstrual [Title/Abstract]) OR (menarche [Title/Abstract]) OR (reproductive [Title/Abstract]) OR (age at first birth [Title/Abstract]) OR (at first pregnancy [Title/Abstract]))
Web of Science	Alcohol	(AB=((breast density OR mammographic density) AND ((alcohol) OR (lifestyle)))) OR (TI=((breast density OR mammographic density) AND ((alcohol) OR (lifestyle))))
	Smoking	(AB=((breast density OR mammographic density) AND ((smoking) OR (tobacco) OR (lifestyle)))) OR (TI=((breast density OR mammographic density) AND ((smoking) OR (tobacco) OR (lifestyle))))
	All reproductive	(AB=((breast density OR mammographic density) AND ((parity) OR (number of births) OR (number of children) OR (menstrual) OR

		(menarche) OR (reproductive) OR (age at first birth) OR (at first pregnancy)))) OR (TI=((breast density OR mammographic density) AND ((parity) OR (number of births) OR (number of children) OR (menstrual) OR (menarche) OR (reproductive) OR (age at first birth) OR (at first pregnancy))))
EMBASE	Alcohol	((breast density or mammographic density) and alcohol).ab. or ((breast density or mammographic density) and alcohol).ti.
	Smoking	((breast density OR mammographic density).ab AND ((smoking) OR (tobacco) OR (lifestyle)).ab) or ((breast density OR mammographic density).ti AND ((smoking) OR (tobacco) OR (lifestyle)).ti)
	All reproductive	(("breast density" or "mammographic density").ti. and (parity or “number of births” or “number of children” or menstrual or menarche or reproductive or "age at first birth" or at "first pregnancy").ti.) or (("breast density" or "mammographic density").ab. and (parity or “number of births” or “number of children” or menstrual or menarche or reproductive or "age at first birth" or at "first pregnancy").ab.)

Appendix 2. Summary of studies included in the systematic review

Study	Country	Population type	Population size	Age	Menopausal status	Method of MD assessment	ALCOHOL	SMOKING	PARITY	AGE AT FIRST BIRTH	AGE AT MENARCHE
Ahern 2021 [324]	USA	Cohort of women high risk for breast cancer (High Risk Breast Program (HRBP) Cohort)	387	median 47 (range 25-76)	all	computer automated					
Peplonska 2021 [325]	Poland	Screening women from two centres	467	range 40-60	all	computer automated					
Lee 2020 [326]	USA	Study cohort of Vietnamese American women	380	mean 54.7 (SD 7.4)	all	visually assessed					
Lee 2020b [187]	Korea	Screening population from Korean National Cancer Screening Program (NCSP)	7,265,584	ages 40+	all	visually assessed					
Moore 2020 [327]	USA	Screening population at single centre	37,839	mean 58	all	visually assessed					
Albeshan 2019 [328]	Saudi Arabia	Women attending screening in Riyadh, multi-centre	792	mean 49.6	all	computer automated					
Kaya 2019 [329]	Turkey	Women attending screening at single centre	494	mean 54.2	post	computer automated					
Moran 2019 [330]	Canada	Women with strong family history of breast cancer at single centre	156	mean 48.3 (range 27-68)	all	computer assisted					
Okamoto 2019 [331]	Japan	Screening women from single centre	785	mean 49.2	all	visually assessed					
Liu 2018 [332]	USA	Controls from nested case-control study within an occupational cohort of nurses (Nurses Health Study II)	1,211	range 25-44	pre	computer assisted					
Sung 2018 [168]	China	Subset of women attending screening in 17 cities, high risk for breast cancer	11,478	mean 54.4 (range 45-69)	all	visually assessed					
Dung 2017 [333]	Vietnam	Mammogram population at two centres in Vietnam	1651	mean 43.8	all	visually assessed					
Jacobsen 2017[144]	Denmark	Women from study cohort attending screening (Danish Diet, Cancer and Health (DCH) cohort)	5,356	mean 56.2 (SD 4.5)	all	visually assessed					
Pereira 2017[334]	Chile	Random sample of women from study cohort (Determinants of Breast Cancer risk cohort)	192	mean 36.6 (SD 6.6)	pre	visually assessed					
Jacobsen 2016[146]	Denmark	Women from study cohort attending screening (Danish Diet, Cancer and Health (DCH) cohort)	5,356	mean 56.2 (SD 4.5)	all	visually assessed					

Study	Country	Population type	Population size	Age	Menopausal status	Method of MD assessment	ALCOHOL	SMOKING	PARITY	AGE AT FIRST BIRTH	AGE AT MENARCHE
McDonald 2016[142]	USA	Women from birth cohorts (Early Determinants of Mammographic Density (EDMD))	697	mean 43.1 (SD 2.3 range 30.4-48.6)	all	computer assisted					
Yang 2016[335]	China	Cohort from screening study (Southern Professional Women Breast Cancer Screening Project)	1,699	mean 43.2 (range 27-55)	pre	visually assessed					
Frydenberg 2015[336]	Norway	Study cohort (Energy Balance and Breast Cancer Aspects study I)	202	mean 30.7	pre	computer assisted					
Quandt 2015[337]	USA	Sample of women from the New York City Multiethnic Breast Cancer Project	187	mean 50 (SD 5.7)	all	computer assisted					
Rice 2015[59]	Mexico	Random sample of women from occupational cohort of teachers (Mexican Teachers' Cohort)	1,607	mean 47.1	all	computer assisted					
Trinh 2015[338]	Sweden	Population-based cohort of women attending one of four units national mammography screening program in Sweden (KARolinska MAmmography Project for Risk reduction of Breast Cancer (KARMA))	53,060	mean 54.8 (SD 9.7)	all	computer automated					
Borugian 2014[339]	Canada	Women attending Screening Mammography Program of British Columbia	299	range 40-49	pre	computer assisted					
Dai 2014[340]	China	Screening trial (Multi-modality Independent Screening Trial (MIST))	28,388	range 45-65	all	visually assessed					
Ahmadinejad 2013[341]	Iran	Women attending screening or diagnostic mammograms at single centre	728	mean 48.12 (range 19-83)	all	visually assessed					
Brand 2013[143]	Sweden	Controls from a nationwide case-control breast cancer study in Sweden (CAHRES study)	1,147	mean 63.7 (SD 6.1)	post	computer assisted					
Obajimi 2012[342]	Nigeria	Screening women from single centre	498	mean 47 (range 20-76)	all	visually assessed					
Qureshi 2012[343]	Norway	Random sample of screening women from two cities from the Norwegian Breast Cancer Screening Program (NBCSP)	2,251	mean 58.4 (SD 5.3)	post	computer assisted					

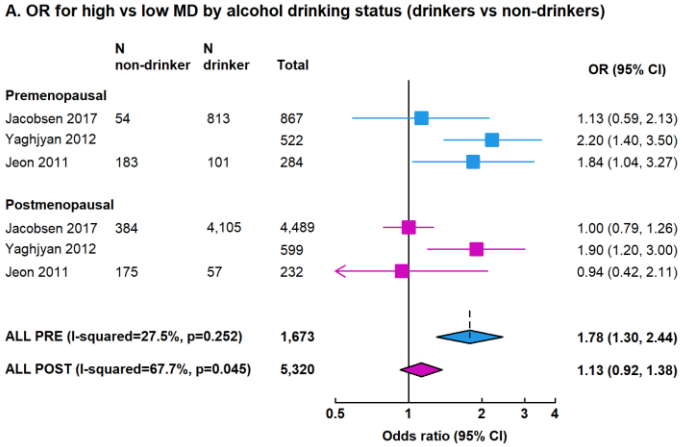
Study	Country	Population type	Population size	Age	Menopausal status	Method of MD assessment	ALCOHOL	SMOKING	PARITY	AGE AT FIRST BIRTH	AGE AT MENARCHE
Yaghjyan 2012[145]	USA	Women from community-based medical surveillance program who lived within five miles of former uranium plant (Fernald Community Cohort)	1,125	mean 51.3	all	visually assessed					
Jeon 2011[344]	Korea	Study cohort of women attending a mammogram in three cities in Korea	516	mean 50.6 (SD 8.8)	all	visually assessed					
Sung 2011[345]	Korea	Monozygotic twin study (Healthy Twin study)	244	mean 39.6 (SD 7.4)	all	computer assisted					
Tehranifar 2011[346]	USA	Sample of women recruited at screening from the New York City Multiethnic Breast Cancer Project	191	mean 50 (SD 5.7 range 39.8-60.9)	all	computer assisted					
Terry 2011[347]	USA	Women from birth cohorts (Early Determinants of Mammographic Density (EDMD))	678	mean 43.2	all	computer assisted					
Tseng 2011[348]	USA	Study cohort of women of Chinese heritage in Philadelphia	201	mean 53.1 (SD 10.2)	all	visually assessed					
Butler 2010[349]	USA	Study cohort of women through menopausal transition (Study of Women's Health Across the Nation (SWAN))	709	mean 47	pre	visually assessed					
McCormack 2010[350]	UK	Random sample of screening women at single centre	645	mean 58.2 (SD 3.3)	all	computer assisted					
Flom 2009[351]	USA	Birth cohort (The New York Women's Birth Cohort)	228	mean 42.38	all	computer assisted					
Stone 2009[352]	Australia	Women at high risk for breast cancer taken at baseline from RCT (International Breast Cancer Intervention Study I trial (IBIS))	799	mean 50.48 (SD 6.16)	all	computer assisted					
Butler 2008[353]	USA	Study cohort of women through menopausal transition (Study of Women's Health Across the Nation (SWAN))	801	mean 47	pre	visually assessed					
Jeffreys 2008[354]	UK	Study cohort of women from Glasgow Alumni Cohort attending breast screening	590	median 54.1 (range 40-71.5)	all	computer automated					
McCormack 2008[355]	UK	Random sample of screening women at single centre	645	mean 57.5	all	computer assisted					

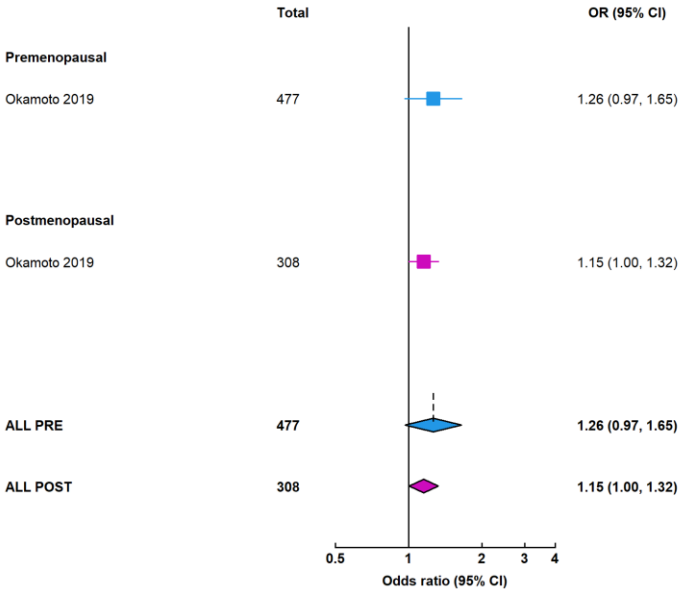
Study	Country	Population type	Population size	Age	Menopausal status	Method of MD assessment	ALCOHOL	SMOKING	PARITY	AGE AT FIRST BIRTH	AGE AT MENARCHE
Tseng 2008 [356]	USA	Cohort of sisters, daughters, nieces, and granddaughters of breast cancer cases (Minnesota Breast Cancer Family Study)	1,286	mean 57	all	computer assisted					
Bremnes 2007[357]	Norway	Women attending Norwegian Breast Cancer Screening Program at a single centre (Tromso Mammography and Breast Cancer Study)	907	range 55-71	post	computer assisted					
Masala 2006[358]	Italy	Women in Florence who were part of the European Prospective Investigation into Cancer and Nutrition study (EPIC)	1,668	mean 53.3	all	visually assessed					
Modugno 2006[359]	USA	Subset of women with screening mammogram from study cohort (Study of Osteoporotic Fractures (SOF))	239	mean 78.6 (range 70-92)	post	computer assisted					
Titus 2006 [166]	USA	Screening population from New Hampshire Mammography Network (NHMN)	144,018	range 30-89	all	visually assessed					
Tseng 2006 [360]	USA	Study cohort of women of Chinese heritage in Philadelphia	212	mean 53.2 (SD 8.6)	all	visually assessed					
Haars 2005[361]	Holland	Random sample of women from population-based screening programme (The Diagnostisch Onderzoek Mammacarcinoom (DOM) project)	418	mean 56.4 (range 49.2-65.8)	post	computer assisted					
Riza 2005[167]	Greece	Screening population of women having first screen (Ormylia Mammography Screening Programme)	4,993	range 40-65	all	visually assessed					
Heng 2004[362]	Singapore	Random sample of women from population-based screening study (Singapore Breast Screening Project)	803	mean 56.68 (SD 4.18)	all	computer assisted					
Jeffreys 2004[363]	UK	Study cohort of women from Glasgow Alumni Cohort attending breast screening	628	median 59	all	visually assessed					
Gapstur 2003[364]	USA	Screening cohort of Hispanic women (Chicago Breast Health Project)	296	range 40-76	all	computer assisted					
Roubidoux 2003[365]	Alaska - USA	Sample of women attending screening at single centre	528	mean 52.1	all	visually assessed					
Warwick 2003[366]	UK	Subset of women at high risk for breast cancer taken at baseline from RCT (International Breast Cancer Intervention Study I trial (IBIS))	102	median 47	all	visually assessed					

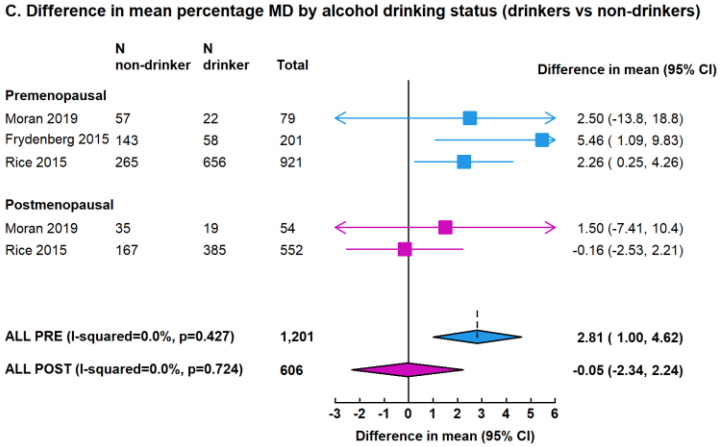
Study	Country	Population type	Population size	Age	Menopausal status	Method of MD assessment	ALCOHOL	SMOKING	PARITY	AGE AT FIRST BIRTH	AGE AT MENARCHE
Maskarinec 2002[367]	Japan & Hawaii - USA	Women attending screening in Gifu, Japan and Honolulu, Hawaii	523	mean 51.8	all	computer assisted					
Soares 2002[368]	Jamaica	Women attending screening or diagnostic mammograms at single centre	891	range 20-88	all	visually assessed					
Maskarinec 2001 [141]	Hawaii - USA	Women attending screening across five centres on Oahu Island, Hawaii	514	mean 53.9 range: 35-85	all	computer assisted					
El-Bastawissi 2000 [184]	USA	Screening cohort in Seattle (Group Health Cooperative of Puget Sound)	28,984	range 20-79	all	visually assessed					
Vachon 2000 [369]	USA	Cohort of sisters, daughters, nieces, and granddaughters of breast cancer cases (Minnesota Breast Cancer Family Study)	1,900	mean 61.7 (SD 11.5)	all	visually assessed					
Jakes 2000[370]	Singapore	Controls from nested case-control within national breast screening study (Singapore Breast Screening Project)	348	range 46-67	all	visually assessed					

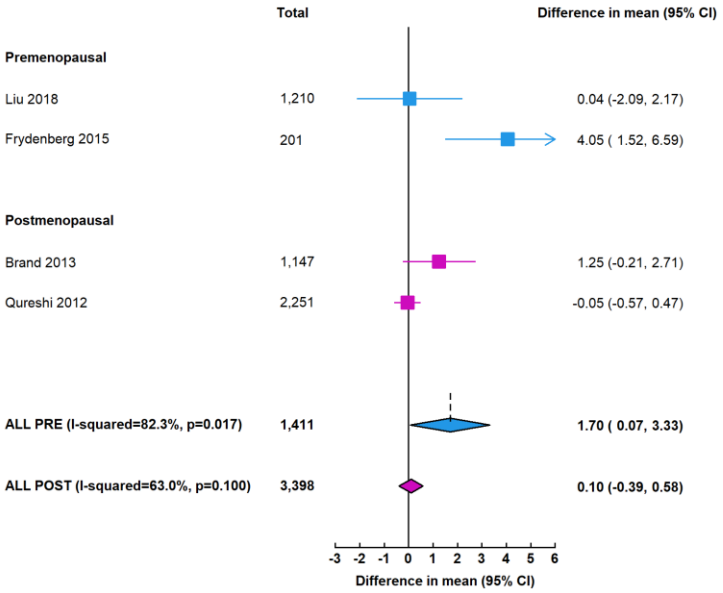
Appendix 3. Relationship of breast cancer risk factors with continuous and categorical measures of mammographic density by menopausal status

Alcohol and mammographic density by menopausal status



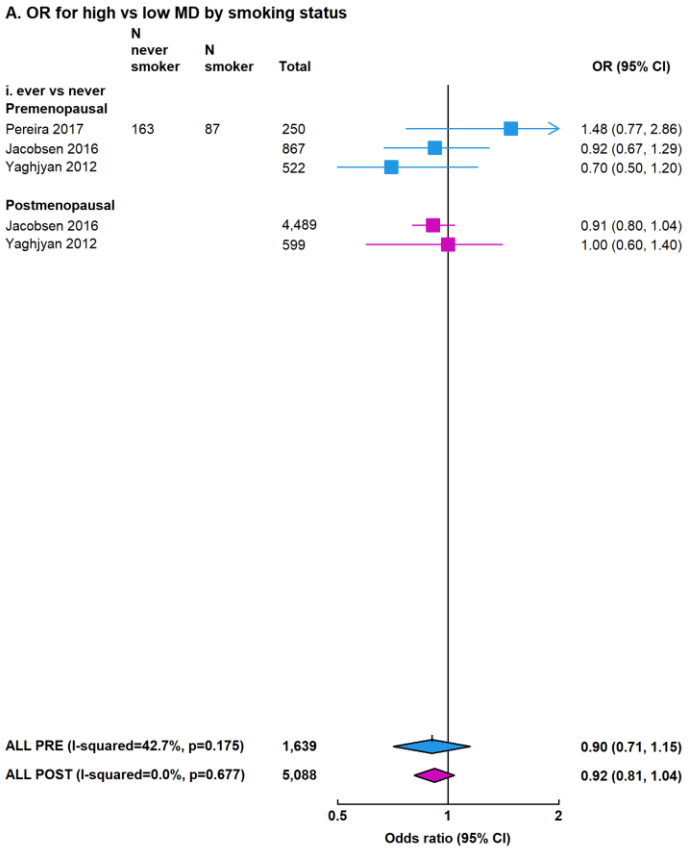


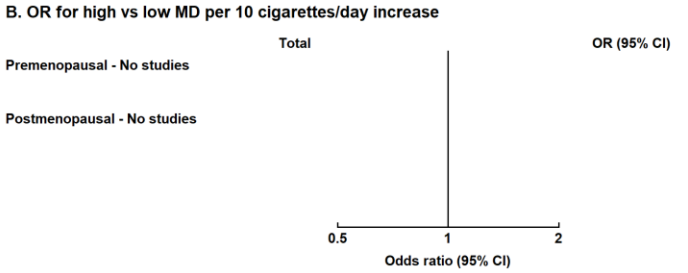


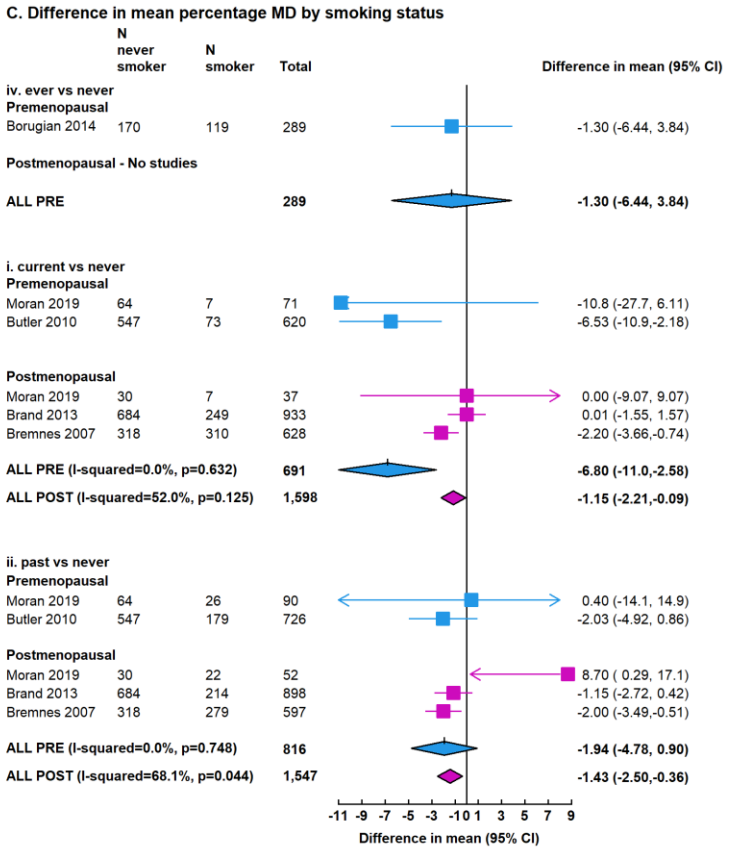


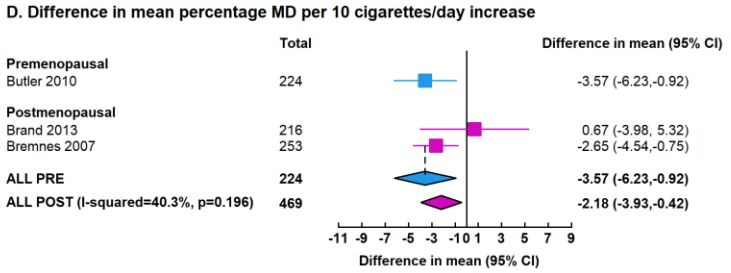
p(heterogeneity for pre- vs postmenopausal): A, 0.017; B, 0.55; C, 0.055; D, 0.065

Smoking and mammographic density by menopausal status



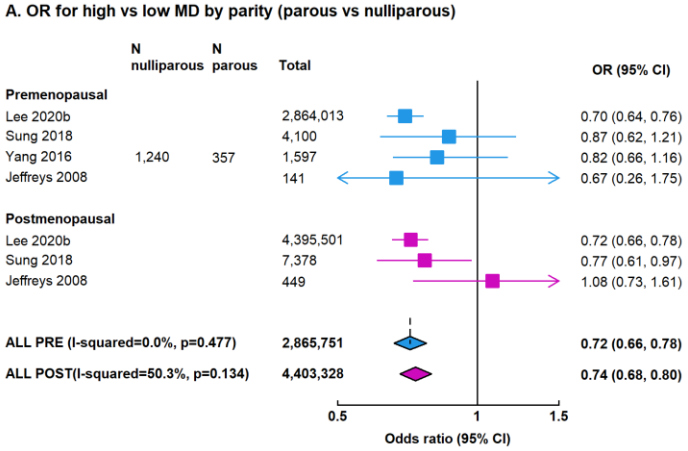


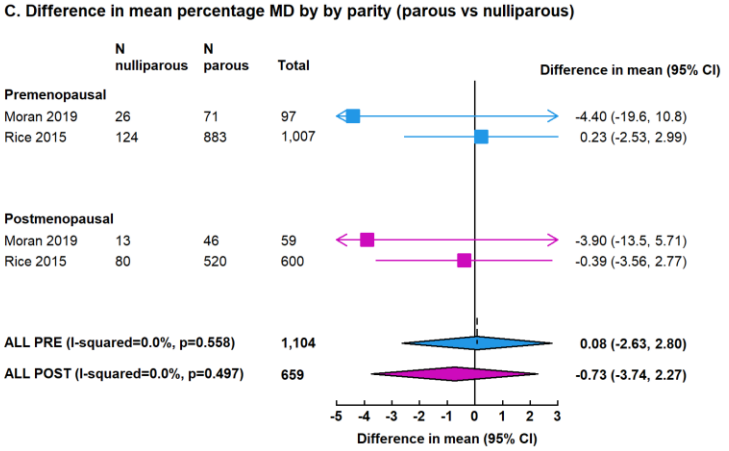


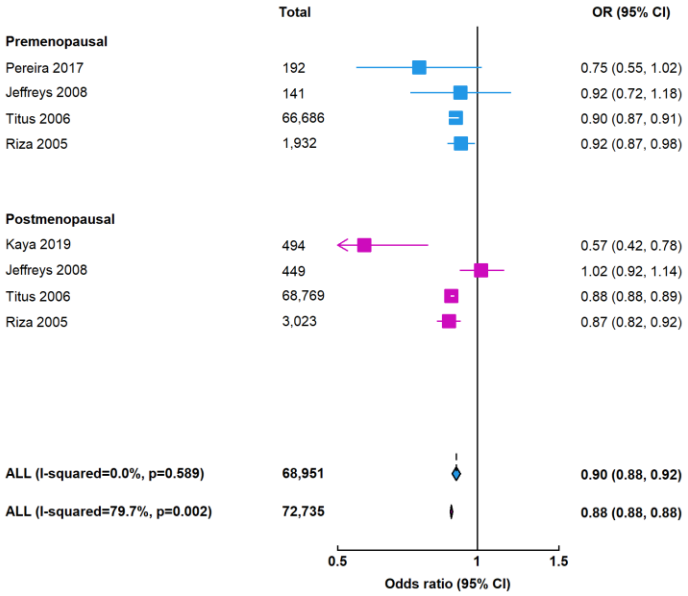


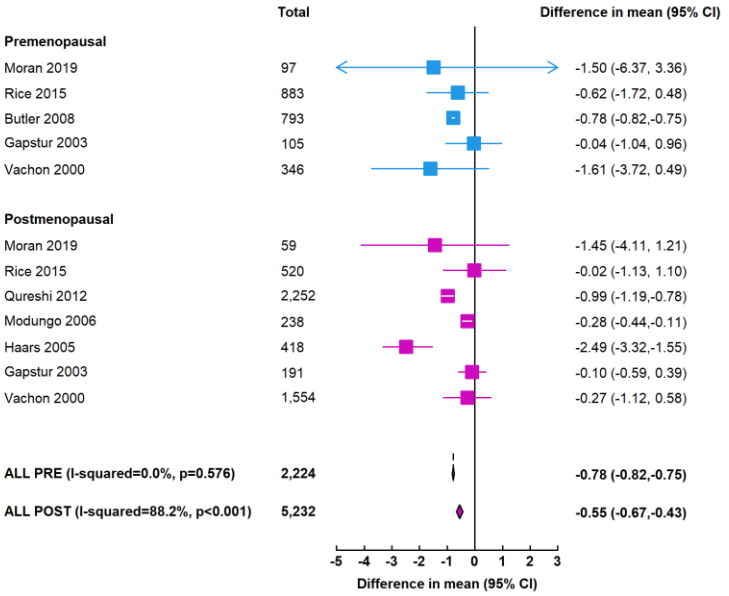
p(heterogeneity for pre- vs postmenopausal): A, 0.87; B, no studies; C, current vs never 0.011; D, 0.39

Parity and mammographic density by menopausal status



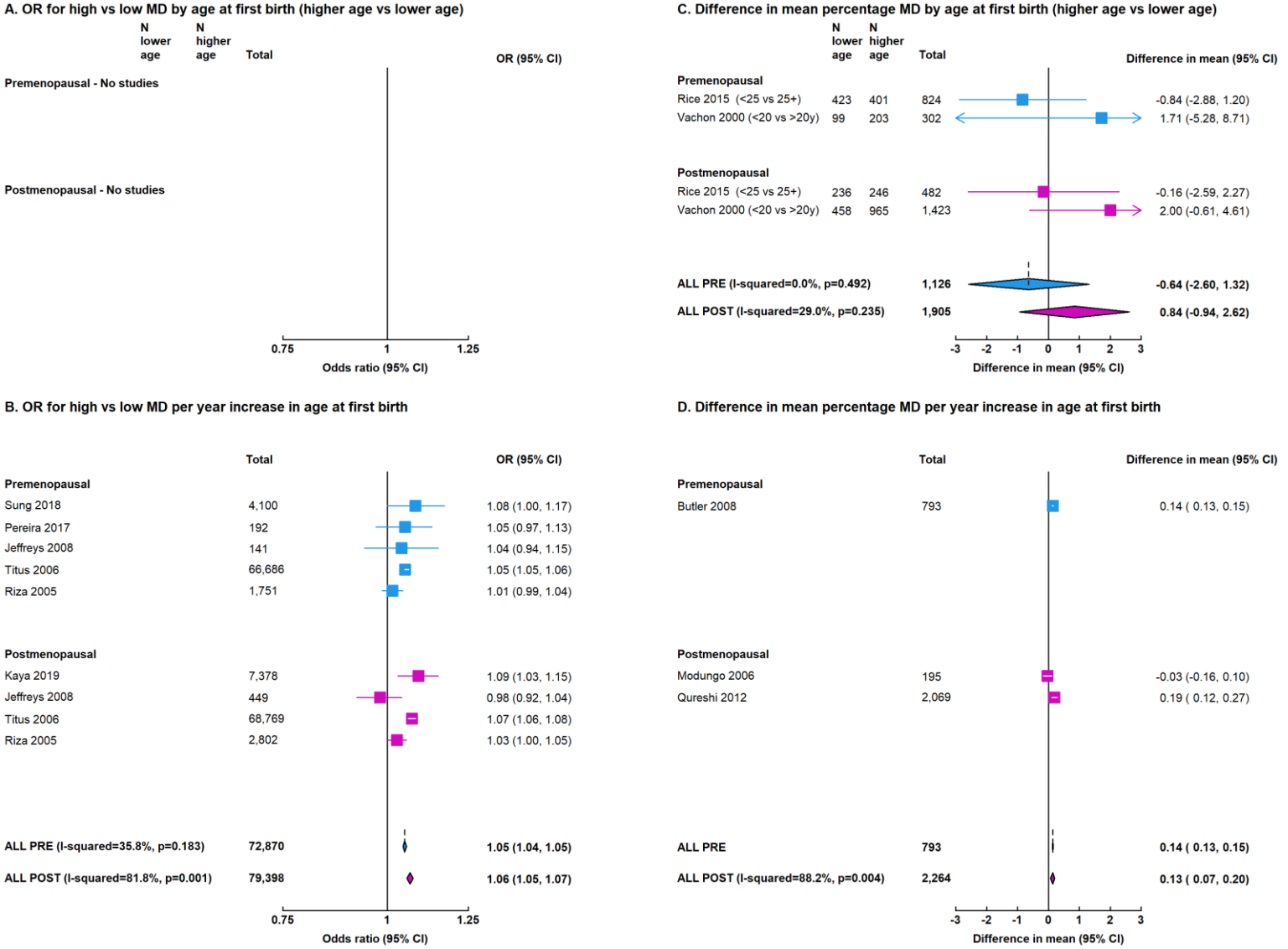






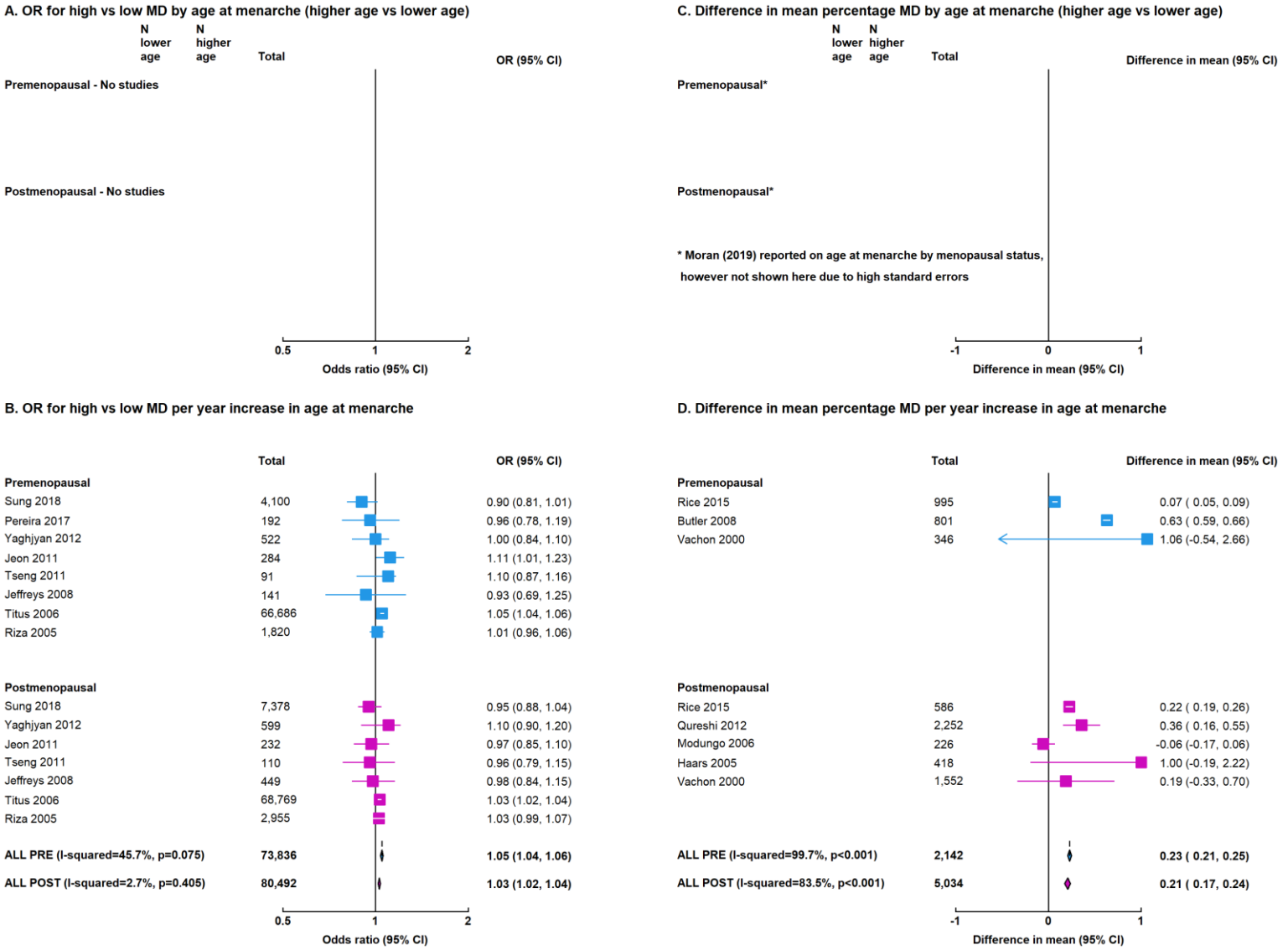
p(heterogeneity for pre- vs postmenopausal): A, 0.65; B, 0.048; C, 0.70; D, <0.001

Age at first birth and mammographic density by menopausal status



p(heterogeneity for pre- vs postmenopausal): A, no studies; B, 0.079; C, 0.27; D, 0.77

Age at menarche and mammographic density by menopausal status



p(heterogeneity for pre- vs postmenopausal): A, no studies; B, 0.006; C, 0.no studies; D, 0.33

Appendix 4. Exploring sources of heterogeneity

- A, B,C,D relate to the different plots in the main figures

- Abbreviations: BMI, body mass index; MHT, menopausal hormone therapy; MS, menopausal status; MD, mammographic density

ALCOHOL	A	Drinkers vs non-drinkers		B	Per 10 g/day		C	Drinkers vs non-drinkers		D	Per 10 g/day	
	N	OR (95% CI)	I^2 (p-value)	N	OR (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)
Adjustments												
Age	1	0.77 (0.36, 1.61)		0			0			1	0.11 (0.01, 0.21)	
+ BMI	2	1.08 (0.95, 1.23)	7.9 (0.30)	0			1	5.46 (1.09, 9.83)		4	3.83 (3.60, 4.07)	95.6 (<0.001)
+ BMI + MHT + MS	3	1.23 (1.09, 1.37)	81.0 (0.005)	2	1.06 (1.02, 1.1)	63.5 (0.098)	2	1.27 (-0.23, 2.77)	0.0 (0.91)	5	-0.08 (-0.12, -0.04)	99.4 (<0.001)
Location												
Africa /Asia/Oceania	2	1.08 (0.95, 1.23)	7.9 (0.298)	1	1.24 (1.02, 1.49)		0			0		
Europe	0			1	1.05 (1.01, 1.09)		1	5.46 (1.09, 9.83)		4	0.10 (0.06, 0.14)	75.1 (0.007)
Americas	3	1.29 (1.13, 1.47)	76.6 (0.014)	0			2	1.27 (-0.23, 2.77)	0.0 (0.91)	6	-0.19 (-0.26, -0.12)	99.7 (<0.001)
Method of MD assessment												
Computer assisted	0			0			0			8	-0.18 (-0.25, -0.11)	99.6 (<0.001)
Computer automated	0			0			0			2	0.10 (0.06, 0.14)	76.8 (0.038)
Visually assessed	6	1.15 (1.06, 1.25)	66.3 (0.011)	2	1.06 (1.02, 1.1)	63.5 (0.098)	3	1.71 (0.29, 3.13)	37.0 (0.20)	0		
ALL STUDIES	6	1.15 (1.06, 1.25)	66.3 (0.011)	2	1.06 (1.02, 1.1)	63.5 (0.098)	3	1.71 (0.29, 3.13)	37.0 (0.20)	10	0.03 (-0.01, 0.06)	99.5 (<0.001)

SMOKING	A	Ever vs never smokers		B	Per 10 cigarettes/day		C	Current vs never smokers		D	Per 10 cigarettes/day	
	N	OR (95% CI)	I^2 (p-value)	N	OR (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	
Adjustments												I^2 (p-value)
Age	0			0			1	-0.05 (-2.79, 2.69)		0		
+ BMI	3	0.89 (0.58, 1.37)	70.4 (0.34)	0			4	-1.12 (-2.00, -0.24)	76.7 (0.005)	3	-0.10 (-0.20, 0.00)	70.1 (0.035)
+ BMI + MHT + MS	3	0.97 (0.88, 1.07)	48.9 (0.14)	1	0.93(0.90, 0.96)		4	-1.20 (-2.25, -0.14)	29.1 (0.24)	1	-2.65 (-4.54, -0.75)	
Location												
Africa /Asia/Oceania	2	1.08 (0.92, 1.27)	38.8 (0.20)	0			1	-4.91 (-8.41, -1.40)		0		
Europe	2	0.89 (0.78, 1.02)	75.5 (0.044)	1	0.93(0.90, 0.96)		3	-0.78 (-1.49, -0.06)	60.3 (0.08)	3	-0.10 (-0.9, -0.002)	71.7 (0.029)
Americas	2	1.00 (0.74-1.36)	42.1 (0.19)	0			5	-2.12 (-3.98, -0.25)	35.8 (0.18)	1	-3.57(-6.23, -0.92)	
Method of MD assessment												
Computer assisted	0			0			6	-1.41 (-2.32, -0.49)	47.4 (0.091)	2	-2.18 (-3.93, -0.42)	40.3 (0.20)
Computer automated	0			0			2	-0.47 (-1.43, 0.49)	0.0 (0.71)	1	-0.09 (-0.19, 0.01)	
Visually assessed	6	0.97 (0.88, 1.07)	53.7 (0.055)	1	0.93(0.90, 0.96)		1	-6.53 (-10.88, -2.18)		1	-3.57 (-6.23, -0.92)	
ALL STUDIES	6	0.97 (0.88, 1.07)	53.7 (0.055)	1	0.93(0.90, 0.96)		9	-1.09 (-1.74, -0.43)	54.8 (0.024)	4	-0.10 (-0.20, -0.010)	78.0 (<0.001)

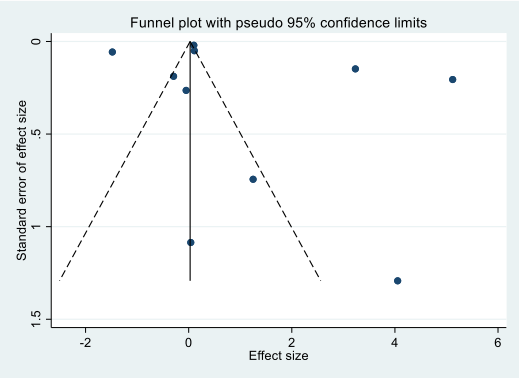
PARITY	A	Parous vs nulliparous		B	Per birth		C	Parous vs nulliparous		D	Per birth	
	N	OR (95% CI)	I^2 (p-value)	N	OR (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	
Adjustments												
Age	2	0.42 (0.19, 0.92)	5.6 (0.30)	2	0.75 (0.68, 0.83)	77 (0.037)	0			0		
+ BMI	3	0.47 (0.36, 0.61)	58.2 (0.92)	8	0.88 (0.86, 0.89)	45.5 (0.076)	2	-1.3 (-2.63, 0.02)	89.6 (0.002)	5	-0.77 (-0.80, -0.73)	92.1 (<0.001)
+ BMI + MHT + MS	2	0.67 (0.62, 0.71)	0.0 (0.72)	5	0.88 (0.88, 0.89)	72.2 (0.006)	3	-1.17 (-2.84, 0.51)	38.8 (0.20)	5	-0.83 (-1.01, -0.65)	74.6 (0.003)
Location												
Africa/Asia/Oceania	4	0.69 (0.65, 0.73)	69.1 (0.012)	5	0.88 (0.86, 0.90)	63 (0.029)	2	-4.05 (-5.98, -2.12)	0 (0.45)	2	-1.25 (-1.9, -0.59)	0 (0.74)
Europe	3	0.76 (0.59, 0.99)	77.5 (0.012)	4	0.89 (0.86, 0.93)	63.2 (0.043)	1	0.01 (-1.56, 1.57)		2	-1.07 (-1.27, -0.86)	90.5 (0.001)
Americas	2	0.67 (0.62, 0.71)	0.0 (0.81)	7	0.88 (0.87, 0.88)	78.8 (<0.001)	3	-2 (-3.57, -0.43)	75 (0.018)	7	-0.76 (-0.79, -0.73)	88.7 (<0.001)
Method of MD assessment												
Computer assisted	0			0			4	-2.25 (-3.64, -0.85)	64.5 (0.038)	8	-0.57 (-0.69, -0.45)	87.3 (<0.001)
Computer automated	1	1.03 (0.72, 1.48)		3	0.92 (0.87, 0.98)	83.7 (0.002)	1	0.01 (-1.56, 1.57)		0		
Visually assessed	9	0.68 (0.65, 0.71)	56.2 (0.019)	13	0.88 (0.87, 0.88)	64.4 (0.001)	0			3	-0.79 (-0.82, -0.75)	73.5 (0.023)
ALL STUDIES	10	0.68 (0.68, 0.71)	61.4 (0.005)	16	0.88 (0.88, 0.88)	68.8 (<0.001)	5	-1.25 (-2.29, -0.21)	68.9 (0.12)	11	-0.77 (-0.8, -0.74)	86.5 (<0.001)

AGE AT FIRST BIRTH	A	Higher vs lower age		B	Per year		C	Higher vs lower age		D	Per year	
	N	OR (95% CI)	I^2 (p-value)	N	OR (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)
Adjustments												
Age	1	2.14 (0.71, 6.46)		1	1.08 (1.02, 1.15)		0			0		
+ BMI	1	1.17 (1.08, 1.27)		6	1.03 (1.01, 1.04)	55.2 (0.048)	1	2.06 (-0.39, 4.50)		4	0.14 (0.13, 0.14)	75.4 (0.007)
+ BMI + MHT + MS	1	1.14 (0.67, 1.92)		4	1.05 (1.04, 1.05)	94.6 (<0.001)	1	-0.56 (-2.12 – 1.00)		3	0.2 (0.13, 0.28)	0.0 (0.55)
Location												
Africa/Asia/Oceania	2	1.17 (1.08, 1.27)	12.5 (0.29)	3	1.04 (1.01, 1.07)	79.8 (0.007)	0			2	0.18 (-0.07, 0.42)	86.6 (0.006)
Europe	0			4	1.02 (1.00, 1.04)	0.0 (0.45)	0			1	0.19 (0.12, 0.27)	
Americas	1	1.14 (0.67, 1.92)		5	1.05 (1.04, 1.05)	92.9 (<0.001)	2	0.2 (-1.12, 1.51)	0.0 (0.48)	4	0.14 (0.13, 0.14)	57.1 (0.072)
Method of MD assessment												
Computer assisted	0			0			1	-0.56 (-2.12 – 1.00)		6	0.13 (0.07, 0.19)	69.8 (0.005)
Computer automated	0			2	0.99 (0.96, 1.02)	0.0 (0.99)	0			0		
Visually assessed	3	1.17 (1.08, 1.27)	0.0 (0.56)	10	1.05 (1.04, 1.05)	86.4 (<0.001)	1	2.06 (-0.39, 4.5)		1	0.14 (0.13, 0.15)	
ALL STUDIES	3	1.17 (1.08, 1.27)	0.0 (0.56)	12	1.04 (1.04, 1.05)	87.1 (<0.001)	2	0.2 (-1.12, 1.51)	0.0 (0.48)	7	0.14 (0.13, 0.14)	63.8 (<0.001)

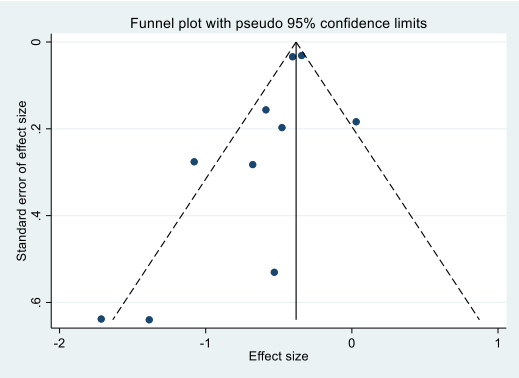
AGE AT MENARCHE	A	Higher vs lower age		B	Per year		C	Higher vs lower age		D	Per year	
	N	OR (95% CI)	I^2 (p-value)	N	OR (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)	N	Diff in mean %MD (95% CI)	I^2 (p-value)
Adjustments												
Age	1	0.75 (0.28, 1.98)		1	1.00 (0.87, 1.16)		0			1	1.10 (0.51, 1.69)	
+ BMI	1	0.97 (0.89, 1.05)		6	1.02 (0.99, 1.05)	65.8 (0.012)	0			6	0.56 (0.53, 0.60)	96.2 (<0.001)
+ BMI + MHT + MS	0			6	1.04 (1.03, 1.05)	0.0 (0.84)	2	0.82 (-1.4, 3.04)	0 (0.65)	4	0.12 (0.1, 0.14)	59.2 (0.061)
Location												
Africa/Asia/Oceania	2	0.97 (0.89, 1.05)	0.0 (0.61)	3	1.01 (0.97, 1.06)	84.3 (0.002)	0			2	0.00 (-0.71, 0.72)	0.0 (0.38)
Europe	0			4	1.02 (0.99, 1.05)	0.0 (0.45)	0			4	0.35 (0.20, 0.50)	0.0 (0.63)
Americas	0			7	1.04 (1.03, 1.05)	0.0 (0.83)	2	0.82 (-1.4, 3.04)	0 (0.65)	5	0.22 (0.21, 0.24)	99.4 (<0.001)
Method of MD assessment												
Computer assisted	0			0			2	0.82 (-1.4, 3.04)	0 (0.65)	8	0.12 (0.1, 0.13)	76.2 (<0.001)
Computer automated	0			2	1.08 (0.99, 1.17)	66.2 (0.085)	0			1	0.27 (0.02, 0.52)	
Visually assessed	2	0.97 (0.89, 1.05)	0.0 (0.61)	12	1.04 (1.03, 1.04)	32.7 (0.13)	0			2	0.63 (0.59, 0.66)	68.0 (0.077)
ALL STUDIES	2	0.97 (0.89, 1.05)	0.0 (0.61)	14	1.04 (1.03, 1.04)	35.3 (0.093)	2	0.82 (-1.4, 3.04)	0 (0.65)	11	0.22 (0.21, 0.24)	98.5 (<0.001)

Appendix 5. Funnel plots (p-value for Egger's test)

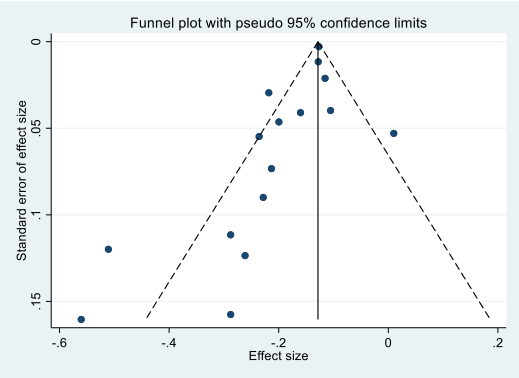
Funnel plot of difference in mean percentage MD per 10 g/day increase in alcohol consumption (p=0.47)



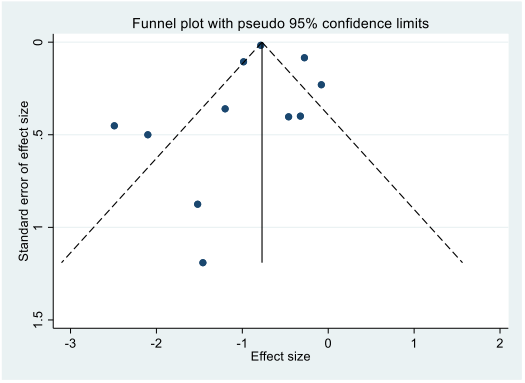
Funnel plot of OR for high vs low MD by parity (p=0.093)



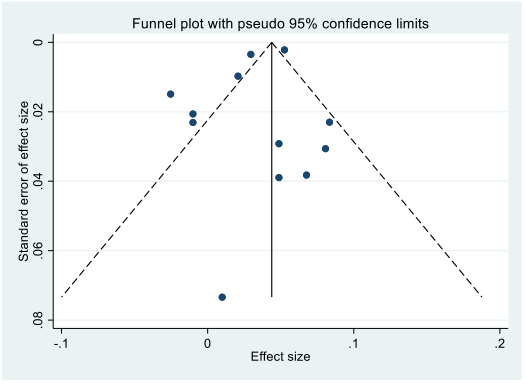
Funnel plot of OR for high vs low MD per birth (p=0.015)



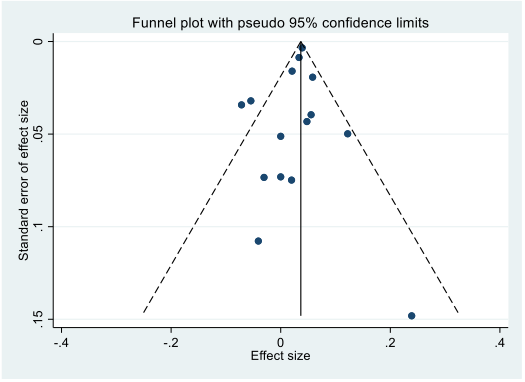
Funnel plot of difference in mean percentage MD per birth (p=0.89)



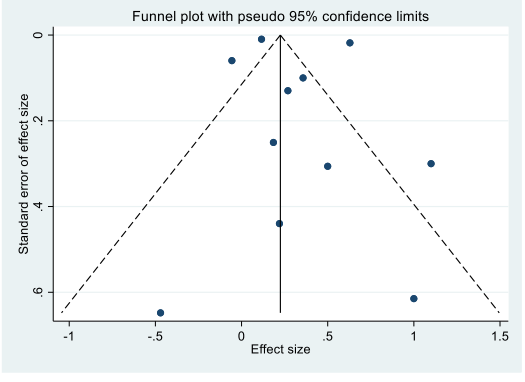
Funnel plot of OR for high vs low MD per year increase in age at first birth (p=0.26)



Funnel plot of OR for high vs low MD per year increase in age at menarche (p=0.31)



Funnel plot of difference in mean percentage MD per year increase in age at menarche (p=0.64)



Chapter 4 appendix

Appendix 6. Sensitivity analysis: breast cancer risk factors and MD in women with known time since menopause

Risk factor	N / mean	Not adjusting for time since menopause	Adjusting for time since menopause
		OR (95% CI)	**OR (95% CI)**
Age			
<65	121	1.00 (ref)	1.00 (ref)
65-69	1654	1.13 (0.7, 1.82)	1.12 (0.7, 1.8)
70+	1520	0.90 (0.55, 1.47)	0.87 (0.54, 1.4)
Per year	70.4	0.97 (0.93, 1.00)	0.96 (0.93, 0.99)
Alcohol (units/week)			
<3	1698	1.00 (ref)	1.00 (ref)
3 to 7	875	1.21 (0.98, 1.51)	1.21 (0.98, 1.51)
>8	713	1.11 (0.88, 1.41)	1.12 (0.88, 1.41)
Per unit	4.7	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Smoking			
Never	1849	1.00 (ref)	1.00 (ref)
Past	1080	1.10 (0.9, 1.35)	1.09 (0.89, 1.34)
Current	333	0.78 (0.57, 1.09)	0.78 (0.56, 1.07)
Exercise (sessions/week)			
<1	659	1.00 (ref)	1.00 (ref)
1	451	1.00 (0.71, 1.42)	1 (0.71, 1.42)
2 to 3	659	1.18 (0.87, 1.61)	1.19 (0.88, 1.61)
4+	927	1.53 (1.16, 2.03)	1.54 (1.16, 2.03)
Family history of breast cancer			
No	2665	1.00 (ref)	1.00 (ref)
Yes	462	1.03 (0.79, 1.33)	1.02 (0.79, 1.33)
Benign breast disease			
No	2895	1.00 (ref)	1.00 (ref)
Yes	384	1.70 (1.31, 2.21)	1.70 (1.31, 2.20)

Risk factor	N / mean	Not adjusting for time since menopause	Adjusting for time since menopause
BMI			
<25	1472	1.00 (ref)	1.00 (ref)
25-29	1174	0.23 (0.19, 0.28)	0.23 (0.19, 0.28)
30+	596	0.08 (0.06, 0.13)	0.08 (0.06, 0.13)
Per unit	26.3	0.74 (0.72, 0.77)	0.74 (0.72, 0.77)
Parity			
0	373	1.00 (ref)	1.00 (ref)
1 to 2	1924	0.53 (0.39, 0.72)	0.53 (0.39, 0.72)
3+	982	0.44 (0.32, 0.62)	0.44 (0.32, 0.62)
Per year	2.1	0.76 (0.68, 0.86)	0.77 (0.68, 0.86)
Age at first birth			
<20	301	1.00 (ref)	1.00 (ref)
20-24	1313	1.23 (0.84, 1.8)	1.23 (0.84, 1.8)
25-29	907	1.57 (1.06, 2.33)	1.56 (1.05, 2.32)
30+	347	1.85 (1.18, 2.89)	1.85 (1.18, 2.89)
Per year	24.4	1.05 (1.02, 1.07)	
Age at menarche			
<12	670	1.00 (ref)	1.00 (ref)
13 to 14	1469	1.02 (0.8, 1.31)	1.02 (0.8, 1.31)
14+	1126	0.91 (0.70, 1.19)	0.91 (0.7, 1.19)
Per year	12.9	0.97 (0.89, 1.06)	0.97 (0.89, 1.06)
Breast feeding			
No		1.00 (ref)	1.00 (ref)
Yes	1344	0.99 (0.74, 1.31)	0.98 (0.74, 1.31)
Oral contraceptive use			
No	1233	1.00 (ref)	1.00 (ref)
Yes	2049	1.09 (0.9, 1.33)	1.09 (0.9, 1.33)
Bilateral oophorectomy			
No	2873	1.00 (ref)	1.00 (ref)
Yes	361	0.66 (0.47, 0.93)	0.66 (0.47, 0.92)

Chapter 5 appendix

Appendix 7. Rules to update MHT use in prospective analysis

Questionnaires answered	Rules for updating MHT use
Recruitment only	• Set never, past or current use to unknown four years after recruitment
Recruitment; 2001 resurvey	• Set never, past or current use to unknown four years after recruitment and update use at 2001 resurvey • Set current use to unknown four years after 2001 resurvey • Never/past users at 2001 resurvey remain in their respective categories until the end of follow-up
Recruitment; 2008 resurvey	• Set never, past or current use to unknown four years after recruitment and update use at 2008 resurvey • Set current use to unknown four years after 2008 resurvey • Never/past users at 2008 resurvey remain in their respective categories until the end of follow-up
Recruitment; 2001 resurvey; 2008 resurvey	• Set never, past or current use to unknown four years after recruitment and update use at 2001 resurvey • Set never, past or current use to unknown four years after 2001 resurvey and update use at 2008 resurvey • Set current use to unknown four years after 2008 resurvey • Never/past users at 2008 resurvey remain in their respective categories until the end of follow-up

Chapter 6 appendix

Appendix 8. Conditional logistic regression model of breast cancer risk factors and breast cancer (time-specific models)

5-years since recruitment

Variable	Coefficient	95% Confidence interval	
MHT status			
Never	ref		
Past	-0.35	-0.45	-0.26
Current	-0.28	-0.38	-0.18
MHT duration (years)			
<1	ref		
1-4	0.10	0.02	0.19
5-9	0.56	0.47	0.66
10-14	1.00	0.90	1.11
≥15	1.24	1.08	1.41
MHT type			
O-only	ref		
O+P	0.27	0.20	0.34
Other	0.38	0.26	0.49
Age at menarche (years)			
<12	ref		
12-13	0.04	-0.02	0.09
14+	0.09	0.04	0.14
Parity			
Nulliparous	ref		
1-2	-0.59	-0.68	-0.50
3+	-0.47	-0.56	-0.39
Age at first birth (years)			
20	ref		
20-24	0.29	0.22	0.36
25-29	0.48	0.40	0.56
30+	0.58	0.49	0.67

Variable	Coefficient	95% Confidence interval	
Age at menopause (years)			
<46	ref		
46-48	0.19	0.11	0.26
49-51	0.25	0.19	0.32
≥52	0.51	0.44	0.58
Hysterectomy	-0.06	-0.16	0.03
Premenopausal	-1.03	-1.18	-0.89
Alcohol units/week			
<3	ref		
3-7	-0.01	-0.06	0.04
>8	0.06	0.01	0.11
Smoking status			
Never	ref		
Past	0.05	0.01	0.10
Current	0.04	-0.01	0.09
Oral contraceptive use			
No	ref		
Yes	-0.42	-0.46	-0.38
Family history breast cancer			
No	ref		
Yes	0.54	0.48	0.60
Benign breast disease			
No	ref		
Yes	0.51	0.46	0.57
BMI			
<25	ref		
25-29	0.15	0.10	0.19
30+	0.20	0.14	0.25
Deprivation quintile			
Q1, least deprived	ref		
Q2	0.01	-0.05	0.07
Q3	0.06	-0.00	0.12
Q4	-0.00	-0.06	0.06
Q5, most deprived	0.04	-0.03	0.10

10-years since recruitment

Variable	Coefficient	95% Confidence interval	
MHT status			
Never	ref		
Past	-0.26	-0.32	-0.20
Current	-0.15	-0.22	-0.09
MHT duration (years)			
<1	ref		
1-4	0.07	0.01	0.12
5-9	0.27	0.22	0.33
10-14	0.49	0.42	0.56
≥15	0.67	0.56	0.78
MHT type			
O-only	ref		
O+P	0.28	0.24	0.33
Other	0.23	0.16	0.31
Age at menarche (years)			
<12	ref		
12-13	0.01	-0.02	0.05
14+	0.02	-0.01	0.06
Parity			
Nulliparous	ref		
1-2	-0.39	-0.45	-0.34
3+	-0.38	-0.44	-0.33
Age at first birth (years)			
20	ref		
20-24	0.14	0.10	0.19
25-29	0.28	0.24	0.33
30+	0.42	0.36	0.47
Age at menopause (years)			
<46	ref		
46-48	0.11	0.04	0.17
49-51	0.18	0.13	0.22
≥52	0.33	0.28	0.38
Hysterectomy	0.04	-0.02	0.10
Premenopausal	-0.24	-0.33	-0.15

Variable	Coefficient	95% Confidence interval	
Alcohol units/week			
<3	ref		
3-7	0.01	-0.02	0.04
>8	0.11	0.08	0.14
Smoking status			
Never	ref		
Past	0.06	0.03	0.09
Current	0.11	0.08	0.15
Oral contraceptive use			
No	ref		
Yes	-0.20	-0.23	-0.18
Family history breast cancer			
No	ref		
Yes	0.51	0.47	0.54
Benign breast disease			
No	ref		
Yes	0.42	0.38	0.45
BMI			
<25	ref		
25-29	0.17	0.14	0.19
30+	0.27	0.23	0.30
Deprivation quintile			
Q1, least deprived	ref		
Q2	0.01	-0.03	0.05
Q3	0.02	-0.02	0.06
Q4	0.01	-0.03	0.04
Q5, most deprived	0.04	-0.00	0.08

Appendix 9. Conditional logistic regression model of breast cancer risk factors and breast cancer (subtype-specific models)

Screen-detected breast cancer model

Variable	Coefficient	95% Confidence interval	
MHT status			
Never	ref		
Past	-0.45	-0.54	-0.36
Current	-0.30	-0.40	-0.21
MHT duration (years)			
<1	ref		
1-4	0.17	0.09	0.25
5-9	0.35	0.26	0.44
10-14	0.38	0.27	0.48
≥15	0.38	0.21	0.56
MHT type			
O-only	ref		
O+P	0.29	0.23	0.35
Other	0.13	0.02	0.23
Age at menarche (years)			
<12	ref		
12-13	-0.01	-0.05	0.04
14+	-0.07	-0.12	-0.02
Parity			
Nulliparous	ref		
1-2	-0.41	-0.49	-0.33
3+	-0.46	-0.54	-0.39
Age at first birth (years)			
20	ref		
20-24	0.14	0.08	0.20
25-29	0.30	0.23	0.37
30+	0.42	0.33	0.50
Age at menopause (years)			
<46	ref		
46-48	0.14	0.05	0.22
49-51	0.24	0.17	0.31
≥52	0.29	0.22	0.36
Hysterectomy	0.07	-0.03	0.17
Premenopausal	0.04	-0.05	0.12

Variable	Coefficient	95% Confidence interval	
Alcohol units/week			
<3	ref		
3-7	0.04	-0.01	0.08
>8	0.15	0.10	0.20
Smoking status			
Never	ref		
Past	0.04	-0.01	0.08
Current	0.10	0.05	0.15
Oral contraceptive use			
No	ref		
Yes	-0.05	-0.08	-0.01
Family history breast cancer			
No	ref		
Yes	0.41	0.35	0.46
Benign breast disease			
No	ref		
Yes	0.31	0.26	0.36
BMI			
<25	ref		
25-29	0.20	0.16	0.24
30+	0.37	0.32	0.42
Deprivation quintile			
Q1, least deprived	ref		
Q2	-0.001	-0.06	0.05
Q3	0.01	-0.05	0.06
Q4	-0.01	-0.07	0.04
Q5, most deprived	-0.05	-0.11	0.01

Interval breast cancer model

Variable	Coefficient	95% Confidence interval	
MHT status			
Never	ref		
Past	-0.34	-0.44	-0.23
Current	-0.03	-0.14	0.08
MHT duration (years)			
<1	ref		
1-4	0.17	0.08	0.27
5-9	0.35	0.26	0.45
10-14	0.54	0.42	0.66
≥15	0.61	0.43	0.79
MHT type			
O-only	ref		
O+P	0.30	0.24	0.37
Other	0.20	0.08	0.32
Age at menarche (years)			
<12	ref		
12-13	-0.01	-0.05	0.04
14+	-0.07	-0.12	-0.02
Parity			
Nulliparous	ref		
1-2	- 0.002	-0.06	0.05
3+	-0.01	-0.06	0.05
Age at first birth (years)			
20	ref		
20-24	0.19	0.12	0.27
25-29	0.31	0.22	0.39
30+	0.47	0.37	0.57
Age at menopause (years)			
<46	ref		
46-48	0.10	-0.04	0.23
49-51	0.15	0.06	0.24
≥52	0.26	0.16	0.36
Hysterectomy	0.05	-0.04	0.14
Premenopausal	0.07	-0.05	0.19

Variable	Coefficient	95% Confidence interval	
Alcohol units/week			
<3	ref		
3-7	0.07	0.01	0.12
>8	0.13	0.07	0.18
Smoking status			
Never	ref		
Past	0.05	-0.01	0.10
Current	-0.04	-0.10	0.02
Oral contraceptive use			
No	ref		
Yes	-0.04	-0.08	0.01
Family history breast cancer			
No	ref		
Yes	0.57	0.51	0.63
Benign breast disease			
No	ref		
Yes	0.52	0.46	0.58
BMI			
<25	ref		
25-29	-0.02	-0.07	0.03
30+	-0.02	-0.08	0.04
Deprivation quintile			
Q1, least deprived	ref		
Q2	0.02	-0.05	0.08
Q3	0.02	-0.04	0.09
Q4	-0.01	-0.07	0.06
Q5, most deprived	-0.06	-0.13	0.01

Low grade breast cancer model

Variable	Coefficient	95% Confidence interval	
MHT status			
Never	ref		
Past	-0.20	-0.25	-0.14
Current	0.08	0.02	0.14
MHT duration (years)			
<1	ref		
1-4	0.04	-0.01	0.09
5-9	0.07	0.02	0.12
10-14	0.12	0.05	0.18
≥15	0.13	0.02	0.23
MHT type			
O-only	ref		
O+P	0.32	0.29	0.36
Other	0.12	0.05	0.19
Age at menarche (years)			
<12	ref		
12-13	-0.02	-0.06	0.01
14+	-0.08	-0.12	-0.05
Parity			
Nulliparous	ref		
1-2	-0.26	-0.31	-0.21
3+	-0.31	-0.36	-0.26
Age at first birth (years)			
20	ref		
20-24	0.02	-0.02	0.06
25-29	0.13	0.08	0.17
30+	0.28	0.23	0.34
Age at menopause (years)			
<46	ref		
46-48	0.11	0.06	0.16
49-51	0.15	0.11	0.20
≥52	0.18	0.13	0.22
Hysterectomy	0.12	0.07	0.16
Premenopausal	0.38	0.32	0.44

Variable	Coefficient	95% Confidence interval	
Alcohol units/week			
<3	ref		
3-7	0.06	0.03	0.09
>8	0.18	0.15	0.21
Smoking status			
Never	ref		
Past	0.07	0.04	0.10
Current	0.24	0.21	0.27
Oral contraceptive use			
No	ref		
Yes	-0.01	-0.03	0.02
Family history breast cancer			
No	ref		
Yes	0.45	0.42	0.49
Benign breast disease			
No	ref		
Yes	0.34	0.31	0.37
BMI			
<25	ref		
25-29	0.13	0.10	0.16
30+	0.27	0.24	0.30
Deprivation quintile			
Q1, least deprived	ref		
Q2	-0.01	-0.04	0.03
Q3	-0.01	-0.04	0.03
Q4	0.003	-0.03	0.04
Q5, most deprived	0.01	-0.02	0.05

High grade breast cancer model

Variable	Coefficient	95% Confidence interval	
MHT status			
Never	ref		
Past	-0.13	-0.22	-0.04
Current	0.02	-0.07	0.11
MHT duration (years)			
<1	ref		
1-4	0.09	0.01	0.17
5-9	0.07	-0.01	0.16
10-14	0.06	-0.05	0.16
≥15	0.17	0.01	0.33
MHT type			
O-only	ref		
O+P	0.14	0.08	0.19
Other	0.04	-0.06	0.14
Age at menarche (years)			
<12	ref		
12-13	-0.03	-0.08	0.02
14+	-0.02	-0.07	0.03
Parity			
Nulliparous	ref		
1-2	-0.26	-0.34	-0.18
3+	-0.30	-0.38	-0.22
Age at first birth (years)			
20	ref		
20-24	0.05	-0.01	0.11
25-29	0.12	0.05	0.19
30+	0.21	0.13	0.30
Age at menopause (years)			
<46	ref		
46-48	0.14	0.06	0.22
49-51	0.14	0.08	0.21
≥52	0.18	0.11	0.25
Hysterectomy	0.11	0.03	0.20
Premenopausal	0.34	0.23	0.45

Variable	Coefficient	95% Confidence interval	
Alcohol units/week			
<3	ref		
3-7	0.02	-0.02	0.07
>8	0.05	- 0.00002	0.10
Smoking status			
Never	ref		
Past	0.06	0.01	0.10
Current	0.15	0.10	0.20
Oral contraceptive use			
No	ref		
Yes	0.05	0.01	0.09
Family history breast cancer			
No	ref		
Yes	0.40	0.34	0.45
Benign breast disease			
No	ref		
Yes	0.29	0.24	0.34
BMI			
<25	ref		
25-29	0.18	0.14	0.23
30+	0.32	0.27	0.37
Deprivation quintile			
Q1, least deprived	ref		
Q2	-0.01	-0.07	0.04
Q3	0.03	-0.03	0.08
Q4	-0.002	-0.06	0.06
Q5, most deprived	0.04	-0.02	0.10