Understanding the role of mammographic density in a population based breast cancer screening program: A step towards stratified screening for breast cancer in Norway?

Nataliia Moshina

Faculty of Medicine

University of Oslo

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-109-1

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Contents

Acknowledgements

First, I would like to thank my research supervisor professor Solveig Hofvind, who offered me the opportunity to become a PhD student and in doing so gave me a great opportunity to develop my scientific thinking and skills. Solveig has been an excellent project and life mentor, supporting me in difficult times and inspiring me to push forward through challenges and uncertainty. I admire Solveig's personality and highly appreciate our research collaboration.

I am very grateful to my co-supervisor professor Giske Ursin for valuable suggestions and feedback, and, well-deserved criticism, encouragement to do additional analyses and revisions.

My statistical analyses would have been impossible without Sofie Sebuødegård, Marta Roman and Kaitlyn Tsuruda. Thank you very much for your guidance and patience!

I would like to thank Gunvor Giplig Waade, an outstanding radiographer and PhD research fellow at the Oslo and Akershus University College of Applied Sciences, for important contributions to and insights on studies on automated density assessment and breast compression.

My deepest gratitude goes to Astri Syse, my master thesis research supervisor, for showing me the way to go and offering support in difficult times.

I am very grateful to the whole Mammography department at the Cancer Registry of Norway for creating the best possible working environment and making the three years of my research unforgettable. I have been very lucky to work and share the most important moments of my life with all of you!

I am thankful to the Cancer Registry of Norway, the institution where the main part of the project has been performed, for giving me the opportunity to acquire knowledge and skills in cancer epidemiology, research and Norwegian working culture.

I would like to thank the Norwegian Breast Cancer Society for applying for funding for this project, and Extrastiftelsen for providing the funding. This research would not be performed without it. I am very grateful to all the representatives of the Norwegian Breast Cancer Society and Extrastiftelsen, who have been communicating with me during the project.

I would like to thank Solveig Roth Hoff, professor Per Skaane and professor Lars A. Akslen for critical reviews and help in solving methodological challenges.

I am grateful to Hilde Trå Hervig, Gry Rosseid, Berit Hanestad and Evy Gran for their willingness to participate in the data collection and valuable assistance in processing the density information for the study on breast compression.

I am very grateful to the University of Oslo for helpful courses in research methods, statistics, scientific writing and career orientation.

Finally, I would like to thank my husband for his patience, understanding and support, as well as for cheering me up when it was essential.

Nataliia Moshina

Oslo, June 2017

Abstract

Mammographic density represents the amount of the epithelium and fibrous tissue in the breast and refers to the radiographic density of the breast visible on mammography. The epithelium and fibrous tissues are radiodense, and appear as white or light gray areas on a mammogram, whereas the fatty tissue is radiolucent, and appears as black or dark gray areas. Mammographic density has been shown to be an independent risk factor for breast cancer. Women with high density (>75% dense tissue) have a four- to six-fold increased risk of breast cancer compared to women with entirely fatty breasts (<25% dense tissue). High mammographic density also significantly decreases the sensitivity of mammography. Furthermore, dense breasts can negatively influence early performance measures of a screening program, resulting in a higher recall rate, missed tumors due to masking and an increased risk of interval cancer.

In this thesis, I examined the role of mammographic density in breast cancer screening in Norway. The goal was to contribute with knowledge that would be helpful in determining whether mammographic density could be used for future stratified screening. The articles in this thesis are based on information about women screened in the Norwegian Breast Cancer Screening Program. We investigated whether mammographic density affects early performance measures of breast cancer screening, which compression parameters are associated with density estimates, and how different mammographic density classifications correspond to each other. Mammographic density was assessed both subjectively, by breast radiologists working in the screening program, and objectively, using a fully automated method.

We found that positive predictive values for recall examinations and invasive procedures decreased with increasing mammographic density among women screened in the program. We also determined that high mammographic density was associated with large (>15 mm) tumor size and positive lymph node status in women with screen-detected invasive breast cancer. Further, we identified correlations between compression force, pressure, compressed breast thickness, breast volume and fibroglandular volume, and volumetric breast density. The strongest associations were observed between compression pressure, breast volume and fibroglandular volume, and between compressed breast thickness and volumetric breast density. We found that subjective mammographic density classifications used by the

screening program in Norway corresponded well to estimates of the fully automated density assessment method.

The results of our studies indicated that mammographic density could be a useful tool in stratification of breast cancer screening. However, the obtained evidence is currently not sufficient to support stratified screening for breast cancer based on mammographic density in Norway.

List of papers

Paper I

Moshina N, Ursin G, Roman M, Sebuødegård S, Hofvind S. Positive predictive values by mammographic density and screening mode in the Norwegian Breast Cancer Screening Program. Eur J Radiol 2016; 85(1):248-54.

Paper II

Moshina N, Ursin G, Hoff SR, Akslen LA, Roman M, Sebuødegård S, Hofvind S. Mammographic density and histopathologic characteristics of screen-detected tumors in the Norwegian Breast Cancer Screening Program. Acta Radiol Open 2015; 4(9) 50.

Paper III

Moshina N, Roman M, Waade GG, Sebuødegård S, Ursin G, Hofvind S. Breast compression parameters and mammographic density in the Norwegian Breast Cancer Screening Program (submitted to European Radiology, January 2017, under revision).

Paper IV

Moshina N, Roman M, Sebuødegård S, Waade GG, Ursin G, Hofvind S. Comparison of subjective and fully automated methods for measuring mammographic density. Acta Radiol 2017 [in press].

Abbreviations

ABUS – automated whole breast ultrasound

BI-RADS – breast imaging-reporting and data system

BMI – body mass index

CC – craniocaudal

CI – confidence interval

DBT – digital breast tomosynthesis

DCIS – ductal carcinoma *in situ*

DICOM – digital imaging and communications in medicine

FFDM – full-field digital mammography

IARC – international agency for research on cancer

MLO – mediolateral oblique

MRI – magnetic resonance imaging

OR – odds ratio

PPV-1 – positive predictive value for recall examinations

PPV-2 – positive predictive value for invasive procedures

REC – regional committees for medical and health research ethics

SD – standard deviation

SFM – screen-film mammography

VDG – Volpara density grade

1. Introduction

1.1 Breast cancer

1.1.1 Breast cancer incidence, survival and mortality

Breast cancer is the most common type of cancer among women worldwide (1). In 2012, the world age-standardized incidence rate of breast cancer was 43.3 per 100,000 person-years (2, 3). In Norway, the age-standardized incidence rate for breast cancer was 128.0 per 100,000 person-years in 2015. That same year, 3,415 Norwegian women were diagnosed with invasive breast cancer and approximately 300 women were diagnosed with Ductal Carcinoma *In Situ* (DCIS) (4).

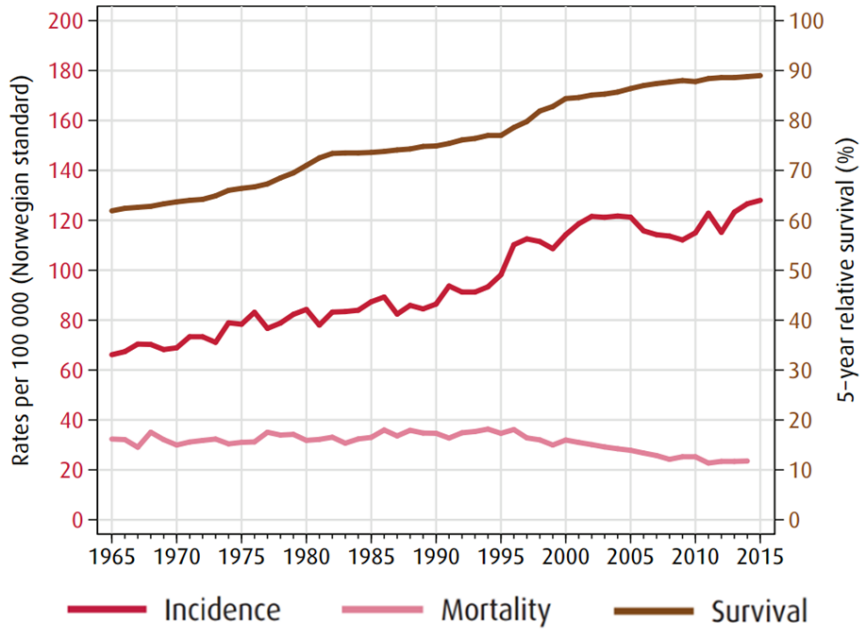

Figure 1. Trends in breast cancer incidence and mortality rates and 5-year relative survival proportions for the Norwegian female population, 1965-2015 (source: Cancer in Norway 2015).

Increased attention to breast cancer coupled with improvements in diagnostics during the last decades has resulted in more women seeking help and advice for breast symptoms, leading to an increased incidence of the disease (1), also in Norway. This has resulted in better secondary prevention and treatment, leading to an increase in 5-year relative survival, and a decrease in breast cancer mortality in Norway (Figure 1). The age-standardized mortality rate for breast cancer in Norway was 23.5 per 100,000 person-years in 2015, and the 5-year relative survival for breast cancer in Norway during the period 2011-2015 was 89.0% (95%CI 88.3-89.7) for all stages and 100.2% for stage I disease.

1.1.2 Breast cancer risk factors

Breast cancer risk factors can be non-modifiable and modifiable. Non-modifiable factors include gender, age, family history, age at menarche, age at menopause, atypical hyperplasia or borderline lesion confirmed histologically, genetic mutations and nucleotide polymorphisms (5-11). Gender represents a very strong risk factor for the disease (4). It is well established that the risk of most cancers, including breast cancer, increases with age (12-14), which makes age another strong non-modifiable breast cancer risk factor (5, 15-17). Studies have shown that early age at menarche and late age at menopause are associated with a higher risk of breast cancer (9, 18). The relative risk of breast cancer associated with selected factors is provided in Table 1.

Modifiable breast cancer risk factors include parity, breastfeeding, age at first birth, mammographic density, use of combined estrogen-progesterone hormonal therapy, body mass index (BMI), alcohol consumption, tobacco smoking, physical activity, diet and exposure to ionizing radiation (9, 19-29). An increase in number of live births and years of breastfeeding are associated with a lower risk of breast cancer (28). Age at first birth of 35 years or more is associated with a higher relative risk of breast cancer (20). Use of combined estrogen-progesterone hormonal therapy for five or more years increases relative risk of breast cancer among postmenopausal women (22, 30). A high body mass index (BMI) (>32 kg/m^2) is related to an increased relative risk of breast cancer among postmenopausal women (24). Alcohol use and tobacco smoking have been reported to be associated with a higher relative risk of breast cancer compared with no use and no smoking, respectively (25, 29). Physical activity, as well as diets including vitamin A, carotenoids and folate, may have protective effect on the risk of breast cancer; however, studies on this topic have been inconsistent (26). Ionizing radiation has been reported to be related to increased breast cancer risk, particularly for women exposed to radiation in young age (27).

Table 1. Relative risk of breast cancer associated with selected risk factors, based on the results of meta-analyses

Risk factor	Relative risk for breast cancer (95% confidence interval)
Breast density (>75% versus <5%) (19)	4.6 (3.6-5.9)
Histologically verified atypical hyperplasia (atypical hyperplasia versus normal breast tissue) (11)	3.9 (3.2-4.8)
Family history of breast cancer (first degree relative with breast cancer versus no family history) (6)	2.1 (2.0-2.2)

Breast cancer risk factors play an important role in risk prediction and have been used to create breast cancer risk models during the last decades (31). Accurate risk models are needed to identify women with the highest risk of developing breast cancer. Several risk models have been described so far (31). Adding new genetic factors and information about mammographic density may improve these models (7, 32), but they all still lack discriminatory power (31).

As many factors can contribute to breast cancer development, it is essential to carry out preventive procedures, which might help decrease breast cancer incidence rates. Primary prevention of breast cancer includes eliminating the risk factors and by that decreasing the probability of the disease occurrence. However, breast cancer risk factors are numerous and often difficult to modify. Therefore, secondary prevention, including detection of the tumor in an early stage of the disease plays an important role in the disease control.

1.2 Screening for breast cancer

1.2.1 Rationale for breast cancer screening: historical screening criteria

In 1951, the United States Commission of Chronic Illness defined screening as "the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly" (28). Further, the World Health Organization (WHO) guidelines published in 1968, often referred to as Wilson's Criteria, defined the essence of screening, including a recognizable latent or early symptomatic stage, the availability of an appropriate screening test and an accepted treatment (33).

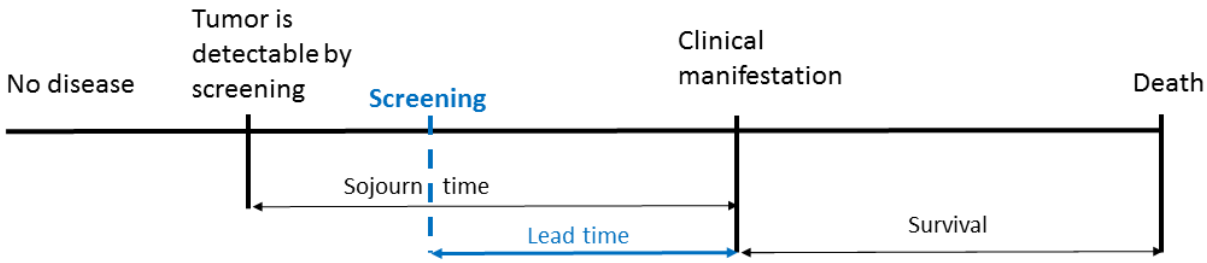

Figure 2. Overview of disease progression and role of screening for breast cancer (34). ***Sojourn time*** *is the period during which a breast tumor can be detected by screening (mammography) and has no signs of clinical symptoms (35).* ***Lead time*** *is the time gained due to detecting a tumor using screening prior to its clinical manifestation (35).*

Evidence of today suggests that breast cancer usually develops from an early stage with a small tumor to a more advanced stage with a large tumor, involvement of lymph nodes and metastases in different organs (36). Prognosis has been shown to be less favorable for late stage compared with early stage breast cancer (37-39). The disease is typically characterized by a long preclinical stage that can be successfully treated, thereby increasing life expectancy due to early detection (Figure 2). The average preclinical stage for breast cancer (sojourn time), has been estimated to be approximately 3 years, ranging from 1 to 8 years (34, 40, 41). Screening for breast cancer is primarily mammography. Several aspects, including acceptable radiation exposure and costs, as well as a high probability of identifying breast cancer on an early stage, make mammography an appropriate screening tool for breast cancer.

1.2.2 Mammographic screening

A number of randomized control trials have demonstrated a reduction in breast cancer mortality associated with mammographic screening (42). Mammographic screening as a secondary prevention for breast cancer is stated to have sufficient evidence of efficacy by several organizations and institutions, including WHO, National Cancer Institute (NCI), International Agency on Cancer Research (IARC), European Commission Initiative on Breast Cancer and European Society of Breast Imaging (1, 43-46). Mammographic screening has thus been established worldwide (28, 47, 48).

Organized mammographic screening represents a service for breast cancer control and has been developed according to the screening criteria of the WHO (49). These criteria include definition of the screening objectives and target population, evidence of effectiveness, integration of education and clinical services, quality assurance, informed choice, equal access, planned evaluation and prevailing of benefits compared with harms (49). In Norway, six more criteria are suggested (Table 2).

Table 2. The Norwegian additional criteria for breast cancer screening (50)

N	Criteria
1	The benefits should outweigh the harms
2	Personal and legal aspects should be ensured
3	The screening program should be acceptable from an ethical point of view
4	Information about the screening program should be evidence based and facilitate an informed choice about participation
5	The screening program should satisfy requirements related to cost effectiveness
6	A plan for administration, quality assurance and evaluation should be available

Organized breast cancer screening implies comprehensive evaluation of the performance and appropriate quality assurance to minimize potential harms (49). Breast cancer screening is a complex multidisciplinary process, involving the evaluation of various performance measures, as the detection rate of breast cancer, sensitivity and specificity (51). Further, participation rate, equipment used and organization of the screening process are factors that greatly influence screening performance (52, 53). Success of a breast cancer screening program is judged not only by the outcome and its impact on public health, but also by its organization, implementation, execution and acceptability to stakeholders (54-56). The evaluation of performance measures represent an essential determinant of effective improvement and future development of mammography screening programs.

1.2.3 Benefits and harms of mammographic screening

Mammography is a non-invasive and readily available method for breast cancer detection, with sensitivity of 70-90% and specificity of 80-100% (55, 57). These two performance measures are, however, difficult to assess because these are not based on an individual level data as for instance positive predictive values (PPVs) or rates of interval cancer.

Mammographic screening provides a benefit of detection of breast cancer in an early stage of the disease, which is associated with favorable prognostic and predictive tumor characteristics (58) and less aggressive treatment (57, 59). Studies have reported 30-40% lower mortality from breast cancer among participants versus non-participants and 20-30% lower mortality for women invited to screening versus non-participants (60-63).

Furthermore, economic analyses have shown screening to be cost-effective, as treatment and disability related to breast cancer detected by screening is associated with lower costs compared to treatment and disability related to breast cancer detected without screening (59, 64, 65).

One of the main harms of mammographic screening is psychological aspects related to false positive screening results (66-68). False positive results have been shown to cause anxiety and distress immediately after the announcement of the result and for at least three years thereafter (69, 70).

Interval cancers, or cancers detected after a normal screening result but before the next screening examination, are considered a serious limitation of mammography as these cancers

are either overlooked at screening (false negative results) or fast growing and thus associated with a poor prognosis (38, 52).

Another negative aspect of mammographic screening is detection of slow growing tumors that never would have caused symptoms during a woman's lifetime if she had not attended screening, or so called overdiagnosis (57, 71). Rates of overdiagnosis have been estimated to range from zero to over 50%, as assumptions and methods for its estimation vary (53, 72, 73).

Radiation exposure is also considered a disadvantage of mammographic screening (47, 57). However, to date, studies on radiation-induced breast cancers have been based on modelled outcomes, reporting that the risk of a radiation-induced breast cancer or breast cancer death is negligible (74).

Pain and discomfort associated with compression of the breast are well-known limitations of screening mammography, which have been reported to be a possible reason for non-attendance in some screening programs (75).

Breast cancer screening has evolved over time. Full-field digital mammography (FFDM) has replaced screen-film mammography (SFM) (76), and mammographic equipment vendors currently offer different set-ups with respect to image acquisition parameters and radiation dose. In addition, several imaging methods have been developed, including hand-held ultrasound, digital breast tomosynthesis (DBT), automated whole breast ultrasound (ABUS) and magnetic resonance imaging (MRI) (76-78). The introduction of these technologies, as well as changes in breast cancer treatment strategies (79), could have affected the evaluation of benefits and harms of breast cancer screening (57).

1.2.4 Mammography: image acquisition and breast compression

Mammography is a low-energy x-ray method for breast cancer detection (80). A mammography image acquisition system is composed of an x-ray tube for the generation of a photon beam, a breast compression paddle and an image receptor system (81). Image acquisition parameters include compression force, compressed breast thickness, x-ray tube current, x-ray tube voltage peak, and anode and beam filtration material (80).

Compression force (newton, N) is the force applied to the breast placed between the image receptor system and a compression paddle of the x-ray machine during the imaging

procedure. Compressed breast thickness (millimeter, mm) refers to the distance between the the image receptor system and the paddle measured at exposure. During the exposure, the x-ray tube current is set as milliampere per second. The x-ray tube voltage peak is the maximum voltage applied across the tube. The x-ray tube anode material and beam filtration material are factors determining the x-ray spectrum (81). Image acquisition parameters differ across exposures and vendors with the aim of reducing the effect of radiation and generating a clear image for the reader (80, 82). X-ray tube current and voltage peak, anode and beam filtration material, and radiation dose are usually set automatically by the automated exposure control, whereas compression force is set by the radiographers conducting mammographic examinations (56).

Breast compression during mammography is argued to be one of the most important prerequisites of image quality (56, 83, 84). Application of compression force to the breast during image acquisition immobilizes the breast and reduces breast thickness, which limits scatter effects and decreases radiation absorbed by the glandular tissue (80).

1.2.5 Breast cancer screening in Norway

About one third of all invasive breast cancers in Norway are detected among participants of the Norwegian Breast Cancer Screening Program. The organized population based program started as a pilot in four counties in 1996 and expanded nationwide, covering all 19 counties by 2005. The program is administered by the Cancer Registry of Norway and serves approximately 600,000 women aged 50-69 years, who are invited to two-view mammography biennially. The transition from SFM to FFDM had been carried out during 2005-2011. As of today, all 30 screening units operating in the program are equipped with FFDM.
The attendance rate is about 75% for each screening round (85). Approximately 3% of women are recalled for further assessment, which includes additional mammograms and potentially ultrasound, MRI, and/or image-guided needle biopsy (85). Breast biopsy is performed in about 40% of women recalled for further assessment after a positive mammogram, and about 50% of women who undergo a biopsy are diagnosed with breast cancer (85).

Two breast radiologists read screening mammograms independently and give a score for each breast indicating the susceptibility of malignancies (86). A score of 1 indicates a normal mammogram; 2 - probably benign; 3 - intermediate; 4 - probably malignant; and 5 - high susceptibility of breast cancer. All cases with a score of 2 or higher by one or both

radiologists are discussed at a consensus meeting, where a decision whether or not to recall the women for further assessment is made.

One of the major advantages of the Norwegian Breast Cancer Screening Program is the national screening databases with availability to complete data and thus possibilities to perform quality assurance and high quality research. Results of early performance measures as well as estimates of mortality, false positive recalls and overdiagnosis based on data from the program are heavily documented both from researchers at the Cancer Registry and from external researchers (55, 60, 61, 63, 87, 88).

1.3 Mammographic density

1.3.1 Breast anatomy and mammographic density

The breast consists of 15-20 lobes. Each lobe has a system of ever-branching ducts ending blindly in a network of terminal ductules. The ducts are lined by the epithelial and basal cells. The breast epithelium is a functional part of the breast. Lobes with ducts and ductules correspond to glandular tissue of the breast. The breast also consists of skin and subcutaneous tissue, fatty and fibrous connective tissue, and stromal elements, such as blood vessels, lymph nodes and vessels, nerves and ligaments (89). The proportions of glandular, fatty, fibrous and stromal components vary among women (90). During menopause, the large amount of epithelium diminishes as it involutes and is replaced by fatty tissue.

Mammographic density reflects the appearance of various tissues presented in the breast on the mammogram (82). Fibroglandular tissue, including fibrous, stromal and glandular components, is radiodense and appears white or light gray on the mammogram, whereas fatty tissue is radiolucent, and appears black or dark gray.

John Wolfe was the first to propose that the mammographic appearance of the breast is related to breast cancer risk (91). Mammographic density is currently a well-established breast cancer risk factor (19). Women with extremely high mammographic density have 4-6-fold higher risk of developing breast cancer compared with women with low mammographic density. In addition, it is more difficult to detect breast cancer among women with high mammographic density compared with low mammographic density, because the tumor has a mammographic appearance similar to fibroglandular tissue (92).

1.3.2 Assessment of mammographic density

In 1976, John Wolfe created the first classification of mammographic density according to risk for breast cancer. This classification included four categories: N1, fatty breast; P1, ≤25% ductal prominence in the breast; P2, >25% ductal prominence in the breast; and DY, dysplastic breast with sheets of dense parenchyma (91). Wolfe reported that the women classified with DY had a 37-fold higher incidence of breast cancer compared with those classified with N1; however these strong results have never been replicated (91).

Methods for mammographic density assessment have undergone various changes over the last 40 years. As of today, two main approaches, qualitative and quantitative, are used for mammographic density assessment. The qualitative approach implies subjective visual evaluation of density on the mammogram assigning it with a score depending on the measurement scale. The Breast Imaging-Reporting and Data System (BI-RADS) classification (93) is the most common qualitative method used in clinical and screening practice for reporting mammographic density (see Figure 4, Chapter 3.1.2). However, subjective assignment is a time-consuming process associated with substantial differences in the scores depending on the reader (82, 94, 95).

The quantitative approach was introduced to eliminate inter-reader variability and increase precision in mammographic density assessment. This approach is characterized by computerized evaluation of mammographic density and includes area-based and volumetric methods. Area-based methods represent a two-dimensional assessment of breast composition (15, 32, 82, 96-98) and frequently estimate mammographic density by means of segmentation of areas on the acquired mammogram in accordance with a reference value determined by the reader or a semi- or fully-automated computer program (96, 99-101). Volumetric methods estimate density using information about x-ray attenuation characteristics (102, 103) or breast thickness from each pixel value of the mammographic image (99, 103). Volumetric methods are fully automated, which eliminates subjectivity, substantially reduces time used for density assessment and allows evaluating the volume of fibroglandular tissue in the breast (82).

In the Norwegian Breast Cancer Screening Program, the radiologists have been assessing mammographic density using two different subjective classification methods: a three-point scale and BI-RADS (see Chapter 4 for further information). In addition, a fully automated method of density assessment has been used at four breast centers for quality assurance and research within a limited time period (104). The subjective mammographic density

classifications have never been validated in relation to a fully automated method and, therefore, the accuracy of the Norwegian radiologists with respect to mammographic density assessment has not previously been investigated in the program.

1.3.3 Mammographic density and screening performance

High mammographic density is associated with decreased mammographic sensitivity (92, 105, 106). However, less attention has been paid to the impact of mammographic density on the performance measures of breast cancer screening programs, including PPV for recall examinations or invasive procedures (107) and histopathologic characteristics of screen-detected tumors (92, 108, 109). PPV is considered a measure of radiologists' performance and thus an indicator of the effectiveness of a screening program (56, 110, 111). Screened women who are recalled for further assessment that turns out to be negative are deemed to have a false positive screening result.

Breast cancer detected in mammographic dense breast is often associated with less favorable histopathologic prognostic tumor characteristics, such as larger tumor size, higher histologic grade and lymph node involvement (108, 109, 112, 113). These associations have not been studied among Norwegian women. Gaining knowledge on the impact of mammographic density on the screening performance measures, including histopathologic tumor characteristics, is needed to maintain and potentially improve the effectiveness of the Norwegian Breast Cancer Screening Program.

1.3.4 Factors affecting mammographic density and its assessment

Mammographic density of a woman can be a dynamic characteristic. It may decrease with age and during menopause due to physiological changes in the breast including involution of ducts and replacement of fibroglandular tissue by fatty tissue (114). Most of the risk factors for breast cancer have been shown to be associated with mammographic density (115). Parity and increased number of live births are associated with low mammographic density, whereas late age at first birth is associated with high mammographic density (23, 115). BMI is highly inversely associated with mammographic density (116). Postmenopausal hormonal therapy with combined estrogen-progesterone increases mammographic density, whereas tamoxifen, a selective estrogen receptor modulator, reduces mammographic density (117, 118). It has also

been shown that several serum growth factors, including insulin growth factor-I, transforming growth factor-β and tumor necrosis factor-α (119-121), as well as variants in several genes in the hormone metabolism, are associated with mammographic density (122, 123).

Parameters related to breast compression and image acquisition (compression force, compressed breast thickness, x-ray tube current and voltage peak and radiation dose) in mammography are hypothesized to affect mammographic density assessment (124) as these parameters may change the representation of density on the mammogram. Previous studies have shown compression force and compressed breast thickness to be correlated with mammographic density estimates obtained from area-based and fully-automated methods of assessment (124, 125). However, further studies including a larger number of examinations are needed to verify these results.

Mammographic density has been studied over the last three decades, and despite the many factors affecting it, the independent association between mammographic density and breast cancer risk has not been disproved (19, 97). Therefore, there is a need to gain more knowledge and thereafter consider mammographic density an important parameter in risk prediction models and stratified breast cancer screening (126-128).

1.3.5 Possible role of mammographic density in stratified breast cancer screening

Stratified breast cancer screening refers to dividing the screening population into groups and aims to intensify screening in a minority of higher risk women by increasing the frequency of mammography or by adjunction of other screening tools to mammography (129). It has been hypothesized that stratified screening could be a relevant approach for improving the screening impact on breast cancer mortality without increasing costs and harms for the majority of women and society (129). Stratification is aimed at reducing the burden of screening in a majority of women with lower risk for breast cancer if they are offered less frequent screening.

Considering the negative effect of density on mammographic sensitivity, the harms of breast cancer screening for women with dense breasts could easily outweigh the benefits (92, 106, 130). New preventive approaches for women with dense breasts have recently been discussed (77, 131). During the last decade, ultrasound (78, 94), DBT (132, 133) and MRI (77) have been tested as additional or substitute mammography screening tools for women with dense

breasts. Moreover, different screening intervals have been proposed for women with high versus low mammographic density (59, 129-131, 134). All these approaches may contribute to improved performance measures, resulting in increased sensitivity and specificity.

An example of using different screening strategies could be found in the United States, where breast density legislation (135) has been enacted in over 25 states. The legislation movement aims to inform women about their mammographic density and the consequences of having dense breasts. Women with dense breasts might thus be able to decide if they need to attend screening more often or if they would like to have an additional breast assessment using ultrasound or MRI. However, the possibilities to do so and the costs are related to the women's health plan and/or insurance. Additional screening or supplementary screening tools are not included in the insurance coverage in most states, and a woman's choice of more frequent assessments or further examinations may be based on her income (135). This implies that participating women can afford breast cancer screening service and have made an informed decision about it. Population based screening programs aim to achieve a full participation of the target population regardless of socioeconomic status, which is associated with high costs for the service providers (56). Population based programs offer the women modern, cost-efficient screening for breast cancer. Stratifying breast cancer screening by risk factors, including mammographic density, might be one possibility to increase the effectiveness of screening programs.

This thesis is devoted to a multifaceted investigation of mammographic density in a population based breast cancer screening program with respect to selected parameters that need to be understood before the Norwegian Breast Cancer Screening Program could proceed in the research towards stratified breast cancer screening.

2. Aims of the thesis

The overall aim of the thesis is to provide knowledge about mammographic density and determine its role in breast cancer screening in Norway with respect to its relevance for potential stratified screening associated with a more effective screening program.

Mammographic density has been shown to be related to decreased effectiveness of mammographic screening (92, 105, 106). However, the association of mammographic density with performance measures, including PPVs for recall examinations or invasive procedures, as well as histopathologic prognostic tumor characteristics, such as tumor size, grade and lymph node involvement, have never been studied among women screened in the Norwegian Breast Cancer Screening Program. Furthermore, the presentation of mammographic density on the mammogram is hypothesized to be affected by breast compression parameters (124). This issue has not been investigated in depth, either internationally or with data from Norway. Moreover, in Norway, mammographic density has been classified using two different subjective classification methods and one fully automated method of density assessment in quality assurance/research mode (104). The subjective mammographic density classifications have never been studied in comparison to the fully automated method and, therefore, subjective density assessment has never been validated. We intended to fill these gaps of knowledge in this thesis and addressed the overall aim in four studies.

These studies have the following objectives:

Study I: To investigate positive predictive value for recall examination (PPV-1) and invasive procedure (PPV-2) by mammographic density and screening mode, including SFM and FFDM, in the Norwegian Breast Cancer Screening Program.

Study II: To investigate the association between mammographic density and histopathologic tumor characteristics among women screened in the Norwegian Breast Cancer Screening Program.

Study III: To explore possible associations between breast compression parameters and mammographic density assessed by an automated software among women screened in the Norwegian Breast Cancer Screening Program.

Study IV: To evaluate the three-point classification of mammographic density and the BI-RADS density classification scale used in the Norwegian Breast Cancer Screening Program with respect to estimates from an automated method of mammographic density assessment.

3. Methods

3.1 Study samples

Information solely from women who attended the Norwegian Breast Cancer Screening Program was used in all four studies. Study I and II were based on data (n = 69,442 recall examinations) obtained from women recalled due to abnormal mammographic findings (n = 62,303 women) from 1996 to 2010 (Figure 6).

Study I was limited to subsequently screened women, or women who had attended the program more than once (n = 39,427 recall examinations; n = 36,130 women). As one of the study aims was to compare performance of SFM and FFDM, we excluded examinations performed in the transition period from SFM to FFDM. As a result, recall examinations performed as part of Oslo I and Oslo II studies (n = 1,038) and during the transition period between SFM and FFDM (n = 5,315) were excluded (76). The main aim of the study was to examine PPV-1 and PPV-2 stratified by mammographic density. Therefore, any recall examinations, which did not contain information about mammographic density, were also excluded (n = 4,248). This left us with data on 28,826 recall examinations from 26,951 women for analyses.

In Study II, we used data pertaining to the first breast cancer cases (n = 10,037) among the women recalled due to abnormal mammographic findings from 1996 to 2010 (Figure 6). As we used screening mode (SFM and FFDM) for adjustment, we excluded cancers detected as part of Oslo I and Oslo II studies (n = 13). Cases without information about mammographic density were excluded (n = 898). This left us with 9,126 cases of breast cancer for analyses; both invasive cancer and DCIS were included in the study population.

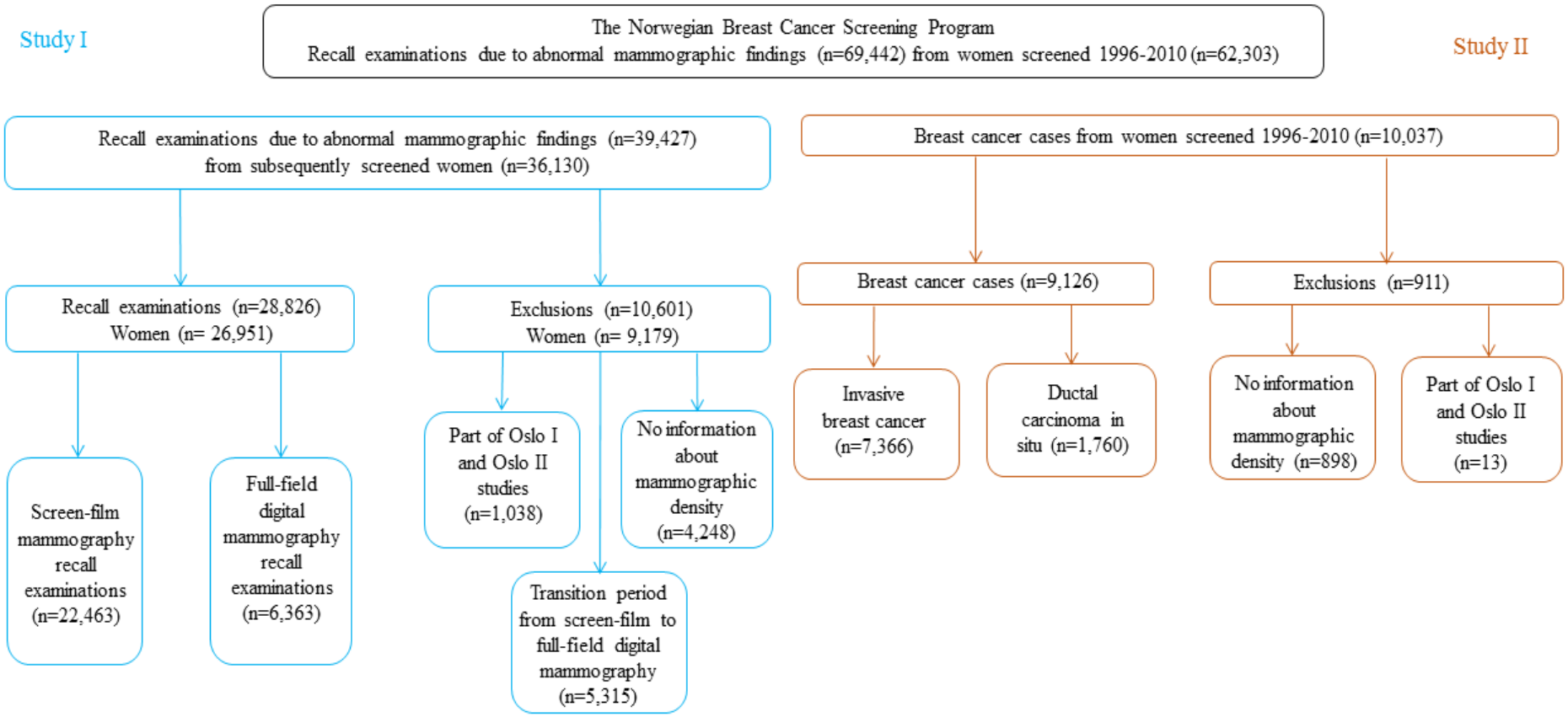

Figure 6. Flowchart describing sample selection for Study I and II

Studies III and IV were based on information obtained from women screened with FFDM in counties of Rogaland and Hordaland during the period 2007-2015 and Akershus during the period 2014-2015 (Figure 7). In study III, we used information about 17,867 screening examinations from women who attended screening units in Rogaland, Hordaland and Akershus in the period 2014-2015. As we aimed to examine the association between breast compression parameters and mammographic density assessed by a fully automated method for each mammographic projection separately, we included data from all four mammographic projections, left and right craniocaudal (CC) and mediolateral oblique (MLO) images per examination, and excluded examinations, which did not consist of four images (n = 1,485). We used BMI in this study and therefore excluded all examinations missing data on height and/or weight, which were used to calculate BMI (n = 3,484). The final dataset for the study consisted of 12,898 screening examinations from the same number of women.

In Study IV, we included information on mammographic density assessed using the three-point scale, BI-RADS and the automated method for density measurement (104), from women screened in Rogaland and Hordaland, 2007-2015 (Figure 7). The automated mammographic density estimates were retrospectively available for all women (n = 110,241 screening examinations); however, information about subjective mammographic density classifications (the three-point scale and BI-RADS) was available solely for women, who had been recalled following a screening examination. The data were divided in two sets for analyses as mammographic density assessment using the three-point scale was available for the period 2007-2012 (n = 2,310 recall examinations), while data on BI-RADS mammographic density classifications (4th edition) were available for 2013-2015 (n = 1,325 recall examinations) (Figure 6). We obtained information on volumetric breast density from the automated software for the period 2007-2015 (n = 3,625 examinations).

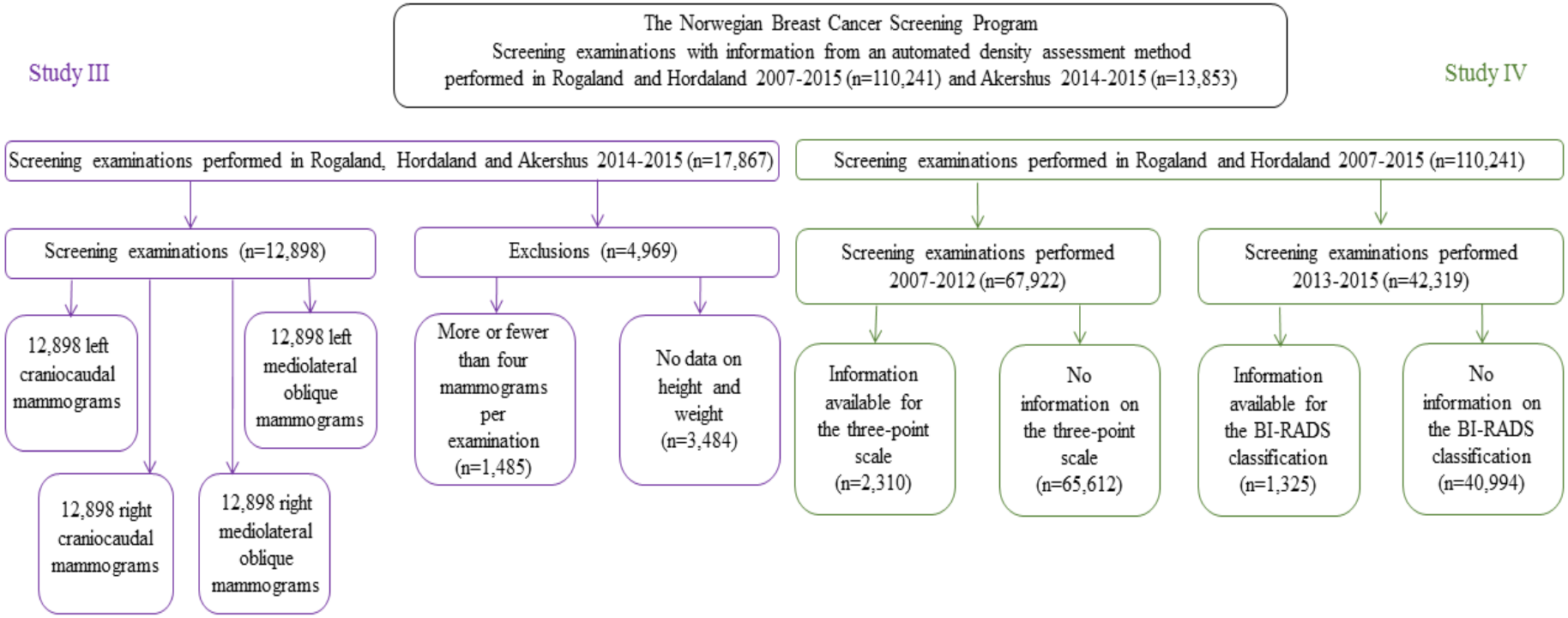

Figure 7. Flowchart describing sample selection for Study III and IV.

3.2 Data collection

Data on mammographic density were obtained from the Cancer Registry databases. The values of the three-point scale and BI-RADS mammographic density assessments were available per each recall examination (per woman). The values derived from the fully automated software were available for each mammographic image within a screening examination and overall per screening examination (average of the values of right and left CC and MLO images).

Information on histopathologic characteristics of breast cancer, including tumor size, histologic type (DCIS, invasive ductal carcinoma, invasive lobular carcinoma and other invasive cancers), grade (I, II and III) and lymph node status (positive versus negative), was available for the majority of breast cancer cases from the Cancer Registry databases. Histologic grade was assigned using the Nottingham system (136, 137).

Measurements of fibroglandular volume, breast volume, volumetric breast density, compression pressure were estimated by the fully automated method of density assessment, whereas data on compression force and compressed breast thickness were retrieved from the Digital Imaging and Communications in Medicine (DICOM) header.

Data on breast cancer risk factors, such as anthropometric parameters, were available from the questionnaire; all women invited to attend the Norwegian Breast Cancer Screening Program received this questionnaire together with an invitation to screening between 2006 and 2015 (see Appendix I). About 70% of women, who attended screening units for mammographic examination, returned a completed questionnaire at their appointments. We used information about self-reported height and weight stated at time of screening to calculate BMI (kg/m^2).

3.3 Mammographic density assessment in Norway

During 1996-2012, Norwegian breast radiologists subjectively classified mammographic density among recalled women using a three-point scale, which includes the following categories: I - fatty (<30% visible fibroglandular tissue on the mammogram), II - medium dense (30-70%) and III - dense breasts (>70%) (Figure 3).

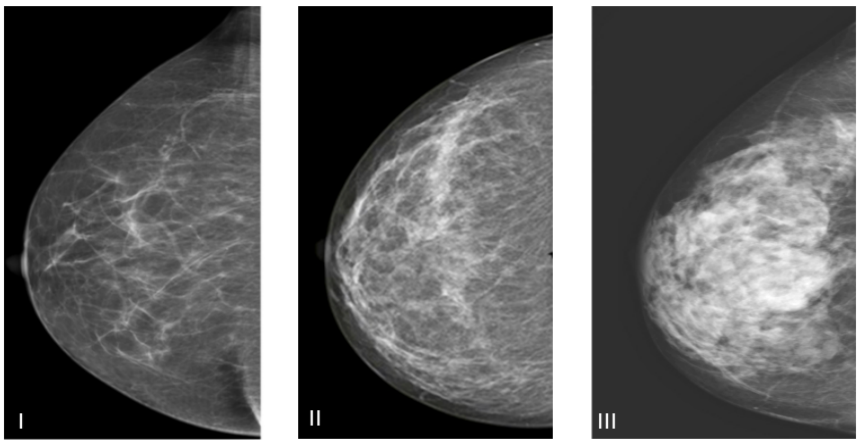

Figure 3. Mammograms assessed as fatty (I), medium dense (II) and dense breast (III) using the three-point scale of the Norwegian Breast Cancer Screening Program (source: Nataliia Moshina. Mammographic density and performance measures in the Norwegian Breast Cancer Screening Program 1996-2010. Virrat Winter Symposium 2016, Virrat, Finland, 29-31 January 2016)

In 2015, the three-point scale was replaced by the 4th edition of the BI-RADS density classification, which includes four categories based on the percent amount of visible fibroglandular tissue: BI-RADS 1 (<25% fibroglandular tissue), BI-RADS 2 (25-50%), BI-RADS 3 (50-75%) and BI-RADS 4 (>75%) (Figure 4) (93).

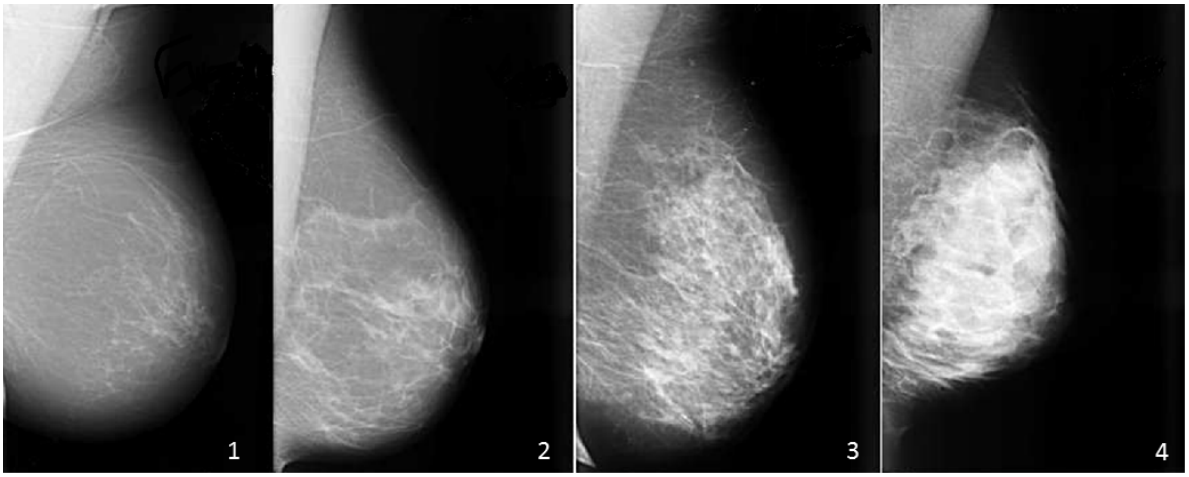

Figure 4. Mammographic density assessed by the BI-RADS density classification, 4th edition, as BI-RADS 1 (<25% fibroglandular tissue), BI-RADS 2 (25-50%), BI-RADS 3 (50-75%) and BI-RADS 4 (>75%) (source: Dave Tahmoush. Image Similarity to Improve the Classification of Breast Cancer Images. Algorithms 2009;2(4): 1503-1525) (138).

The 5th edition of the BI-RADS density classification scale has gradually been implemented in the program from 2016 onward. The 5th edition of the BI-RADS density classification includes four categories; a) almost entirely fatty, b) scattered areas of fibroglandular density, c) heterogeneously dense, which can obscure small masses, and d) extremely dense, which

lowers the sensitivity of mammography (93). The 5th edition is focused on possible masking effect of mammographic density and patterns corresponding to b and even a category can be categorized as c if an area of dense tissue, which can obscure small masses, is present on the mammogram (93).

Further, as a part of quality assurance and improvement activities within the screening program, Volpara (version 1.5.0) (Figure 5), a fully automated software for mammographic density assessment (104), was installed at four of 30 screening units in 2015. The software automatically detects a point of entirely fatty tissue in the breast and selects this as a reference level (139). Further, the software uses information on the compressed breast thickness over each pixel in the image (99, 100, 139). The reference level value and pixel-wise compressed breast thickness are compared with the intensity of each pixel in the image to determine the amount of fibroglandular tissue in that pixel. The pixel-wise proportions of fatty and dense tissue are used to create a density map showing the volume of dense tissue, or fibroglandular volume, in the breast (99). After adding up all the pixel values in the density map, the software extracts the total amount of fibroglandular tissue (cm^3). The software calculates the total volume of the breast (cm^3) by multiplying the area of the breast by the recorded breast thickness. The ratio between these two volumes determines the volumetric breast density (%), the percentage of dense volume of the total volume of the breast. Based on the volumetric breast density, a Volpara Density Grade (VDG) designed to be similar to the BI-RADS density classification, is provided by the software (104). The categories of VDG correspond to the following ranges of volumetric breast density in Volpara (version 1.5.0); VDG 1: <4.5%; VDG 2: 4.5-7.49%; VDG 3: 7.5-15.49%; and VDG 4: ≥15.5% (104).

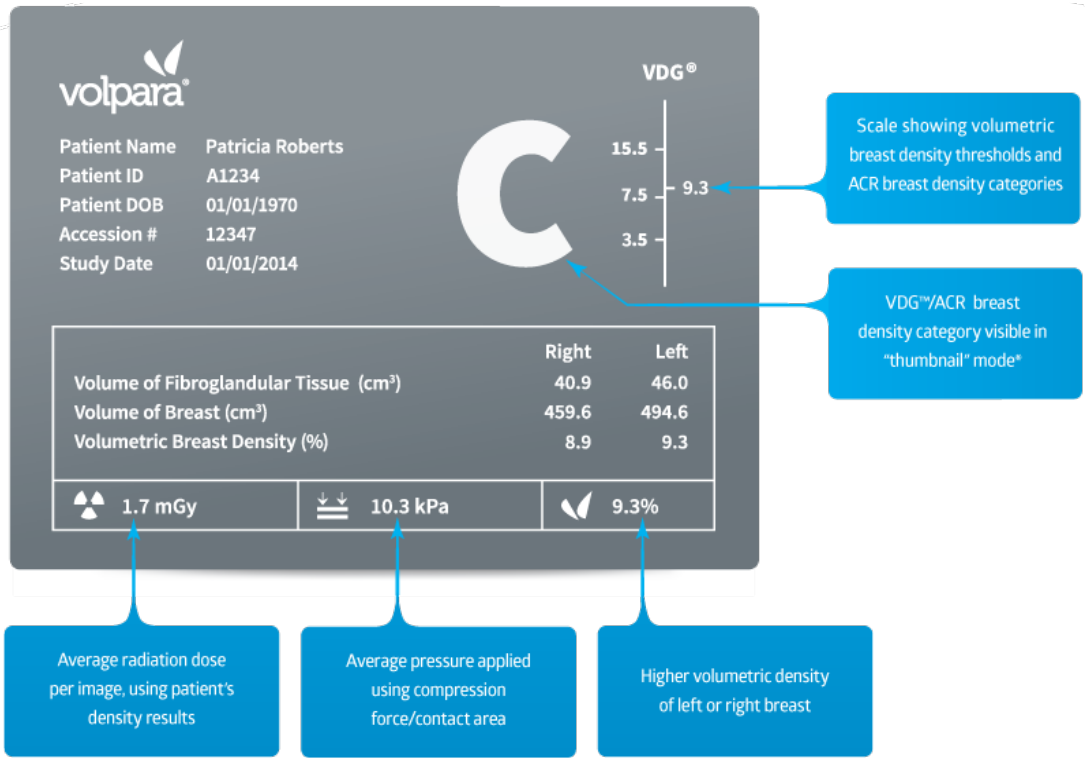

Figure 5. Automated sample patient report from Volpara™, a fully automated breast density assessment method (source: http://volparasolutions.com/our-products/volparadensity/).

3.4 Statistical analyses

In Study I, we calculated PPVs and their inverses for subsequently screened women, who were recalled following their screening examination. We estimated PPV-1 (%) as the number of screen-detected breast cancers (DCIS or invasive breast cancer) divided by the total number of recall examinations due to abnormal mammographic findings. PPV-2 (%) was estimated as the number of screen-detected breast cancers divided by the number of recall examinations including an invasive procedure (fine-needle aspiration cytology or core needle biopsy). Inverse PPVs were used to estimate the number of women needed to be recalled and the number of women needed to undergo an invasive procedure to detect one breast cancer (1/PPV-1 and 1/PPV-2, respectively). Results for these four outcomes (PPV-1, PPV-2, 1/PPV-1 and 1/PPV-2) with associated 95% confidence intervals (95% CIs) were presented, stratified by age (50-54, 55-59, 60-64, and 65-69 years), mammographic density category and screening mode (SFM and FFDM). We used a two-sample test of proportions to identify any differences in PPV-1 and PPV-2 by mammographic density and screening mode. A test for trend was used to determine whether an increase or decrease in PPV-1 and PPV-2 was observed across categories of mammographic density by screening mode and

age. Lastly, we estimated the odds of detecting breast cancer at screening among recalled women for varying levels of mammographic density, adjusting for age and screening mode, and presented the resulting odds ratios (ORs) and 95% CIs.

In Study II, we used a chi-square test to compare the distribution of histologic type and to compare tumor size (≤15 mm versus >15 mm), histologic grade (I versus II and III), and lymph node status (positive versus negative) of invasive cancers by mammographic density, as classified by the three-point scale. All tests were two sided with a 5% significance level. We then used logistic regression to estimate the odds of the aforementioned (binary) histopathologic tumor characteristics of invasive cancers associated a dichotomous measure of mammographic density (fatty versus medium dense and dense). These models were adjusted for age (50-54, 55-59, 60-64, and 65-69 years) and screening mode (SFM and FFDM).

In Study III, we measured the correlation between compression force, compression pressure, compressed breast thickness, breast volume, fibroglandular volume and volumetric breast density, stratified by mammographic view (CC and MLO), using the Spearman correlation coefficient (ρ). Additionally, scatterplots with locally weighted smoothing were used to display associations between these parameters. We then used linear regression to study the association between breast compression parameters (compression force, pressure, compressed breast thickness and breast volume) and natural log transformed fibroglandular volume and volumetric breast density, adjusting for age (continuous) and BMI (continuous). All variables included in the regression models were standardized so that the estimated regression coefficients represented the change in standard deviations (SDs) of the natural log transformed outcome variables (fibroglandular volume and volumetric breast density) associated with one SD change in breast compression parameters.

In Study IV, we presented a descriptive analysis of the distribution of mammographic density classified using the three-point scale, BI-RADS and VDG, stratified by age (50-54, 55-59, 60-64, and >64 years). We also compared the distributions of mammographic density assessed using the three-point scale and BI-RADS with VDG. A quadratically weighted kappa (k_w) was used to identify the agreement between BI-RADS and VDG. Agreement between the density measures was assessed using the scale: slight: 0.00–0.20; fair: 0.21–0.40; moderate: 0.41–0.60; substantial: 0.61–0.80; and almost perfect: 0.81–1.00 (140). Further, we compared the mean values of fibroglandular volume, breast volume and volumetric breast density for

the categories of mammographic density assessed by the three-point scale and the BI-RADS density classification, using Bonferroni adjustments for multiple comparisons (three comparisons for the three-point scale and six comparisons for BI-RADS). We graphically presented the categories of the subjective classifications of mammographic density (the three-point scale and BI-RADS) in relation to the estimates of volumetric breast density.

All analyses were performed with Stata (versions 13 and 14, *Stata*Corp, College Station, TX, USA).

3.5 Ethical considerations

We used solely de-identified data for all four studies. Only data from women, who have not explicitly notified that they refuse the Cancer Registry to use data about their screening examinations for quality assurance and research, were used for analyses. The Regional Committees for Medical and Health Research Ethics (REC) approved all four studies (reference number 2014/1526 for Study I and II, and 2016/938 for Study III and IV).

4. Results of the studies

4.1 Study I

The overall PPV-1 and PPV-2 decreased with increasing mammographic density for both SFM and FFDM (p for trend <0.05) (Table 3). PPV-1 was statistically significantly higher for FFDM compared with SFM for women with fatty breasts. PPV-2 was statistically significantly higher for FFDM compared with SFM for women with fatty and medium dense breasts. PPV-1 and PPV-2 increased by age regardless of mammographic density or screening mode (p for trend <0.05 for all).

When data from both screening modes were combined, the number of women needed to be recalled or undergo an invasive procedure to detect one breast cancer was statistically significantly lower for women with fatty (4.9 or 2.0, respectively) compared with medium dense (5.8 or 2.1, respectively) and dense breasts (6.6 or 2.2, respectively).

Table 3. Positive predictive values (PPV-1 and PPV-2) stratified by mammographic density (fatty, medium dense and dense) and five-year age groups among subsequently screened women in the Norwegian Breast Cancer Screening Program, 1996-2010

Age groups (years)	PPV-1 (%, 95% confidence interval)		
	Fatty (n=7,548)	Medium dense (n=18,219)	Dense (n=3,059)
50-54	12.5 (10.5-14.7)	9.9 (9.0-10.8)	10.7 (8.9-12.7)
55-59	18.0 (16.4-19.8)	15.9 (15.0-16.9)	15.1 (13.0-17.4)
60-64	21.7 (20.0-23.4)	21.2 (20.1-22.4)	19.7 (16.8-23.1)
65-69	25.7 (23.9-27.7)	24.3 (22.9-25.9)	21.4 (17.3-25.9)
Overall	20.5 (19.6-21.5)	17.4 (16.8-17.9)[a]	15.3 (14.0-16.6)[a]
	PPV-2 (%, 95% confidence interval)		
	Fatty (n=3,079)	Medium dense (n=6,682)	Dense (n=1,037)
50-54	40.1 (34.7-45.6)	32.1 (29.6-34.7)	36.1 (30.8-41.8)
55-59	46.4 (43.0-49.9)	44.6 (42.5-46.8)	44.9 (39.6-50.3)
60-64	51.4 (48.2-54.5)	54.0 (51.7-56.3)	50.4 (43.9-56.9)
65-69	55.0 (52.8-59.2)	57.2 (54.5-59.8)	55.6 (48.1-64.0)
Overall	50.3 (48.6-52.1)	47.4 (46.2-48.6)[b]	45.0 (42.0-48.1)[c]

[a] P value <0.001 for comparison between fatty and medium dense, and between fatty and dense.
[b] P value 0.006 for comparison between fatty and medium dense.
[c] P value 0.003 for comparison between fatty and dense.

Among women recalled in the Norwegian Breast Cancer Screening Program, the odds of breast cancer decreased with increasing mammographic density (Table 4). Compared with women with fatty breasts, the odds of breast cancer were 10% lower for those with medium

dense and 15 % lower for those with dense breasts after adjustment for screening mode and age. Compared with recalled women aged 50-54 years, the odds of breast cancer were almost three times higher for those aged 65-69 years.

Table 4. Crude and adjusted odds ratio (OR) with 95% confidence interval (95%CI) of breast cancer among women recalled after a subsequent screening examination (n=28,826) in the Norwegian Breast Cancer Screening Program, 1996-2010

	Crude OR (95%CI)	Adjusted [a] OR (95%CI)
Mammographic density		
Fatty	Reference	reference
Medium dense	0.81 (0.76-0.87)	0.90 (0.84-0.96)
Dense	0.68 (0.62-0.78)	0.85 (0.76-0.95)
Age groups (years)		
50-54	Reference	reference
55-59	1.68 (1.52-1.85)	1.67 (1.51-1.84)
60-64	2.31 (2.10-2.56)	2.28 (2.07-2.52)
65-69	2.81 (2.54-3.15)	2.75 (2.48-3.04)
Screening mode		
Screen-film mammography	Reference	reference
Full-field digital mammography	1.12 (1.07-1.20)	1.17 (1.06-1.23)

[a] Adjusted for mammographic density, screening mode and age.

4.2 Study II

DCIS represented 15.8% and 22.0% of the cancers among women with fatty and dense breasts ($p<0.001$), respectively, while the proportions of invasive lobular carcinoma were 6.8% and 11.1%, respectively ($p<0.001$) (Table 5). The mean and median tumor size of invasive breast cancers was 13.8 mm (95%CI: 13.4-14.1) and 12 mm, respectively, for women with fatty breasts. These values were 16.2 mm (95% CI: 15.4-17.0) and 14 mm, respectively, for women with dense breasts. The percentage of tumors >15 mm was 28.1% among women with fatty breasts and 37.6% among those with dense breasts ($p<0.001$). There were no statistically significant differences in histologic grade by mammographic density (data not shown). Lymph node positive tumors were less common in women with fatty breasts (20.6%) compared with women with dense breasts (27.2%).

Table 5. Histopathologic characteristics of screen-detected breast cancers in the Norwegian Breast Cancer Screening Program, 1996-2010, stratified by mammographic density (fatty, medium dense and dense breasts)

	Total n (%)	Fatty n (%)	Medium dense n (%)	Dense n (%)	p-value[a]	p-value[b]
Histologic type	N=9,126	N=2,721	N=5,538	N=867		
Ductal carcinoma *in situ*	1,760 (19.3)	429 (15.8)	1,140 (20.6)	191 (22.0)	<0.001	<0.001
Invasive ductal carcinoma	6,176 (67.7)	1,959 (72.0)	3,660 (66.1)	557 (64.2)	<0.001	<0.001
Invasive lobular carcinoma	755 (8.3)	184 (6.8)	475 (8.6)	96 (11.1)	0.004	<0.001
Other invasive cancers	435 (4.8)	149 (5.5)	263 (4.8)	23 (2.7)	0.154	0.001
Invasive breast cancers	N=7,366	N=2,292	N=4,398	N=676		
Tumor size[c]						
Mean, mm	14.5	13.8	14.7	16.2		
Median (mm)	13	12	13	14		
≤15 mm	4,821 (65.5)	1,601 (69.9)	2,834 (64.4)	386 (57.1)	<0.001	<0.001
>15 mm	2,342 (31.8)	645 (28.1)	1,443 (32.8)	254 (37.6)	<0.001	<0.001
Lymph nodes[c]						
Positive	1,753 (23.8)	472 (20.6)	1,097 (24.9)	184 (27.2)	<0.001	<0.001

[a] Fatty versus medium dense breasts
[b] Fatty versus dense breasts
[c] Cases with missing information are not shown

Compared to women with fatty or medium dense breasts, women with dense breasts had higher odds of large tumors (OR 1.44, 95% CI: 1.18-1.73) and lymph node positive tumors (OR 1.26, 95% CI: 1.05-1.51), after adjustment for age and screening mode. Including screening mode in the model did not change the observed estimates.

4.3 Study III

Compression force, compressed breast thickness and breast volume were positively correlated with fibroglandular volume ($\rho = 0.20$, 0.27 and 0.53 for CC and $\rho = 0.14$, 0.33 and 0.45 for MLO, respectively), while compression pressure was inversely correlated with fibroglandular volume ($\rho = -0.48$ for CC and $\rho = -0.28$ for MLO). Compression force, compressed breast thickness and breast volume were inversely correlated with volumetric breast density ($\rho = -0.12$, -0.55 and -0.55 for CC and $\rho = -0.18$, -0.60 and -0.63 for MLO, respectively), while compression pressure was positively correlated with volumetric breast density ($\rho = 0.30$ for CC and $\rho = 0.33$ for MLO).

In the linear regression models, after adjustment for age and BMI, the strongest associations were observed between compression pressure and fibroglandular volume, as well as breast volume and fibroglandular volume (Table 6). Compressed breast thickness had the strongest association with volumetric breast density.

Table 6. Associations[a] between breast compression parameters (compression force, pressure, compressed breast thickness and breast volume) and fibroglandular volume or volumetric breast density among 12,898 left craniocaudal (CC) and left mediolateral oblique (MLO) mammograms

	CC (n=12,898)			MLO (n=12,898)		
	Beta [b] (95% CI)	P-value	R^2	Beta [b] (95% CI)	P-value	R^2
Fibroglandular volume						
Compression force	0.21 (0.19; 0.23)	<0.001	0.37[c]	0.13 (0.11; 0.15)	<0.001	0.25[c]
Compression pressure	-0.33 (-0.35;-0.31)	<0.001		-0.35 (-0.39;-0.30)	<0.001	
Compressed breast thickness	-0.01 (-0.02; 0.04)	0.43		0.03 (0.01; 0.06)	0.02	
Breast volume	0.35 (0.31; 0.39)	<0.001		0.31 (0.28; 0.35)	<0.001	
Volumetric breast density						
Compression force	-0.06 (-0.08;-0.04)	<0.001	0.40[c]	-0.11 (-0.13;-0.09)	<0.001	0.44[c]
Compression pressure	0.09 (0.07; 0.11)	<0.001		0.30 (0.26; 0.34)	<0.001	
Compressed breast thickness	-0.50 (-0.52;-0.48)	<0.001		-0.43 (-0.46;-0.41)	<0.001	
Breast volume	0.03 (-0.01; 0.06)	0.14		-0.02 (-0.05; 0.01)	0.22	

[a] Models were adjusted for compression force, compression pressure, compressed breast thickness, breast volume, body mass index (BMI) and age; adjusted covariates are not shown for BMI and age

[b] Beta coefficients represent the difference in fibroglandular volume or volumetric breast density represented standard deviations (SDs) on the natural log transformed scale associated with one SD change in the explanatory variable, after adjustment for other covariates

[c] R-squared for the model

4.4 Study IV

The proportion of screening examinations classified as dense breasts by the three-point scale or as VDG 4 by Volpara decreased with increasing age ($p<0.05$ for trend for both methods). However, this trend was not observed for the BI-RADS density classification, and the proportions of BI-RADS 2, 3 and 4 did not vary significantly by age for women between 50 and 64 years old (p for trend = 0.72; 0.74; and 1.00 for BI-RADS 2, 3 and 4, respectively).

According to the three-point scale, 23% of the screening examinations were classified as fatty, 67% as medium dense and 10% as dense (Table 7). Among the examinations classified as fatty by the three-point scale, 60% were classified as VDG 1 and none was classified as VDG 4. Among the examinations classified as medium dense by the three-point scale, 78% were classified as VDG 2 or 3. The proportion of examinations classified as fatty by the three-point scale was lower than VDG 1 (23% versus 26%, $p<0.05$). Furthermore, the proportion of examinations classified as dense by the three-point scale was significantly higher than VDG 4 (10% versus 7%, $p<0.001$).

Proportions of examinations classified as BI-RADS 1 and 4 were lower compared to VDG 1 and 4, respectively (16% and 4% versus 29% and 6 %, $p<0.05$ for all). The agreement between BI-RADS and VDG was moderate ($k_w = 0.50$, 95% CI: 0.47-0.53; $p<0.001$).

Table 7. Distribution of mammographic density based on the three-point scale, BI-RADS and Volpara Density Grade (VDG) among women recalled for further assessment due to abnormal mammographic findings in the Norwegian Breast Cancer Screening Program, 2007-2015

Study period 2007-2012 (n=2310)				
Three-point scale	Fatty	Medium Dense		Dense
	535 (23%)	1,538 (67%)		237 (10%)
VDG	1	2	3	4
	606 (26%)	879 (38%)	674 (29%)	151 (7%)
Study period 2013-2015 (n=1325)				
BI-RADS	1	2	3	4
	215 (16%)	743 (56%)	321 (24%)	46 (4%)
VDG	1	2	3	4
	382 (29%)	465 (35%)	402 (30%)	76 (6%)
Study period 2007-2015 (n=3635)				
VDG	1	2	3	4
	988 (27%)	1,344 (37%)	1,076 (30%)	227 (6%)

Mean fibroglandular volume and volumetric breast density for each category of the three-point scale were statistically significantly different ($p<0.001$ for all) and increased by increasing density category (p for trend <0.001). The same pattern was observed for the BI-RADS categories (p for trend <0.001). Box-plots for mammographic density assessed with the three-point scale and the BI-RADS density classification by volumetric breast density indicated that volumetric breast density increased alongside both subjective classification scores (Figure 8).

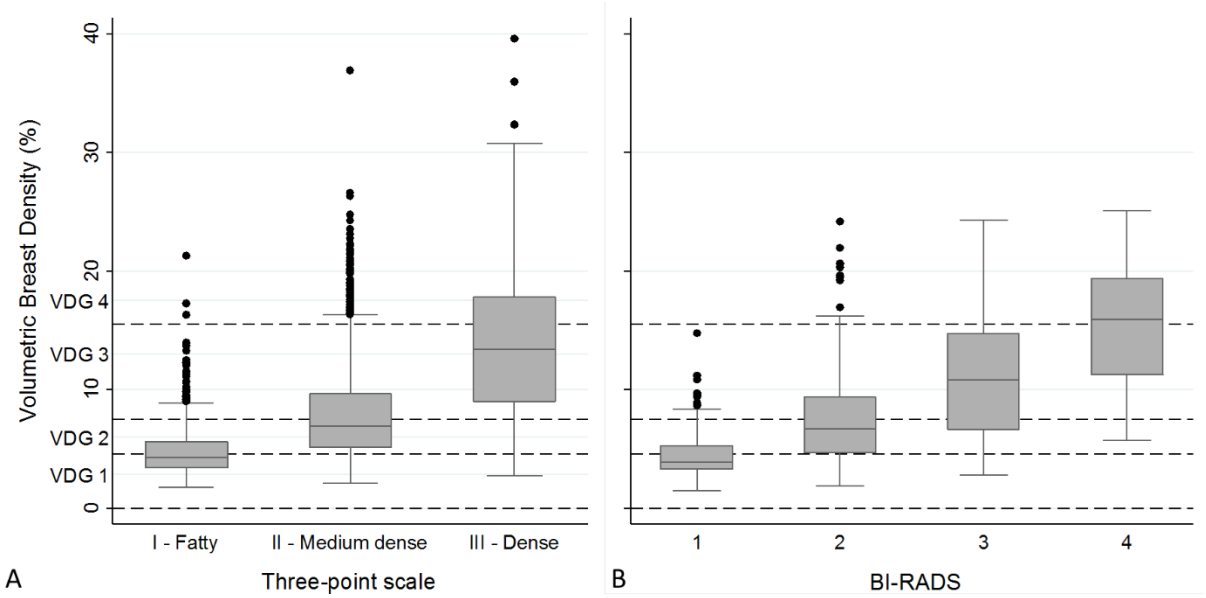

Figure 8. Box-plots of mammographic density assessed with the three-point scale and BI-RADS classifications corresponding to volumetric breast density (%).
For each category, the horizontal line shows the median. Each box contains 50% of the centrally observed data from the 25th to 75th percentile. The whiskers of the boxes contain the remaining 25% cases at each size. Extreme values (outliers) are depicted by black dots. The horizontal dashed lines indicate the ranges of the four categories of Volpara Density Grade (VDG).
Panel A: mammographic density distribution for women recalled in the Norwegian Breast Cancer Screening Program, 2007-2012 (n=2,310).
Panel B: mammographic density distribution for women recalled in the Norwegian Breast Cancer Screening Program, 2013-2015 (n=1,325).

5. Discussion

5.1 Interpretation of main findings

We found that PPVs decreased with increasing mammographic density among women recalled for further assessment in the screening program in Norway. High mammographic density was positively associated with larger tumor size (>15mm) and lymph node involvement among women with screen-detected invasive breast cancer. We showed that breast compression parameters were associated with mammographic density estimates obtained from the fully automated software. We observed that the Norwegian three-point scale and the BI-RADS density classification corresponded well with the estimates of volumetric breast density obtained from the fully automated software.

5.1.1 Study I

The probability of detecting breast cancer among recalled women and among women, who underwent an invasive procedure, decreased with increasing mammographic density, and increased with age. Our results on PPV-1 are similar to results from other studies (107, 141, 142). PPV-2 was not reported in any of the referred three studies. Lower PPV-1 among women with dense breasts might be explained by the phenomenon of superimposition (143). Superimposition is related to overlapping tissue, which could present as an area with suspicious lesion(s) formed by a summation of normal breast tissue layers. Superimposition has to be distinguished from masking by mammographic density, wherein a tumor can be hidden in mammographic dense tissue and cannot be identified by the radiologist.

Compared with SFM, mammograms acquired with FFDM were associated with higher PPVs among women with fatty breasts, who were recalled or underwent an invasive procedure in the age range 50-60 years. These findings were expected; however, we did not observe significant differences between SFM and FFDM for women with dense breasts. Previous studies have shown that FFDM is associated with higher PPV and sensitivity among women with dense breasts (144, 145).

5.1.2 Study II

We found that high mammographic density was positively associated with large tumor size (>15 mm) and lymph node positive tumors. Our results support the findings from other

studies, where various methods were used to classify mammographic density (108, 109, 112, 113, 146). Several studies have reported a larger tumor size in dense compared to fatty breasts (108, 109, 112, 113, 146). We did not identify any association between mammographic density and histologic grade of screen-detected tumors. Similar results have been reported in a study from Sweden (108). However, other studies have demonstrated either positive (146) or negative (113, 147) associations between mammographic density and histologic grade. A positive association between mammographic density and grade may reflect a biological relationship between a high amount of breast glandular tissue and a low degree of tumor differentiation (or high histologic grade) (146). On the other hand, a negative association between mammographic density and histopathologic grade could be related to the tissue microenvironment in fatty breasts, which might be more conductive to high-grade tumors (108).

5.1.3 Study III

In this study, we identified associations between breast compression parameters and fibroglandular volume and volumetric breast density. The strongest associations were observed between compression pressure and fibroglandular volume, breast volume and fibroglandular volume, as well as between compressed breast thickness and volumetric breast density.

The associations between breast compression parameters and fibroglandular volume and volumetric breast density may be partially explained by the impact of breast volume. Women with large breasts have higher fibroglandular volume and considerably higher volume of fatty tissue and thus lower volumetric breast density, compared with women with small breasts (148). Given that women with large breasts tend to receive higher compression force compared to women with small breasts (149, 150), higher compression force is thus related to higher fibroglandular volume and lower volumetric breast density.

Compression pressure is estimated as the compression force divided by the contact area of the breast during compression. Compression pressure was inversely correlated with fibroglandular volume and positively correlated with volumetric breast density. Women with large breasts have lower pressure compared with women with small breasts and the lower pressure is thus associated with high fibroglandular volumes. Volumetric breast density is lower in large versus small breasts, which might indicate a relationship between low pressure

and low volumetric breast density. In a study investigating the association between breast stiffness (the ratio of compression force and deformation of the breast due to compression) and the risk of breast cancer, it was shown that dense volume, percent dense volume and breast tissue stiffness were positively associated with the risk of breast cancer (152). This might indicate that low compressibility of breast tissue, corresponding to a high level of compression pressure, is associated with high volumetric breast density.

Our results on the association of breast compression parameters and fibroglandular volume and volumetric breast density are in line with results from previous studies (124, 125, 152-155).

5.1.4 Study IV

The mean values of an automated volumetric breast density corresponding to the three-point scale and BI-RADS increased with increasing the density category. Moderate agreement was observed between VDG and BI-RADS density scores.

No studies have previously examined the three-point scale in relation to estimates of volumetric breast density obtained by the fully automated software. However, several studies have compared fully automated methods for breast density assessment to BI-RADS (85, 96, 102, 103, 105, 156-162), indicating a clear increase in volumetric breast density as the BI-RADS density category increased (103) or a strong relationship between volumetric breast density and the BI-RADS categories (96). Similar trends have been reported in other studies (102, 157-159, 162). Furthermore, k_w for the association between VDG and the BI-RADS classification has been shown to vary from 0.26 (102) to 0.80 (103). Our findings regarding BI-RADS are thus in line with the results of the prior studies.

The lack of trends and differences between the proportions of examinations classified as BI-RADS 2, 3 and 4 by age for women 50-64 years old is likely related to the difficult transition from the three-point scale to BI-RADS mammographic density classifications.

The proportions of examinations classified as fatty by the three-point scale and VDG 1, as well as those classified as dense and VDG 4, differed significantly. In addition, 17% of the exams classified as dense by the three-point scale and 13% of those classified as BI-RADS 4 were shown to have VDG 1 or 2. This discrepancy might be related to the use of different images for the subjective versus automated mammographic density assessments. Post-

processed ("for presentation") images were used for the subjective assessment, whereas raw ("for processing") images were used for the automated measurement. The differences between the post-processed and raw images have been shown to affect the estimation of percent breast density (163-165). Furthermore, the disparity between the subjective and automated density assessment might be noticeable because radiologists tend to assign density using the highest value based on the available mammographic projections, whereas volumetric density of a woman is estimated as an average value based on the values of two mammographic projections (93, 103). This regression toward the mean density value will result in lower volumetric estimates of density when compared to radiologists. In addition, mammograms with mammographic density that is close to thresholds between various VDG categories may be classified into the upper or the lower to the corresponding BI-RADS category, due to inter- or intra-reader variability (162).

In Norway, two qualitative classifications for mammographic density (the three-point scale and BI-RADS) were used in non-overlapping periods solely for recalled women. It is thus not possible to compare the three-point scale with the BI-RADS density classification. The distribution of volumetric breast density estimates for the categories of the subjective classifications for the recalled population might help identify upper and lower thresholds for each mammographic density category and reproduce hypothetical categorical distribution of mammographic density for the screened population. The distribution of mammographic density is assumed to be different for the recalled and screened population because women with dense breasts are more often recalled than women with fatty breasts (77, 78, 95).

5.2 Methodological considerations

5.2.1 Data quality at the Cancer Registry of Norway

Data on breast cancer used for the studies in this project were obtained from the Cancer Registry of Norway databases. Cancer reporting is mandatory by law and the Cancer Registry has registered cancer cases since 1952 (166). The cancer incidence registry contains the basic data collected from clinicians, pathologists, administrative patient discharge records and mortality sources. The incidence registry is updated continuously with information on both new cases and cases diagnosed during previous years (4). The Cancer Registry databases are 99% complete for breast cancer (167). The databases also contain data on screening invitations, attendance, recall, radiologists' assessment and

pathology reports with cytology and/or histology for all findings from screening recalls, including benign and malignant findings and interval cancers. We also used women's responses to a questionnaire (administered 2006-2016) on sociodemographic and anthropometric parameters, as well as breast cancer risk factors (Appendix I).

5.2.2 Selection bias

Selection bias is a systematic error, which occurs when the association between the exposure and the outcome is different for those who were included and those who were not included in the study (168).

In Study I, we excluded women who participated in the Oslo I (169) and Oslo II (170) studies, women screened in the transition period between SFM and FFDM, and women with no information about mammographic density. The question was whether the PPVs across categories of mammographic density differed between those women who were included and those who were excluded. Oslo I and Oslo II lasted for only four years, and the number of examinations excluded was small (n=1038, Figure 6) considering the entire number of screening examinations (n=28,826). Therefore, the exclusion of the women participated in Oslo I and Oslo II studies was unlikely to bias substantially the outcome for PPVs across categories of mammographic density.

The women screened in the transition period between SFM and FFDM had a higher detection rate of screen-detected breast cancer compared with those screened with either SFM or FFDM (171). However, we have performed sensitivity analyses and determined that mammographic density distribution, trends in PPVs and the association between mammographic density and breast cancer among women excluded due to being screened in the transition period were similar to the women included in the study. Therefore, the exclusion of the women screened during the transition period could not bias the outcome of the study. It is also unlikely that exclusion of women with no information about mammographic density could lead to differences in PPVs across categories of mammographic density.

In Study II, we excluded women diagnosed with breast cancer if they participated in the Oslo I and Oslo II studies or did not have information about mammographic density. The number of women excluded due to participation in Oslo I and Oslo II was 13 (Figure 6). We

assume that it is unlikely that the distribution of histopathologic tumor characteristics (tumor size, lymph node status and histologic grade) across mammographic density categories or association between mammographic density and histopathologic tumor characteristics differed between the included and excluded women.

In Study III, we excluded women with fewer or more than four mammograms or without data on height and/or weight. Women with very large breasts often require more than four mammograms in order to image the entire breast. However, we see no obvious reason why the association between breast compression parameters and density estimates (volumetric breast density and fibroglandular volume) should differ for excluded and included women, as there are no studies on whether the relationship between compression parameters and density estimates is different for women with large versus small breasts.

The bias associated with inclusion of women who participated versus not participated in the screening program is described in part 5.2.5. External validity and generalizability.

5.2.3 Information bias

Information bias occurs when there is a misclassification of the outcome or/and exposure variable in the study sample (168, 172). Information bias includes differential misclassification, which occurs when the probability of misclassification is systematically different for one of the groups being compared, and non-differential misclassification, which occurs if the extent of misclassification is random or not systematic across comparisons group (168).

In Study I, differential misclassification related to the identification of the outcome (breast cancer) may have occurred if some breast cancer cases, specifically those of a small size, had been missed among women with dense breasts. Similarly, misclassification of one of the outcome variables (tumor size) could have been possible in Study II, if a substantial amount of small tumors had been missed among women with dense breasts, as the association between a large tumor size and high mammographic density might not have been found. However, women with dense breasts having any suspicious findings on their mammograms are more often recalled for further assessment and undergo additional examinations including MRI and/or an invasive procedure. The probability to miss a (small) cancer is thus very low. The studies on the interval cancer in the Norwegian Breast Cancer Screening Program showed

that the rate of interval breast cancer, or breast cancer diagnosed among screened women between a negative screening examination and the next scheduled screening examination, for the period 1996-2004 was 19 per 10,000 screened women (173). Therefore, although we cannot exclude the possibility that some small tumors were overlooked in dense breasts, we think the effect is likely to be minimal, and unlikely to greatly have affected either the outcome of PPVs by mammographic density in Study I or the association between mammographic density and histopathologic tumor characteristics in Study II.

Non-differential misclassification of the exposure variable in Studies I and II may have occurred, as mammographic density might have been misclassified by radiologists. However, although the radiologists could have an inclination as to the tumor size, s/he would not know all the histopathological characteristics. Therefore, although this could have been partially differential, we think this is more likely to have resulted in non-differential error. Due to known differences in inter- and intra-reader reliability, it is obvious that at least some non-differential misclassification would have occurred on the subjectively assessed density (174). Such misclassification would most likely have resulted in a bias toward the null, which could have weakened the associations between exposure and outcome variables in Study I and II.

We used density estimates from the fully automated density assessment software in Study III and IV. We assumed the estimates produced by the automated density measurement software to be a reference standard, because the software has shown very good correlation with measurements obtained by MRI (159, 175). However, this measure is prone to measurement errors in compressed breast thickness, which will result in over-or underestimates of volumetric breast density (154). This is likely to be non-differential, and we would therefore have tended to underestimate the effects, if anything.

The Norwegian radiologists started assessing mammographic density with BI-RADS in 2013, after a long history of using the three-point scale and the BI-RADS scores might have been inconsistent and appeared to be a subject of non-differential misclassification. This misclassification could have weakened the agreement between BI-RADS and VDG. A previous study from Norway using the dichotomized BI-RADS density classification (1 and 2 versus 3 and 4) showed that the agreement (measured by kappa with quadratic weights) between reader's median classification and an individual score of five radiologists ranged between 0.76 and 0.93 (176). Thus, even though strong, this agreement is not perfect, and we

may therefore have underestimated the agreement between BI-RADS and VDG somewhat in Study IV.

5.2.4 Confounding

Confounding occurs when the exposure-outcome association is influenced by some other factors, which affect the outcome and differ between the exposure groups (168). Confounding can influence the internal validity of a study. The lack of adjustment for possible confounders could be a concern for studies I, II and III.

In Study I and II we were not able to adjust for the independent breast cancer risk factors, including family history, use of postmenopausal hormonal therapy or BMI. Family history of breast cancer (at least one first-degree relative affected) and use of postmenopausal hormonal therapy (especially combined estrogen-progesterone) are associated with higher risk of breast cancer and high mammographic density (177, 178).

In study II, variables that affect both mammographic density and histopathologic criteria could represent confounders. Use of postmenopausal estrogen-progesterone therapy has been reported to be associated with both favorable and less favorable histopathologic tumor characteristics (179), and it is therefore not clear how adjustment of this would have affected our results. High BMI is associated with low percent mammographic density (23, 116) and less favorable histopathologic tumor characteristics, including tumor size and histologic grade (180, 181). Therefore, adjustment for BMI would if anything have strengthened the association between mammographic density and tumor characteristics. It should be mentioned that in both Study I and Study II there may be residual confounding due to unknown confounding factors.

In Study III, to calculate fibroglandular volume, breast volume and volumetric breast density, the fully automated software for density assessment uses values of compressed breast thickness (99, 154). A study describing the development of a phantom to test the software showed that application of compression force and reducing the phantom thickness caused both reduction and increase in phantom's volumetric breast density estimates (155). The influence of the estimates of compressed breast thickness on the values of fibroglandular volume and volumetric breast density could have confounded the associations between the rest of breast compression parameters and fibroglandular volume and volumetric breast

density. The impact of breast thickness might have been overestimated, while the impact of compression force, compression pressure and breast volume underestimated in the analyses.

In Study III, to calculate BMI (one of the adjustment variables) we used information on self-reported weight and height. One previous study including Norwegian women showed that despite a small but statistically significant underreporting of weight (0.6 kg, $p<0.05$), there was substantial agreement between self-reported and measured BMI values (182). If this level of reliability were present in our study sample, we may have had some residual confounding due to misclassification of BMI, but we do not assume this a major issue.

In Study III, the associations between breast compression parameters and the outcome variables could have been influenced by the differences in mammography systems. Two different types of mammography equipment were used in Rogaland and Hordaland (GE Senographe Essential) and Akershus (Philips MicroDose SI). In addition, we did not include information about women's and radiographers' preferences with respect to compression force in the regression analyses. Studies on women's preferences and/or experience of pain and discomfort at mammography have never been performed in Norway. The compression force used has been shown to vary between breast centers, mammography systems and radiographers (151), which suggests that confounding may be present in the form of the radiographer's preferences and site-specific compression force. It is possible that the radiographers' preference to use a compression force above the average for women with small breasts and high stiffness, who will presumably have high mammographic density (152). The recommendations of the quality assurance manual regarding application of compression force ranging from 108 to 177 N have been shown to be followed in approximately 60% of acquired mammograms (151, 183). Because a large proportion of mammograms are acquired using compression force outside of the recommended limits, it is possible that the observed effects in Study III are not as strong as they may have been, had we been able to control for different factors associated with the acquisition of the mammogram (e.g. radiographer and mammography system).

5.2.5 External validity and generalizability

External validity indicates the consistency of study sample characteristics (184). If characteristics of a selected study sample differ consistently from the characteristics of those

who were not selected, a sample may not be representative of the study population (172, 185, 186).

Women included in our study samples were selected from the population of women attending the Norwegian Breast Cancer Screening Program. These women may not be representative with respect to early performance measures, risk of breast cancer and mammographic density of the female population of screening age (50-69 years). Women of low socioeconomic status and/or foreign background tend not to attend screening programs (187). We do not know the extent of differences in breast cancer risk and mammographic density among these women in Norway. A recent study reported that breast cancer incidence is higher in higher versus lower educated women in Norway (188), but the association between education and participation in screening in Norway is not known.

The design of the Study I, II and IV was built upon the knowledge that information about subjectively assessed mammographic density was available and, therefore, collected only for women recalled in the Norwegian Breast Cancer Screening Program. Currently, the program has no information on the distribution of mammographic density according to the three-point scale or BI-RADS among those not recalled. Based on a sensitivity analysis of the data available from two counties, we assume that women with medium dense and dense breasts are recalled more frequently than women with fatty breasts. Therefore, the results of our descriptive analyses in Study I, II and IV may not be generalizable to either the screened population or the Norwegian female population aged 50-69 years.

In Study I, our sample consisted solely of women who attended screening more than once (subsequently screened), which might have led to exclusion of a group of women aged 50-52 years i.e. the exclusion of pre- and perimenopausal women with high mammographic density (114). Moreover, if we had included prevalently screened women, we could also have had more recalled women in the age group 50-54 years, because women participating in the first screening round have a significantly higher recall rate (189). However, other performance measures of breast cancer screening such as PPVs and rates of screen-detected cancer are different for prevalently and subsequently screened women and should not be mixed if one aims to accurately estimate screening performance (56). The performance measures for prevalently screened women are largely dependent on the extent of opportunistic screening and regional incidence rates of breast cancer, while the performance measures for subsequently screened women demonstrate the impact and effectiveness of the screening

program (56). The results of descriptive analyses in Study I are therefore not generalizable to prevalently screened women.

5.3 Clinical implications

Our findings from Study I indicate that mammography is a better screening tool for women with fatty breasts compared with those with dense breasts. The latter group is more likely subjected to harm of both false negative and false positive mammograms, including anxiety and distress (111, 190). Therefore, thoughtful planning of informing women about their mammographic density and consideration of additional screening tools for women with dense breasts might have both psychological and diagnostic benefits for breast cancer screening for this group of women.

Our findings from Study II suggest that masking by mammographic density is associated with larger tumor size in the Norwegian Breast Cancer Screening Program. It has previously been shown that masking is related to missed tumors, resulting in missed cancers of large size with lymph node invasion (112, 113). Stratifying breast cancer screening by mammographic density, either by reducing the screening interval (for example, introducing a one year interval instead of two years) or by offering additional screening tools to women with dense breasts, might increase the sensitivity of mammographic screening and therefore improve breast cancer prognosis for this subgroup of women (77, 92).

Our findings from Study III indicate that compression pressure, compressed breast thickness and breast volume might influence mammographic density estimates obtained from the automated software. This suggests that mammographic appearance of breast tissue is modifiable, but that the assessment of mammographic density is primarily dependent on characteristics of the breast such as breast volume, amount of fatty tissue, or stiffness.

The results of Study IV in combination with findings from Study I and II (191, 192) might indicate that the scores of mammographic density assigned by the radiologists could be used to identify women who most likely need adjunctive (other than mammography) screening. According to the three-point scale, these women have been classified with dense breasts.

5.4 Relevance of stratified breast cancer screening based on mammographic density

The results of the performed studies show that mammographic density affects breast cancer screening performance. Further, the subjective density classifications used in Norway have a potential to become a basis for screening stratified by mammographic density. Stratification by offering women ultrasound or MRI in addition to standard mammography is considered beneficial for women with dense breasts (77, 78, 193, 194) and might be relevant in Norway (191, 192). Use of additional screening tools has been reported to increase the sensitivity of the screening test in women with high mammographic density (77, 78, 92). However, a recent review from the United States has shown that harms of supplemental screening with ultrasound or MRI for women with dense breasts include higher recall and biopsy rates when compared with FFDM alone (195).

No gold standard for defining or classifying mammographic density has been stated (195). Subjective assessment of mammographic density among screened women is time-consuming and, therefore requires more breast radiologists to perform reading of the mammograms, which may result in increased costs related to stratified breast cancer screening. The intra- and inter-reader reliability of the subjective density assessment is known to be a substantial limitation in using mammographic density for stratification of breast cancer screening (196-198).

Most modern discussions on breast cancer screening consider mammographic density assessed by an automated software an important parameter for breast cancer risk prediction and potential screening stratification (126, 156, 175, 199-201). However, installation and use of the automated breast density assessment are associated with high costs. Furthermore, changes in density over time have been insufficiently investigated in association with possible influential factors (hormonal therapy, changes in BMI, diet, genetic modification, etc.) (31). The factors affecting mammographic density over time showed inconsistent impact on it, possibly due to the fact that both these factors and mammographic density might be modified (202). Despite the decreased mammographic sensitivity in women with dense breasts, the outcome of the performance measures of the organized breast cancer screening programs in Europe (203, 204) and in Norway is encouraging (86, 171, 205). Therefore, we consider that the current evidence is not sufficient to support the use of mammographic density for stratified breast cancer screening in Norway.

6. Conclusions and future perspectives

6.1 Conclusions

The main results of the studies included in this thesis are:

1) PPVs decreased with increasing mammographic density regardless of screening mode (SFM or FFDM) in the Norwegian Breast Cancer Screening Program. More women with dense breasts compared with fatty breasts were recalled or underwent an invasive procedure to detect one breast cancer. Our findings indicate that mammography is a better screening tool for women with fatty compared with dense breasts. The latter group is thus more likely subjected to harm of both false negative and false positive mammogram.

2) Among women with screen-detected cancers, high mammographic density was positively associated with larger tumor size (>15mm) and positive lymph node status. Masking by mammographic density might result in less favorable prognosis for the women and a decreased sensitivity of mammography.

3) Of the breast compression parameters, compression pressure and breast volume had the strongest association with fibroglandular volume, while compressed breast thickness had the strongest association with volumetric mammographic density.

4) The Norwegian three-point scale and the BI-RADS density classification showed reasonable correspondence with the estimates of volumetric breast density obtained by the automated method for density assessment.

In this thesis, we have shown that mammographic density plays an important role in breast cancer screening performance in Norway. However, considering currently available knowledge, the use of mammographic density for stratified breast cancer screening in Norway is not relevant.

6.2 Future perspectives

Several aspects associated with mammographic density and possibilities of breast cancer screening stratified by mammographic density in Norway need to be investigated in future studies. Use of randomization is preferable. The first step is to gain information about mammographic density assessment of all women screened. There is also a need to determine the risk of breast cancer associated with mammographic density in the screened and non-screened population, including information about breast cancer risk factors.

Radiation dose and parameters of image acquisition (x-ray current and tube voltage, filtration and anode material) may influence mammographic density assessment (124, 125) and should, therefore, be explored in association with mammographic density estimates.

Another important aspect is longitudinal changes in mammographic density observed before breast cancer diagnosis. To investigate this issue, a mammographic density assessment should be performed among women attended several subsequent screening examinations.

In order to consider the use mammographic density for stratified breast cancer screening, the benefits of ultrasound, including ABUS, and/or MRI as additional screening tools available at screening examination for women with dense breasts should outweigh possible harms/limitations of such approach.

Recent studies have shown superior radiologists' performance for women with dense breasts when screening is performed by DBT (206). In the Norwegian context, DBT should be examined as an alternative to mammography screening tool for women with dense breasts.

As far as we are aware, no studies have been published on important longer-term clinical outcomes of supplemental screening or use of different screening intervals for women with high versus low mammographic density. Therefore, research on breast cancer mortality and overdiagnosis associated with stratified breast cancer screening based on mammographic density should be performed to create a basis for potential recommendations on this issue in the future.

References

1. DeSantis CE, Bray F, Ferlay J, Lortet-Tieulent J, Anderson BO, Jemal A. International Variation in Female Breast Cancer Incidence and Mortality Rates. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2015;24(10):1495-506.
2. Breast Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012 2016 [updated 28.11.2016; cited 2016 November 28]. Available from: http://globocan.iarc.fr/old/FactSheets/cancers/breast-new.asp.
3. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer. 2015;136(5):E359-86.
4. Cancer in Norway 2015 - Cancer incidence, mortality, survival and prevalence in Norway. Cancer Registry of Norway, 2016.
5. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. Nature. 1983;303(5920):767-70.
6. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. International journal of cancer. 1997;71(5):800-9.
7. Howell A, Anderson AS, Clarke RB, Duffy SW, Evans DG, Garcia-Closas M, et al. Risk determination and prevention of breast cancer. Breast cancer research : BCR. 2014;16(5):446.
8. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. The Lancet Oncology. 2012;13(11):1141-51.
9. Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. Breast cancer research : BCR. 2006;8(4):R43.
10. Petrucelli N DM, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. . Seattle (WA): University of Washington, Seattle1998 [cited 2017 January 23]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1247/.
11. Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. Breast cancer research and treatment. 2015;149(3):569-75.
12. Byrne C, Schairer C, Wolfe J, Parekh N, Salane M, Brinton LA, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. Journal of the National Cancer Institute. 1995;87(21):1622-9.
13. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 1998;7(12):1133-44.
14. Anisimov VN. Biology of aging and cancer. Cancer control : journal of the Moffitt Cancer Center. 2007;14(1):23-31.
15. Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. Journal of the National Cancer Institute. 1995;87(9):670-5.
16. Henderson IC. Risk factors for breast cancer development. Cancer. 1993;71(6 Suppl):2127-40.
17. Colditz GA, Willett WC, Hunter DJ, Stampfer MJ, Manson JE, Hennekens CH, et al. Family history, age, and risk of breast cancer. Prospective data from the Nurses' Health Study. Jama. 1993;270(3):338-43.
18. Rice MS, Bertrand KA, Lajous M, Tamimi RM, Torres G, Lopez-Ridaura R, et al. Reproductive and lifestyle risk factors and mammographic density in Mexican women. Annals of epidemiology. 2015;25(11):868-73.

19. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006;15(6):1159-69.
20. Ewertz M, Duffy SW, Adami HO, Kvale G, Lund E, Meirik O, et al. Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries. International journal of cancer. 1990;46(4):597-603.
21. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. Lancet (London, England). 2002;360(9328):187-95.
22. Collins JA, Blake JM, Crosignani PG. Breast cancer risk with postmenopausal hormonal treatment. Human reproduction update. 2005;11(6):545-60.
23. Brand JS, Czene K, Eriksson L, Trinh T, Bhoo-Pathy N, Hall P, et al. Influence of lifestyle factors on mammographic density in postmenopausal women. PloS one. 2013;8(12):e81876.
24. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. American journal of epidemiology. 2000;152(6):514-27.
25. Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath CW, Jr., 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. British Journal of Cancer. 2002;87(11):1234-45.
26. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. The Lancet Oncology. 2001;2(3):133-40.
27. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. Breast cancer research : BCR. 2005;7(1):21-32.
28. Morabia A, Zhang FF. History of medical screening: from concepts to action. Postgraduate Medical Journal. 2004;80(946):463-9.
29. Hiatt RA, Klatsky AL, Armstrong MA. Alcohol consumption and the risk of breast cancer in a prepaid health plan. Cancer research. 1988;48(8):2284-7.
30. Roman M, Sakshaug S, Graff-Iversen S, Vangen S, Weiderpass E, Ursin G, et al. Postmenopausal hormone therapy and the risk of breast cancer in Norway. International journal of cancer Journal international du cancer. 2016;138(3):584-93.
31. Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Kasamesup V, Thakkinstian A. Risk prediction models of breast cancer: a systematic review of model performances. Breast cancer research and treatment. 2012;133(1):1-10.
32. Li J, Szekely L, Eriksson L, Heddson B, Sundbom A, Czene K, et al. High-throughput mammographic-density measurement: a tool for risk prediction of breast cancer. Breast cancer research : BCR. 2012;14(4):R114.
33. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. Boletin de la Oficina Sanitaria Panamericana Pan American Sanitary Bureau. 1968;65(4):281-393.
34. Weedon-Fekjaer H, Vatten LJ, Aalen OO, Lindqvist B, Tretli S. Estimating mean sojourn time and screening test sensitivity in breast cancer mammography screening: new results. Journal of medical screening. 2005;12(4):172-8.
35. Shen Y, Zelen M. Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2001;19(15):3490-9.
36. UICC. Global Cancer Control 2016 [cited 2016 November 29]. Available from: http://www.uicc.org/resources/tnm/about.
37. NCCN Clinical Practice Guidelines in Oncology: Breast Cancer. 2015 [cited 2016 November 22]. Available from: http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf.
38. Kalager M, Tamimi RM, Bretthauer M, Adami HO. Prognosis in women with interval breast cancer: population based observational cohort study. BMJ : British Medical Journal. 2012;345:e7536.

39. Richards MA, Smith P, Ramirez AJ, Fentiman IS, Rubens RD. The influence on survival of delay in the presentation and treatment of symptomatic breast cancer. British Journal of Cancer. 1999;79(5-6):858-64.
40. Taghipour S, Banjevic D, Miller AB, Montgomery N, Jardine AK, Harvey BJ. Parameter estimates for invasive breast cancer progression in the Canadian National Breast Screening Study. British Journal of Cancer. 2013;108(3):542-8.
41. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. Statistics in medicine. 1995;14(14):1531-43.
42. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;380(9855):1778-86.
43. Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, et al. Breast-Cancer Screening — Viewpoint of the IARC Working Group. New England Journal of Medicine. 2015;372(24):2353-8.
44. Recommendations on Breast Cancer Screening 2016 [updated 24/11/2016; cited 2016 December 27]. Available from: http://ecibc.jrc.ec.europa.eu/recommendations/details/4.
45. National Cancer Institute. Breast Cancer Screening (PDQ®)–Health Professional Version 2017. Available from: https://www.cancer.gov/types/breast/hp/breast-screening-pdq#section/_40.
46. European Society of Breast Imaging 2016 [cited 2016 Decemebr 26]. Available from: http://www.eusobi.org/cms/website.php?id=/en/society/committees.htm.
47. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2008;19(4):614-22.
48. IARC Handbook of Cancer Prevention Volume 7. Breast Cancer Screening. : IARC Press; 2002. Available from: http://www.iarc.fr/en/publications/pdfs-online/prev/handbook7/.
49. Andermann A, Blancquaert I, Beauchamp S, Déry V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bulletin of the World Health Organization. 2008;86(4):317-9.
50. Styringsstruktur og strategi for nasjonale screeningprogrammer. Høringsutkast. Helsedirektoratet. (June 25, 2014, 2014). Available from: http://legeforeningen.no/PageFiles/182477/H%C3%B8ringsutkast%20fra%20Helsedirektoratet%20-%20Styringsstruktur%20og%20strategi%20for%20nasjonale%20screeningprogrammer.pdf.
51. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. 4 ed. Luxemburg: European Communities; 2006.
52. Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. Breast cancer research : BCR. 2015;17(1).
53. Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Annals of internal medicine. 2012;156(7):491-9.
54. Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, et al. Mammographic screening programmes in Europe: organization, coverage and participation. Journal of medical screening. 2012;19 Suppl 1:72-82.
55. Paci E, Broeders M, Hofvind S, Duffy SW. The benefits and harms of breast cancer screening. Lancet (London, England). 2013;381(9869):800-1.
56. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. Annals of Oncology. 2008;19(4):614-22.
57. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. British journal of cancer. 2013;108(11):2205-40.

58. Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, et al. Breast-cancer screening--viewpoint of the IARC Working Group. The New England journal of medicine. 2015;372(24):2353-8.
59. Mittmann N, Stout NK, Lee P, Tosteson AN, Trentham-Dietz A, Alagoz O, et al. Total cost-effectiveness of mammography screening strategies. Health reports. 2015;26(12):16-25.
60. Broeders M, Paci E. The balance sheet of benefits and harms of breast cancer population-based screening in Europe: outcome research, practice and future challenges. Women's health (London, England). 2015;11(6):883-90.
61. Hofvind S, Ursin G, Tretli S, Sebuodegard S, Moller B. Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. Cancer. 2013;119(17):3106-12.
62. Johns LE, Coleman DA, Swerdlow AJ, Moss SM. Effect of population breast screening on breast cancer mortality up to 2005 in England and Wales: an individual-level cohort study. British Journal of Cancer. 2016.
63. Weedon-Fekjær H, Romundstad PR, Vatten LJ. Modern mammography screening and breast cancer mortality: population study. BMJ : British Medical Journal. 2014;348.
64. Norum J. Breast cancer screening by mammography in Norway. Is it cost-effective? Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 1999;10(2):197-203.
65. Moger TA, Kristiansen IS. Direct and indirect costs of the Norwegian Breast Cancer Screening Program. 2012;3. Available from: https://www.med.uio.no/helsam/forskning/nettverk/hero/publikasjoner/skriftserie/2012/hero2012-3.pdf.
66. Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. Journal of epidemiology and community health. 1997;51(6):705-10.
67. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Annals of internal medicine. 2007;146(7):502-10.
68. Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S, et al. Systematic review of the psychological consequences of false-positive screening mammograms. Health technology assessment (Winchester, England). 2013;17(13):1-170, v-vi.
69. Heleno B, Siersma VD, Brodersen J. Diagnostic invasiveness and psychosocial consequences of false-positive mammography. Annals of family medicine. 2015;13(3):242-9.
70. Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. Annals of family medicine. 2013;11(2):106-15.
71. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. Journal of medical screening. 2012;19 Suppl 1:42-56.
72. Beckmann KR, Lynch JW, Hiller JE, Farshid G, Houssami N, Duffy SW, et al. A novel case-control design to estimate the extent of over-diagnosis of breast cancer due to organised population-based mammography screening. International journal of cancer. 2015;136(6):1411-21.
73. Heywang-Köbrunner SH, Hacker A, Sedlacek S. Advantages and Disadvantages of Mammography Screening. Breast Care. 2011;6(3):199-207.
74. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of Breast Cancer Screening: Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. Annals of internal medicine. 2016;164(4):256-67.
75. Whelehan P, Evans A, Wells M, Macgillivray S. The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. Breast (Edinburgh, Scotland). 2013;22(4):389-94.
76. Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. Acta radiologica (Stockholm, Sweden : 1987). 2009;50(1):3-14.

77. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. Jama. 2012;307(13):1394-404.
78. Kelly KM, Dean J, Comulada WS, Lee SJ. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. European radiology. 2010;20(3):734-42.
79. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. New England Journal of Medicine. 2005;353(17):1784-92.
80. Yaffe MJ. Basic Physics of Digital Mammography. In: Bick U, Diekmann F, editors. Digital Mammography. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010. p. 1-11.
81. Kotre CJ, dos Reis CS. Mammography Equipment. In: Hogg P, Kelly J, Mercer C, editors. Digital Mammography: A Holistic Approach: Springer; 2015. p.125-41.
82. Kopans DB. Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk. Radiology. 2008;246(2):348-53.
83. Faulconer LS, Parham CA, Connor DM, Kuzmiak C, Koomen M, Lee Y, et al. Effect of Breast Compression on Lesion Characteristic Visibility with Diffraction-Enhanced Imaging. Academic Radiology. 2010;17(4):433-40.
84. Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography: how much is enough? Australasian radiology. 2003;47(2):121-6.
85. Kreftregisteret. Resultater fra Mammografiprogrammet 2006-2014. Available from: http://kreftregisteret.no/no/Generelt/Publikasjoner/Mammografiprogrammet/Mammografiprogrammet-Resultater-fra-prosessindikatorer-2006-201314/.
86. Hofvind S, Geller B, Vacek PM, Thoresen S, Skaane P. Using the European guidelines to evaluate the Norwegian Breast Cancer Screening Program. European journal of epidemiology. 2007;22(7):447-55.
87. Research-based evaluation of the Norwegian Breast Cancer Screening Program. Final report. 2015. Available from: https://www.regjeringen.no/contentassets/444d08daf15e48aca5321f2cefaac511/mammografirapport-til-web.pdf.
88. Falk R, Hofvind S, Skaane P, Haldorsen T. Overdiagnosis among women attending a population-based mammography screening program. International journal of cancer. 2013;133(3):705-12.
89. Ellis H. Anatomy of the breast. Surgery (Oxford). 2004;22(7):145-7.
90. Kopans DB. Mammography and the Normal Breast. Breast imaging. 3 ed: Lippincott Williams & Wilkins; 2006. p. 357-63.
91. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer. 1976;37(5):2486-92.
92. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. Journal of the National Cancer Institute. 2000;92(13):1081-7.
93. Sickles E, D'Orsi CJ, Bassett LW, et al. ACR BI-RADS® Mammography. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reston, VA: American College of Radiology; 2013.
94. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR American journal of roentgenology. 2000;174(6):1769-77.
95. Kerlikowske K, Grady D, Barclay J, Frankel SD, Ominsky SH, Sickles EA, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. Journal of the National Cancer Institute. 1998;90(23):1801-9.
96. Jeffreys M HJ, Highnam R. Comparing a New Volumetric Breast Density Method (Volpara) to Cumulus. 2010. In: Lecture Notes in Computer Science: 10th International Workshop on Digital Mammography [Internet]. Girona, Spain: Springer-Verlag; [408–13].

97. Sovio U, Li J, Aitken Z, Humphreys K, Czene K, Moss S, et al. Comparison of fully and semi-automated area-based methods for measuring mammographic density and predicting breast cancer risk. British Journal of Cancer. 2014;110(7):1908-16.
98. Ursin G, Astrahan MA, Salane M, Parisky YR, Pearce JG, Daniels JR, et al. The detection of changes in mammographic densities. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 1998;7(1):43-7.
99. Highnam R BM, Yaffe MJ, Karssemeijer N, Harvey J. Robust Breast Composition Measurement - Volpara™ In: Martí J OA, Freixenet J, Martí R., editor. Lecture Notes in Computer Science: 10th International Workshop on Digital Mammography. Berlin Heidelberg: Springer-Verlag 2010. p. 342–9.
100. Hartman K, Highnam R, Warren R, Jackson V. Volumetric Assessment of Breast Tissue Composition from FFDM Images. In: Krupinski EA, editor. Digital Mammography: 9th International Workshop, IWDM 2008 Tucson, AZ, USA, July 20-23, 2008 Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 33-9.
101. He W, Juette A, Denton ERE, Oliver A, Marti R, Zwiggelaar R. A Review on Automatic Mammographic Density and Parenchymal Segmentation. International journal of breast cancer. 2015;2015: 276217.
102. Ko SY, Kim EK, Kim MJ, Moon HJ. Mammographic density estimation with automated volumetric breast density measurement. Korean journal of radiology. 2014;15(3):313-21.
103. van der Waal D, den Heeten GJ, Pijnappel RM, Schuur KH, Timmers JM, Verbeek AL, et al. Comparing Visually Assessed BI-RADS Breast Density and Automated Volumetric Breast Density Software: A Cross-Sectional Study in a Breast Cancer Screening Setting. PloS one. 2015;10(9):e0136667.
104. Matakina Technology Limited. VolparaSolutions – Volpara Density [cited 2015 October 1]. Available from: http://volparasolutions.com/solutions/volparadensity/.
105. Pisano ED, Gatsonis CA, Yaffe MJ, Hendrick RE, Tosteson AN, Fryback DG, et al. American College of Radiology Imaging Network digital mammographic imaging screening trial: objectives and methodology. Radiology. 2005;236(2):404-12.
106. Wanders JO, Holland K, Veldhuis WB, Mann RM, Pijnappel RM, Peeters PH, et al. Volumetric breast density affects performance of digital screening mammography. Breast cancer research and treatment. 2017;162(1):95-103.
107. van Gils CH, Otten JD, Verbeek AL, Hendriks JH, Holland R. Effect of mammographic breast density on breast cancer screening performance: a study in Nijmegen, The Netherlands. Journal of epidemiology and community health. 1998;52(4):267-71.
108. Sartor H, Borgquist S, Hartman L, Zackrisson S. Do pathological parameters differ with regard to breast density and mode of detection in breast cancer? The Malmo Diet and Cancer Study. Breast (Edinburgh, Scotland). 2015;24(1):12-7.
109. Bertrand KA, Tamimi RM, Scott CG, Jensen MR, Pankratz V, Visscher D, et al. Mammographic density and risk of breast cancer by age and tumor characteristics. Breast cancer research : BCR. 2013;15(6):R104.
110. Roman M, Skaane P, Hofvind S. The cumulative risk of false-positive screening results across screening centres in the Norwegian Breast Cancer Screening Program. European journal of radiology. 2014;83(9):1639-44.
111. Elmore JG, Jackson SL, Abraham L, Miglioretti DL, Carney PA, Geller BM, et al. Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy. Radiology. 2009;253(3):641-51.
112. Nickson C, Kavanagh AM. Tumour size at detection according to different measures of mammographic breast density. Journal of medical screening. 2009;16(3):140-6.
113. Aiello EJ, Buist DS, White E, Porter PL. Association between mammographic breast density and breast cancer tumor characteristics. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2005;14(3):662-8.

114. Boyd N, Martin L, Stone J, Little L, Minkin S, Yaffe M. A longitudinal study of the effects of menopause on mammographic features. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2002;11(10 Pt 1):1048-53.
115. Maskarinec G, Pagano I, Lurie G, Kolonel LN. A longitudinal investigation of mammographic density: the multiethnic cohort. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006;15(4):732-9.
116. Baglietto L, Krishnan K, Stone J, Apicella C, Southey MC, English DR, et al. Associations of mammographic dense and nondense areas and body mass index with risk of breast cancer. American journal of epidemiology. 2014;179(4):475-83.
117. Cuzick J, Warwick J, Pinney E, Duffy SW, Cawthorn S, Howell A, et al. Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: a nested case-control study. Journal of the National Cancer Institute. 2011;103(9):744-52.
118. Brisson J, Brisson B, Cote G, Maunsell E, Berube S, Robert J. Tamoxifen and mammographic breast densities. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2000;9(9):911-5.
119. Yang WT, Lewis MT, Hess K, Wong H, Tsimelzon A, Karadag N, et al. Decreased TGFbeta signaling and increased COX2 expression in high risk women with increased mammographic breast density. Breast cancer research and treatment. 2010;119(2):305-14.
120. Reeves KW, Weissfeld JL, Modugno F, Diergaarde B. Circulating levels of inflammatory markers and mammographic density among postmenopausal women. Breast cancer research and treatment. 2011;127(2):555-63.
121. dos Santos Silva I, Johnson N, De Stavola B, Torres-Mejia G, Fletcher O, Allen DS, et al. The insulin-like growth factor system and mammographic features in premenopausal and postmenopausal women. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006;15(3):449-55.
122. Ellingjord-Dale M, Lee E, Couto E, Ozhand A, Qureshi S, Hofvind S, et al. Polymorphisms in hormone metabolism and growth factor genes and mammographic density in Norwegian postmenopausal hormone therapy users and non-users. Breast cancer research : BCR. 2012;14(5):R135.
123. Ellingjord-Dale M, Grotmol T, Lee E, Van Den Berg DJ, Hofvind S, Couto E, et al. Breast cancer susceptibility variants and mammographic density phenotypes in norwegian postmenopausal women. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2014;23(9):1752-63.
124. Olson JE, Sellers TA, Scott CG, Schueler BA, Brandt KR, Serie DJ, et al. The influence of mammogram acquisition on the mammographic density and breast cancer association in the Mayo Mammography Health Study cohort. Breast cancer research : BCR. 2012;14(6):R147.
125. Khan-Perez J, Harkness E, Mercer C, Bydder M, Sergeant J, Morris J, et al. Volumetric Breast Density and Radiographic Parameters. In: Fujita H, Hara T, Muramatsu C, editors. Breast Imaging: 12th International Workshop, IWDM 2014, Gifu City, Japan, June 29 – July 2, 2014 Proceedings. Cham: Springer International Publishing; 2014. p. 265-72.
126. Brand JS, Czene K, Shepherd JA, Leifland K, Heddson B, Sundbom A, et al. Automated measurement of volumetric mammographic density: a tool for widespread breast cancer risk assessment. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2014;23(9):1764-72.
127. Evans DGR, Howell A. Breast cancer risk-assessment models. Breast cancer research : BCR. 2007;9(5):213.

128. Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. Journal of the National Cancer Institute. 2010;102(10):680-91.
129. Desreux J, Bleret V, Lifrange E. Should we individualize breast cancer screening? Maturitas. 2012;73(3):202-5.
130. Kerlikowske K, Zhu W, Hubbard RA, Geller B, Dittus K, Braithwaite D, et al. Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. JAMA internal medicine. 2013;173(9):807-16.
131. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Annals of internal medicine. 2011;155(1):10-20.
132. Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH, et al. Digital tomosynthesis in breast imaging. Radiology. 1997;205(2):399-406.
133. Lang K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmo Breast Tomosynthesis Screening Trial, a population-based study. European radiology. 2016;26(1):184-90.
134. Berg CD. Breast Cancer Screening Interval: Risk Level May Matter. Annals of internal medicine. 2016;165(10):737-8.
135. Slanetz PJ, Freer PE, Birdwell RL. Breast-Density Legislation — Practical Considerations. New England Journal of Medicine. 2015;372(7):593-5.
136. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat. 1992;22(3):207-19.
137. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19(5):403-10.
138. Tahmoush D. Image Similarity to Improve the Classification of Breast Cancer Images. Algorithms. 2009;2(4):1503.
139. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE transactions on medical imaging. 2006;25(3):273-82.
140. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
141. Lehmann U. Lobular breast cancer--the most common special subtype or a most special common subtype? Breast Cancer Res. 2015;17(1):99.
142. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006;98(17):1204-14.
143. Sickles EA. Findings at mammographic screening on only one standard projection: outcomes analysis. Radiology. 1998;208(2):471-5.
144. Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD, et al. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. Annals of internal medicine. 2011;155(8):493-502.
145. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. The New England journal of medicine. 2005;353(17):1773-83.
146. Harrison DA, Duffy SW, Sala E, Warren RM, Couto E, Day NE. Deterministic models for breast cancer progression: application to the association between mammographic parenchymal pattern and histologic grade of breast cancers. Journal of clinical epidemiology. 2002;55(11):1113-8.
147. Masarwah A, Auvinen P, Sudah M, Rautiainen S, Sutela A, Pelkonen O, et al. Very low mammographic breast density predicts poorer outcome in patients with invasive breast cancer. European radiology. 2015;25(7):1875-82.

148. Stuedal A, Ma H, Bernstein L, Pike MC, Ursin G. Does breast size modify the association between mammographic density and breast cancer risk? Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2008;17(3):621-7.
149. Mercer CE, Hogg P, Szczepura K, Denton ERE. Practitioner compression force variation in mammography: A 6-year study. Radiography. 2013;19(3):200-6.
150. Mercer CE, Hogg P, Lawson R, Diffey J, Denton ER. Practitioner compression force variability in mammography: a preliminary study. The British journal of radiology. 2013;86(1022):20110596.
151. Waade GG, Moshina N, Saebuodegard S, Hogg P, Hofvind S. Compression forces used in the Norwegian Breast Cancer Screening Program. The British journal of radiology. 2017:20160770.
152. Boyd NF, Li Q, Melnichouk O, Huszti E, Martin LJ, Gunasekara A, et al. Evidence that breast tissue stiffness is associated with risk of breast cancer. PloS one. 2014;9(7):e100937.
153. Khan-Perez J, Mercer C, Bydder M, Sergeant J, Morris J, Maxwell A, et al. PB.10: Breast compression, compressed breast thickness and volumetric breast density. Breast Cancer Research. 2013;15(1):P10.
154. Waade GG, Highnam R, Hauge IHR, McEntee MF, Hofvind S, Denton E, et al. Impact of errors in recorded compressed breast thickness measurements on volumetric density classification using volpara v1.5.0 software. Medical physics. 2016;43(6):2870-6.
155. Waade GG, Hofvind S, Thompson JD, Highnam R, Hogg P. Development of a phantom to test fully automated breast density software - A work in progress. Radiography. 2013;23(1):e14-e19.
156. Eng A, Gallant Z, Shepherd J, McCormack V, Li J, Dowsett M, et al. Digital mammographic density and breast cancer risk: a case-control study of six alternative density assessment methods. Breast cancer research : BCR. 2014;16(5):439.
157. Gweon HM, Youk JH, Kim JA, Son EJ. Radiologist assessment of breast density by BI-RADS categories versus fully automated volumetric assessment. AJR American journal of roentgenology. 2013;201(3):692-7.
158. Seo JM, Ko ES, Han BK, Ko EY, Shin JH, Hahn SY. Automated volumetric breast density estimation: a comparison with visual assessment. Clinical radiology. 2013;68(7):690-5.
159. Gubern-Merida A, Kallenberg M, Platel B, Mann RM, Marti R, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms: a validation study. PloS one. 2014;9(1):e85952.
160. Tagliafico A, Tagliafico G, Tosto S, Chiesa F, Martinoli C, Derchi LE, et al. Mammographic density estimation: comparison among BI-RADS categories, a semi-automated software and a fully automated one. Breast (Edinburgh, Scotland). 2009;18(1):35-40.
161. Osteras BH, Martinsen AC, Brandal SH, Chaudhry KN, Eben E, Haakenaasen U, et al. Classification of fatty and dense breast parenchyma: comparison of automatic volumetric density measurement and radiologists' classification and their inter-observer variation. Acta radiologica (Stockholm, Sweden : 1987). 2016.
162. Sartor H, Lang K, Rosso A, Borgquist S, Zackrisson S, Timberg P. Measuring mammographic density: comparing a fully automated volumetric assessment versus European radiologists' qualitative classification. European radiology. 2016.
163. Li D GS, Conant E, et al. Comparison of breast percent density estimation from raw versus processed digital mammograms. 2011. In: SPIE [Internet]. Lake Buena Vista, Florida; [79631e6].
164. Gastounioti A, Oustimov A, Keller BM, Pantalone L, Hsieh MK, Conant EF, et al. Breast parenchymal patterns in processed versus raw digital mammograms: A large population study toward assessing differences in quantitative measures across image representations. Medical physics. 2016;43(11):5862.
165. Vachon CM, Fowler EE, Tiffenberg G, Scott CG, Pankratz VS, Sellers TA, et al. Comparison of percent density from raw and processed full-field digital mammography data. Breast cancer research : BCR. 2013;15(1):R1.

166. Lov om helseregistre og behandling av helseopplysninger (helseregisterloven). Oslo: Helse- og omsorgsdepartementet. 2001. Available from: https://lovdata.no/dokument/NL/lov/2014-06-20-43?q=Lov%20om%20helseregistre%20og%20behandling.
167. Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. European journal of cancer (Oxford, England : 1990). 2009;45(7):1218-31.
168. Rothman KJ GS, Lash TL. Modern Epidemiology. Third Edition ed. 530 Walnut Street Philadelphia, PA 19106 USA: Lippincott Williams & Wilkins; 2008.
169. Skaane P, Skjennald A, Young K, Egge E, Jebsen I, Sager EM, et al. Follow-up and final results of the Oslo I Study comparing screen-film mammography and full-field digital mammography with soft-copy reading. Acta radiologica (Stockholm, Sweden : 1987). 2005;46(7):679-89.
170. Skaane P, Hofvind S, Skjennald A. Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: follow-up and final results of Oslo II study. Radiology. 2007;244(3):708-17.
171. Hofvind S, Skaane P, Elmore JG, Sebuødegård S, Hoff SR, Lee CI. Mammographic Performance in a Population-based Screening Program: Before, during, and after the Transition from Screen-Film to Full-Field Digital Mammography. Radiology. 2014;272(1):52-62.
172. Silva IdS. Cancer Epidemiology: Principles and Methods. Lyon, France: International Agency for Research on Cancer; 1999.
173. Hofvind S, Bjurstam N, Sorum R, Bjorndal H, Thoresen S, Skaane P. Number and characteristics of breast cancer cases diagnosed in four periods in the screening interval of a biennial population-based screening programme. Journal of medical screening. 2006;13(4):192-6.
174. Sprague BL, Conant EF, Onega T, Garcia MP, Beaber EF, Herschorn SD, et al. Variation in Mammographic Breast Density Assessments Among Radiologists in Clinical Practice: A Multicenter Observational Study. Annals of internal medicine. 2016;165(7):457-64.
175. Wang J, Azziz A, Fan B, Malkov S, Klifa C, Newitt D, et al. Agreement of Mammographic Measures of Volumetric Breast Density to MRI. PloS one. 2013;8(12).
176. Osteras BH, Martinsen AC, Brandal SH, Chaudhry KN, Eben E, Haakenaasen U, et al. Classification of fatty and dense breast parenchyma: comparison of automatic volumetric density measurement and radiologists' classification and their inter-observer variation. Acta radiologica (Stockholm, Sweden : 1987). 2016;57(10):1178-85.
177. Martin LJ, Minkin S, Boyd NF. Hormone therapy, mammographic density, and breast cancer risk. Maturitas. 2009;64(1):20-6.
178. Martin LJ, Melnichouk O, Guo H, Chiarelli AM, Hislop TG, Yaffe MJ, et al. Family history, mammographic density, and risk of breast cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2010;19(2):456-63.
179. Roman M, Graff-Iversen S, Weiderpass E, Vangen S, Sakshaug S, Hofvind S, et al. Postmenopausal Hormone Therapy and Breast Cancer Prognostic Characteristics: A Linkage between Nationwide Registries. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2016;25(11):1464-73.
180. Kamineni A, Anderson ML, White E, Taplin SH, Porter P, Ballard-Barbash R, et al. Body mass index, tumor characteristics, and prognosis following diagnosis of early stage breast cancer in a mammographically-screened population. Cancer causes & control : CCC. 2013;24(2):305-12.
181. Daling JR, Malone KE, Doody DR, Johnson LG, Gralow JR, Porter PL. Relation of body mass index to tumor markers and survival among young women with invasive ductal breast carcinoma. Cancer. 2001;92(4):720-9.
182. Skeie G, Mode N, Henningsen M, Borch KB. Validity of self-reported body mass index among middle-aged participants in the Norwegian Women and Cancer study. Clinical epidemiology. 2015;7:313-23.

183. Kvalitetsmanual. Mammografiprogrammet. Retningslinjer for radiograffaglig arbeid 2011. Available from: https://www.kreftregisteret.no/globalassets/mammografiprogrammet/publikasjoner-og-brosjyrer/kval-man-radiograf_v1.0_innholdsfortegnelse.pdf.
184. Pearl J, Bareinboim E. External Validity: From Do-Calculus to Transportability Across Populations. 2014:579-95.
185. Heckman JJ. Sample Selection Bias as a Specification Error. Econometrica. 1979;47(1):153-61.
186. Cortes C, Mohri M, Riley M, Rostamizadeh A. Sample Selection Bias Correction Theory. In: Freund Y, Györfi L, Turán G, Zeugmann T, editors. Algorithmic Learning Theory: 19th International Conference, ALT 2008, Budapest, Hungary, October 13-16, 2008 Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 38-53.
187. Jensen LF, Pedersen AF, Andersen B, Vedsted P. Identifying specific non-attending groups in breast cancer screening--population-based registry study of participation and socio-demography. BMC cancer. 2012;12:518.
188. Trewin CB, Strand BH, Weedon-Fekjaer H, Ursin G. Changing patterns of breast cancer incidence and mortality by education level over four decades in Norway, 1971-2009. European journal of public health. 2017;27(1):160-6.
189. Schouten LJ, de Rijke JM, Huveneers JA, Verbeek AL. Rising incidence of breast cancer after completion of the first prevalent round of the breast cancer screening programme. Journal of medical screening. 2002;9(3):120-4.
190. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Annals of internal medicine. 2007;146(7):502-10.
191. Moshina N, Ursin G, Hoff SR, Akslen LA, Roman M, Sebuodegard S, et al. Mammographic density and histopathologic characteristics of screen-detected tumors in the Norwegian Breast Cancer Screening Program. Acta radiologica open. 2015;4(9):2058460115604340.
192. Moshina N, Ursin G, Roman M, Sebuodegard S, Hofvind S. Positive predictive values by mammographic density and screening mode in the Norwegian Breast Cancer Screening Program. European journal of radiology. 2016;85(1):248-54.
193. Weinstein SP, Localio AR, Conant EF, Rosen M, Thomas KM, Schnall MD. Multimodality screening of high-risk women: a prospective cohort study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(36):6124-8.
194. Okello J, Kisembo H, Bugeza S, Galukande M. Breast cancer detection using sonography in women with mammographically dense breasts. BMC medical imaging. 2014;14:41.
195. Melnikow J, Fenton JJ, Whitlock EP, Miglioretti DL, Weyrich MS, Thompson JH, et al. U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews. Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the US Preventive Service Task Force. Rockville (MD): Agency for Healthcare Research and Quality (US); 2016.
196. Wengert GJ, Helbich TH, Woitek R, Kapetas P, Clauser P, Baltzer PA, et al. Inter- and intra-observer agreement of BI-RADS-based subjective visual estimation of amount of fibroglandular breast tissue with magnetic resonance imaging: comparison to automated quantitative assessment. European radiology. 2016;26(11):3917-22.
197. Redondo A, Comas M, Macià F, Ferrer F, Murta-Nascimento C, Maristany MT, et al. Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. The British journal of radiology. 2012;85(1019):1465-70.
198. Spayne MC, Gard CC, Skelly J, Miglioretti DL, Vacek PM, Geller BM. Reproducibility of BI-RADS Breast Density Measures Among Community Radiologists: A Prospective Cohort Study. The breast journal. 2012;18(4):326-33.
199. Damases CN, Brennan PC, Mello-Thoms C, McEntee MF. Mammographic Breast Density Assessment Using Automated Volumetric Software and Breast Imaging Reporting and Data System (BIRADS) Categorization by Expert Radiologists. Academic Radiology.23(1):70-7.

200. Shepherd JA, Kerlikowske K, Ma L, Duewer F, Fan B, Wang J, et al. Volume of mammographic density and risk of breast cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2011;20(7):1473-82.
201. Destounis S, Johnston L, Highnam R, Arieno A, Morgan R, Chan A. Using Volumetric Breast Density to Quantify the Potential Masking Risk of Mammographic Density. AJR American journal of roentgenology. 2017;208(1):222-7.
202. Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. Journal of the National Cancer Institute. 2010;102(16):1224-37.
203. Hofvind S, Bennett RL, Brisson J, Lee W, Pelletier E, Flugelman A, et al. Audit feedback on reading performance of screening mammograms: An international comparison. Journal of medical screening. 2016.
204. Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. Journal of medical screening. 2012;19 Suppl 1:57-66.
205. Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM. Comparing screening mammography for early breast cancer detection in Vermont and Norway. Journal of National Cancer Institute. 2008;100(15):1082-91.
206. Destounis SV, Morgan R, Arieno A. Screening for dense breasts: digital breast tomosynthesis. AJR American journal of roentgenology. 2015;204(2):261-4.

Appendix I

Study III: Questionnaire on various breast cancer risk factors, sent to all women together with the invitation to participate in the Norwegian Breast Cancer Screening Program between 2006 and 2016 (in Norwegian)

Helseforhold før 50-års alder

Dette skjemaet inneholder spørsmål om forhold som stort sett ikke forandres, f.eks utdanning. Du vil motta skjemaet kun én gang. Du kan delta i Mammografiprogrammet selv om du ikke leverer spørreskjemaet.

Informasjonen du gir vil lagres ved Kreftregisteret, uten tidsbegrensning. Den blir behandlet konfidensielt. Du har innsynsrett til egne opplysninger, og du kan til enhver tid be om at svarene dine slettes. Innlevering av skjema anses som samtykke til at svarene kan brukes slik det her er redegjort for. Spørsmål kan rettes til Kreftregisteret, telefon **22 45 13 00**

Hvordan fylle ut skjemaet?
Skjemaet leses maskinelt

- Sett kryss ☒ eller tall 2
- Bruk svart eller blå penn
- Ikke bruk desimaler
- Ved feil avkryssing: Fargelegg hele boksen ■

Noen spørsmål kan være vanskelig å besvare. Anslå likevel så godt du kan. Husk å ta med utfylt skjema når du møter til mammografi. **Takk for hjelpen!**

2.1

1 Fødested

Hvor er du født? 3413118405

☐ Norge ☐ Asia
☐ Europa utenfor Norge ☐ Afrika
☐ Nord-Amerika ☐ Oceania/Australia
☐ Latin-Amerika

2 Skolegang/utdanning

Hva er din høyeste fullførte skolegang/utdanning?

☐ Ingen fullført skolegang
☐ Grunnskole (barne-, ungdoms-, framhalds-, realskole)
☐ Videregående (gymnas, yrkes-, handels-, husmorskole)
☐ Universitets-/høgskoleutdanning inntil fire år
☐ Universitets-/høgskoleutdanning mer enn fire år

3 Høyde og vekt som barn og ungdom

Sett ett kryss i hver kolonne

7-års alder: (1. klasse) Vekt	Høyde		15-års alder: (konfirmasjon) Vekt	Høyde
☐	☐	Mye under middels	☐	☐
☐	☐	Noe under middels	☐	☐
☐	☐	Middels	☐	☐
☐	☐	Noe over middels	☐	☐
☐	☐	Mye over middels	☐	☐

Hva var din fødselsvekt? ☐☐☐☐ gram

4 Menstruasjon

Hvor gammel var du ved første menstruasjon?

☐☐ år gammel ☐ Aldri hatt menstruasjon

5 Sterilisering

Har du blitt sterilisert?

☐ Nei ☐ Ja, da jeg var ☐☐ år gammel

6 P-piller og hormonspiral

Har du brukt p-piller?

☐ Nei
☐ Ja, jeg begynte å bruke p-piller første gang da jeg var ☐☐ år gammel

Bruker du p-piller nå? ☐ Nei ☐ Ja

Til sammen har jeg brukt p-piller i ☐☐ år ☐☐ måneder

Har du brukt hormonspiral (Levonova/Mirena)?

☐ Nei
☐ Ja, jeg begynte å bruke hormonspiral første gang da jeg var ☐☐ år gammel

Bruker du hormonspiral nå? ☐ Nei ☐ Ja

Til sammen har jeg brukt hormonspiral i ☐☐ år ☐☐ måneder

7 Svangerskap og amming

Har du gjennomført svangerskap med varighet lenger enn seks måneder?

☐ Nei
☐ Ja, jeg har gjennomført (tvillinger regnes som ett svangerskap) ☐☐ svangerskap
og har født ☐☐ levende barn

Hvor gammel var du ved starten av første svangerskap som varte lenger enn seks måneder?

☐☐ år gammel

Har du ammet?

☐ Nei
☐ Ja, til sammen har jeg ammet i ☐☐ måneder

Helseforhold i dag

1 Mammografi

Har du tatt mammografi tidligere?

☐ Ja ☐ Nei ☐ Vet ikke

Hvis "JA":

Hvor tok du mammografi sist? *(kun ett kryss)*

☐ I det offentlige Mammografiprogrammet
☐ På et privat røntgeninstitutt
☐ På et sykehus (ikke det offentlige Mammografiprogrammet)

Når tok du denne undersøkelsen?

☐ For mindre enn 1 år siden
☐ For 1-2 år siden
☐ For mer enn 2 år siden

2 Høyde og vekt

Hvor høy er du i dag? ☐☐☐ (hele) cm

Hvor mye veier du i dag? ☐☐☐ (hele) kg

3 Medisinbruk

Bruker du regelmessig medisiner, foreskrevet av lege, for noen av følgende sykdommer?

☐ Nei
☐ Leddgikt (reumatisk sykdom)
☐ Benskjørhet (osteoporose)
☐ Diabetes (sukkersyke)
☐ Sykdom i skjoldbruskkjertelen
☐ Depresjon/angst
☐ Høyt kolesterol
☐ Høyt blodtrykk
☐ Hjertesykdom
☐ Astma

Bruker du Albyl-E, Globoid, Aspirin eller Dispril som fast medikasjon?

☐ Ja ☐ Nei

Bruker du Ibux, Brexidol, Voltaren, Ibumetin, Naproxen eller Diclofenac som fast medikasjon?

☐ Ja ☐ Nei

4 Røykevaner

Røyker du?

☐ Nei, har aldri røykt
☐ Nei, sluttet for ☐☐ år ☐☐ måneder siden
☐ Ja, jeg røyker om lag ☐☐☐ sigaretter pr. **uke**

5 Alkohol

Angi gjennomsnittlig alkoholforbruk per **måned**

☐ Drikker ikke alkohol
☐☐ glass rødvin/hvitvin
☐☐ halvlitere øl
☐☐ glass hetvin/brennevin

1744377777

6 Mosjon og trening

Hvor mye mosjonerer du ukentlig?

Mosjon: Lette gå- og sykkelturer, arbeid i hagen, snømåking og lignende

☐ Mosjonerer ikke
☐ 0-1 time pr. uke
☐ 2-3 timer pr. uke
☐ 4-5 timer pr. uke
☐ 6+ timer pr. uke

Hvor mye trener du ukentlig?

Trening: Regelmessige aktiviteter med **høy** intensitet av minst 1/2 times varighet hver gang, f. eks. aerobic, løping, sykling

☐ Trener ikke
☐ 0-1 time pr. uke
☐ 2-3 timer pr. uke
☐ 4-5 timer pr. uke
☐ 6+ timer pr. uke

7 Inngrep i bryst og underliv

Har du gjennomført brystreduksjon?

☐ Nei ☐ Ja, da jeg var ☐☐ år gammel

Har du operert inn brystprotese?

☐ Nei ☐ Ja, da jeg var ☐☐ år gammel

Har du fjernet begge eggstokkene?

☐ Nei
☐ Nei, jeg fjernet én eggstokk
☐ Ja, jeg fjernet **begge** da jeg var ☐☐ år gammel
☐ Vet ikke om én eller begge er fjernet

Har du fjernet livmoren?

☐ Nei
☐ Ja, da jeg var ☐☐ år gammel
☐ Vet ikke

8 Menstruasjon

Har du fortsatt menstruasjon?

Om menstruasjonen er regulert av hormonpreparater, svar "Ja"

☐ Ja
☐ Vet ikke, menstruerer uregelmessig
☐ Nei, menstruasjonen stoppet da jeg var ☐☐ år gammel

9 Mammografiprogrammet

Hvordan vurderer du informasjonen om Mammografi-programmet i invitasjonsbrevet og brosjyren?

☐ Veldig god ☐ Mangelfull
☐ God ☐ Veldig mangelfull

Vil du anbefale andre kvinner å delta i det offentlige Mammografiprogrammet?

☐ Ja ☐ Nei ☐ Vet ikke