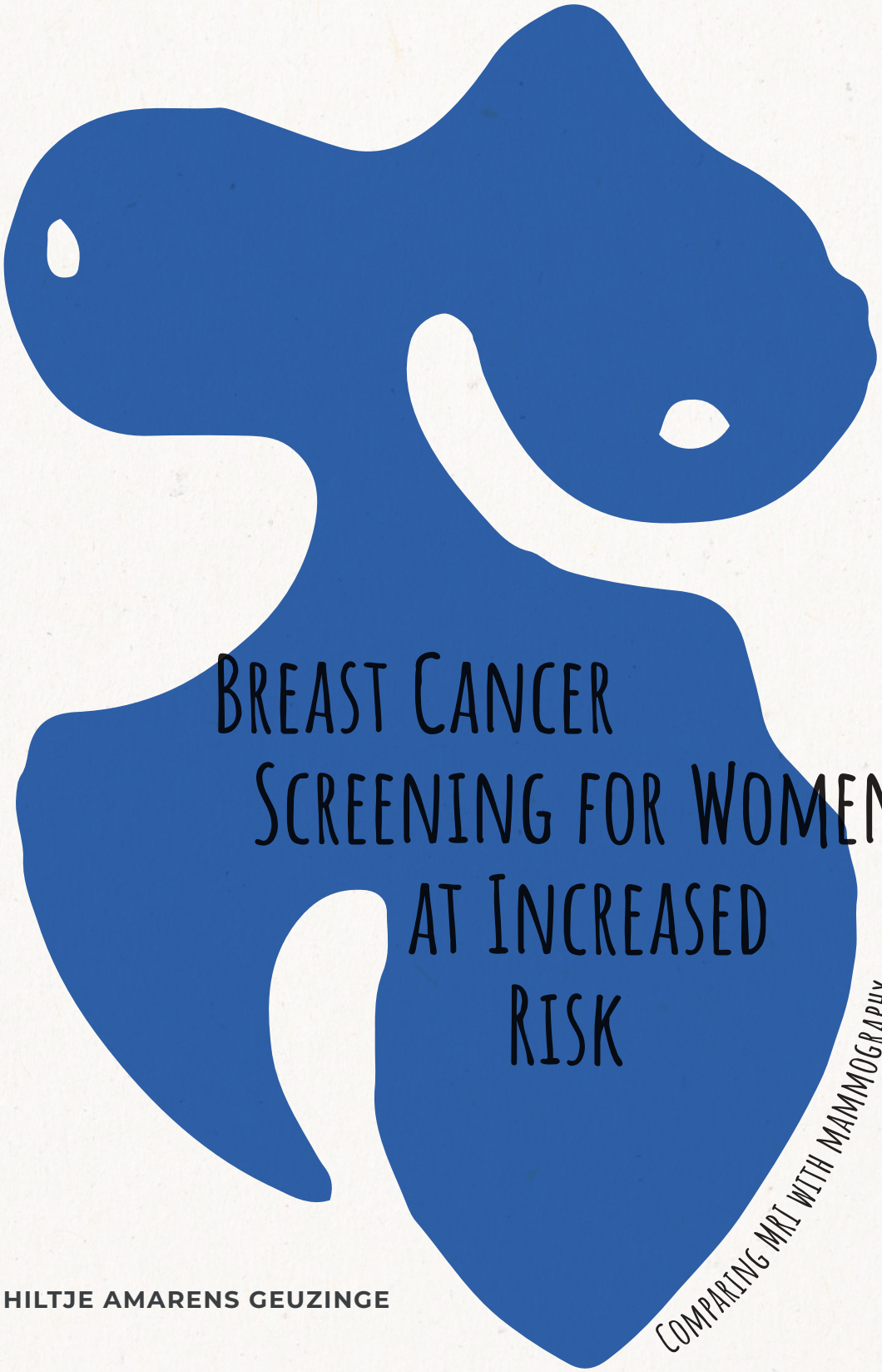


Breast Cancer Screening for Women at Increased Risk

Comparing MRI with mammography

HILTJE AMARENS GEUZINGE

Breast Cancer Screening for Women at Increased Risk

Comparing MRI with Mammography

Hiltje Amarens Geuzinge

ISBN/EAN: 978-94-6361-662-1

Layout, cover design and printing: Optima Grafische Communicatie, Rotterdam, the Netherlands
Cover art work: Mireia L.C.

Financial support for this thesis was kindly provided by the Erasmus University Rotterdam and the Department of Public Health, Erasmus Medical Center.

Breast Cancer Screening for Women at Increased Risk
Comparing MRI with Mammography

Borstkankerscreening voor vrouwen met een verhoogd risico
Een vergelijking van MRI met mammografie

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus
Prof. dr. A.L. Bredenoord
en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op
dinsdag 24 mei 2022 om 15.30 uur

door

Hiltje Amarens Geuzinge
geboren te Zwolle.

Erasmus University Rotterdam

PROMOTIECOMMISSIE:

Promotoren:	Prof. dr. H.J. de Koning Prof. dr. C. Verhoef
Overige leden:	Prof. dr. C.A. Uyl - De Groot Prof. dr. S. Siesling Prof. dr. M.L. Smidt
Copromotor:	Dr. E.A.M. Heijnsdijk

CONTENTS

Chapter 1

General introduction

GENERAL INTRODUCTION

1. Breast anatomy and breast cancer

The female breast consists of three types of tissue: 1) glandular tissue, which includes lobes and ducts; 2) fibrous connective tissue; and 3) adipose tissue (fat). Milk is produced in the lobes, and carried by the ducts to the nipple.

1.1 Breast density

The distribution of fibrous and glandular (fibroglandular) tissue relative to the amount of adipose tissue, is defined as breast density. If a breast contains a relatively high amount of fibroglandular tissue, and a relatively low amount of adipose tissue, it is defined as a dense breast. When it is the other way around, so the breast consists of a relatively large amount of adipose tissue, we call it a fatty breast. Adipose tissue is radiolucent, so it appears black on a mammogram. On the contrary, fibroglandular tissue appears white on a mammogram.

Breast density can be categorized into four categories, as defined in the Breast Imaging Reporting and Data System (BI-RADS) by the American College of Radiology[1]: A) Almost entirely fatty; B) Scattered fibroglandular; C) Heterogeneously dense; D) Extremely dense. Breast density can be estimated by a radiologist, and it can be measured automatically with Volpara Density software, both using mammography. Volpara density grades (VDG1-4) correspond to the BI-RADS breast density categories (A-D).

Breast density is generally high in young women, and it often decreases with increasing age.[2] At age 50, approximately 55% of Dutch women have dense breast tissue (VDG3-4), while at age 75 this percentage is approximately 18%.[3]

1.2 Breast cancer

Breast cancer can occur in different areas of the breast: the ducts, lobes or in the tissue in between. In most cases, breast cancer starts in the ducts. If a tumor remains within a duct, it is called ductal carcinoma in situ (DCIS), which is considered a non-invasive breast cancer. Once the tumor cells are also found in surrounding tissue, it is called invasive breast cancer. A common system for defining invasive breast cancer stages is the TNM-classification system.[4] The 'T' refers to 'Tumor', the size of the tumor at the moment of diagnosis. The 'N' refers to 'Node', which describes whether the cancer has spread to lymph nodes. The 'M' refers to 'Metastasis' which describes whether the cancer has spread to other parts of the body.

2. Breast cancer incidence

Breast cancer is the most common cancer in women in the Netherlands: 26% of all diagnosed cancers among women are breast cancers.[5] One out of seven Dutch women get a breast cancer diagnosis during their lifetime.[5] Compared to many other countries, the incidence of breast cancer is relatively high in the Netherlands, and it is still increasing as shown in **figure 1**.[5,6] This increase is caused by several factors such as the ageing population, improved screening, lifestyle factors, and reproductive factors such as increased age at first birth and decreasing parity.[7,8] In 2019, a total of 17,148 women were diagnosed with breast cancer, which equals a crude incidence rate of 196 per 100,000 women.[5] Approximately 86% of these cases were invasive breast cancers and 13% were DCIS.[5]

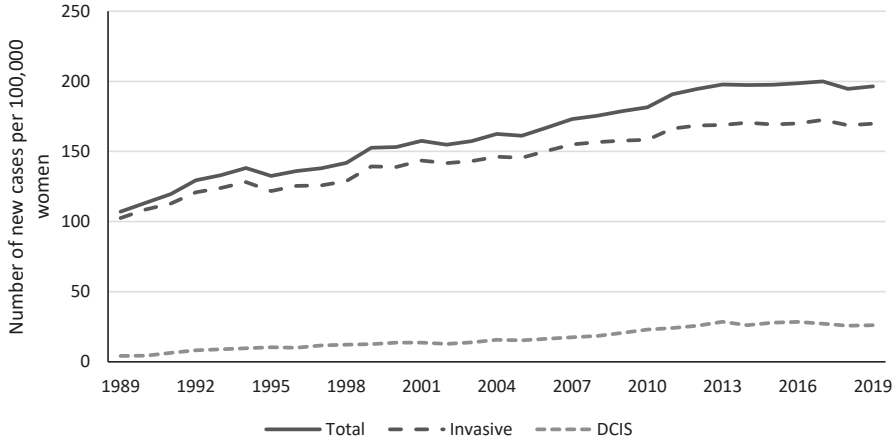


Figure 1. Breast cancer incidence in the Netherlands from 1989 until 2019
Data obtained from IKNL[5]

3. Risk factors

Several factors are associated with the risk of developing breast cancer. Some of these risk factors can be changed, while others are biologically determined.

A strong risk factor for developing breast cancer is having a family member who has previously been diagnosed with breast cancer. This is especially the case when a first-degree family member – so the mother or a sister – has had a breast cancer diagnosis. Approximately 25-30% of the breast cancers occur in women with a pathogenic variant in a penetrant breast cancer gene.[9,10] Most hereditary breast cancers are related to BRCA1/BRCA2 pathogenic variants. Women carrying a pathogenic variant in BRCA1 or BRCA2 are at high risk of developing breast cancer: they have a lifetime risk of approximately 50%.[11] Other genes in which pathogenic variants are known to be associated with a moderate

breast cancer risk are most often CHEK2, ATM and PALB2.[11] Since all aforementioned genes play a role in DNA repair, mutations in those genes can lead to abnormal cell growth.[12-14] In women with a family history without having a known pathogenic variant in a causative gene, it is known that single nucleotide polymorphisms (SNPs) play a role in the increased risk of developing breast cancer.[15]

Furthermore, breast cancer risk is known to increase with age. Of all breast cancers detected in the Netherlands in 2018, 87% occurred in women of 50 years and older.[5]

Breast density plays also an important role in the risk of developing breast cancer. Women with extremely dense breast tissue (BI-RADS category D) have a three- to sixfold higher risk of developing breast cancer than women with entirely fatty breasts (BI-RADS category A).[16] Molecular differences between dense breasts and fatty breasts probably explain the differences in breast cancer risk, but the exact mechanisms are not yet fully understood.[17] Approximately 8% of the Dutch screening population has extremely dense breasts.[18]

Also reproductive factors, such as age at first childbirth, breast feeding, age at menarche, and age at menopause play a role in the risk of developing breast cancer.[19-21] Especially menopausal status is a well-known risk factor: women at late menopause (>55 years old) have a twofold higher risk than women at early menopause (<40 years old).[19] Other risk factors are overweight, alcohol consumption, and use of hormone replacement treatments.[22]

4. Survival

Approximately 1 of 27 women in the Netherlands die from breast cancer.[5] The survival is highly associated with the breast cancer stage at detection: 95% of the women with a small localized tumor (i.e. <T2) is still alive after 10 years, whereas only 12% of the women with an advanced tumor at the moment of diagnosis which has spread to other parts of the body, is still alive 10 years after the diagnosis.[5] Due to screening and improved treatment options, breast cancer survival has improved substantially over the past decades.[23]

5. Breast cancer screening

The goal of breast cancer screening is to reduce breast cancer mortality by detecting the disease in an early stage. Most high-income countries offer national breast cancer screening programs.[6] In most of the countries, mammography is offered with an interval of two years. A mammogram is an X-ray image of the breast. Starting ages of the screening programs vary from 40 to 60 and stopping ages vary from 69 to 75. National breast cancer screening programs are targeted at women with an average risk of developing breast cancer.[6] Women with an elevated risk, for example due to pathogenic variants

in causative genes or a breast cancer family history, are often counseled and screened outside national screening programs, by offering them more intensive screening.[24-26]

5.1 Benefits and harms of screening

A meta-analysis on breast cancer mortality showed a 20% relative risk reduction of dying from breast cancer for women invited to screening, compared with controls.[27,28] Besides a mortality reduction, the detection of tumors in an earlier stage also leads to less advanced treatment and a gain in life years.

Unfortunately screening also causes harms. One of the most controversial harms is overdiagnosis. Overdiagnosis is the detection of tumors that would not have been diagnosed during a woman's lifetime if she was not being screened. A wide range of estimates of overdiagnosis in the Dutch mammography screening program have been published, ranging from 2% to 22%.[29,30] There are two explanations why some tumors do not become clinically detected in a situation without screening. The first one is that some tumors progress very slowly or progress not at all. This is especially the case for DCIS lesions,[31,32] but this is also possible for invasive breast cancer. The second reason is that women may die from other causes early in life, while having a progressive tumor. Since almost all breast cancer cases get treatment (also DCIS), overdiagnosis automatically results in overtreatment.

Other important downsides of screening are false-positive and false-negative screening results. A false-positive result is a positive screening result after which no breast cancer is diagnosed. This may cause distress and anxiety due to the thought of having breast cancer and due to additional diagnostic imaging and biopsies.[33] A false-negative screening result could results in a false reassurance: a woman gets a negative screening result while having breast cancer. False-negative screening outcomes could result in interval cancers: cancers being clinically detected in between two screening rounds due to the presence of symptoms. Interval cancers can also occur if the onset of the disease starts after a screening examination followed by rapid tumor growth. Approximately 24% of the cancers in women attending the Dutch national screening program are interval cancers,[34] which are often detected in an advanced stage and are therefore associated with worse prognosis than screen-detected cancers.[35]

5.2 Breast cancer screening in the Netherlands

In the Netherlands a nation-wide breast cancer screening program was introduced in 1989. Initially, women aged 50-69 years were invited to undergo mammography every two years. In 1998 the stopping age was extended from 69 to 74 years. From the introduction of the national screening program until 2008, analog screen-film mammography was used. In the period of 2008 until 2010, analog screen-film mammogram was fully replaced by digital mammography, which resulted in a higher breast cancer detection

rate.[36,37] The screening takes place in accredited centers and in mobile mammography units, and it is free of charge. Every year, approximately 1.3 million women are invited for a mammogram and approximately 76% of these women attend.[34] Since the beginning of 2021 the screening interval is temporarily extended to an interval of three years. This was caused by a lack of screening personnel and due to consequences of COVID-19 for the time needed to perform screening.[38]

In the Netherlands, screening outside the national screening program is indicated for several groups of women at increased breast cancer risk. Those women are offered yearly screening within a hospital setting, often starting at a younger age than in the national screening program. Women with a pathogenic variant in BRCA1/2, TP53, PALB2, CDH1, PTEN, ATM (c.7271T>G) or STK11 and women with a 50% chance of having a pathogenic variant in BRCA1/2 are offered yearly MRI screening.[26] Most of the carriers of other pathogenic variants and women with a breast cancer family history (without a known pathogenic variant) are offered mammography, sometimes in combination with clinical breast examination (a physical examination of the breast and underarm area).[26]

5.3 Breast cancer screening in the United States

In the United States no such breast cancer screening program as in the Netherlands is offered, and there are different sets of screening guidelines. Based on the United States Preventive Services Task Force (USPSTF) guidelines, women aged 50 to 74 are recommended to undergo mammography screening every two years. Women aged 40 to 49 years are recommended to discuss with their medical doctor whether they should already undergo mammography screening.[39] The American Cancer Society recommends annual mammography screening starting at age 45, and biennial mammography from age 54.[40] For women with a pathogenic variant in BRCA1/2, ATM, PALB2 or CHEK2, annual MRI screening is recommended by the National Comprehensive Cancer Network (NCCN).[41] However, studies evaluating the optimal screening approach for the latter three genes are scarce.

5.4 Screening and breast density

The performance of mammography is affected by breast density. As explained in **paragraph 1.1**, fibroglandular tissue appears white on a mammogram, which is also the case for a lesion. Thus, dense tissue can mask a tumor on a mammogram and it can therefore be missed. A previous Dutch study showed a sensitivity estimate of 61% in women with extremely dense breasts (VDG 4), compared to a sensitivity estimate of 86% for women with almost entirely fatty breasts (VDG 1).[18] This means that more interval cancers occur in women with extremely dense breasts, as shown in **figure 2**. The figure also shows that a higher breast cancer incidence in these women compared the women

with lower breast density.[18] This reflects the higher breast cancer risk among women with extremely dense breasts, as explained in **paragraph 3**.

In the Netherlands, breast density is not measured within the national breast cancer screening program. Therefore, most women in the Netherlands do not know whether they have dense breast tissue. In contrast, in the USA, breast density is repeatedly measured during screening, and women with dense breast tissue are informed about this and about the limited sensitivity of mammography. The reason why breast density is not measured within the Dutch screening program is the fact that no alternative screening strategy can be offered to these women.

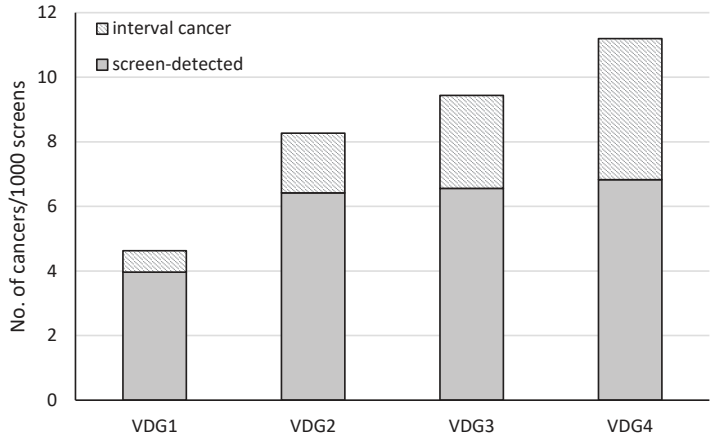


Figure 2. Number of screen-detected and interval cancers by VDG per 1000 screens
VDG: Volpara Density Grade
Volpara density grades (VDG1-4) correspond to the BI-RADS breast density categories (A-D)
Data obtained from Wanders et al[18]

5.5 MRI breast cancer screening

Besides mammography, MRI can be used for breast cancer screening. MRI uses a strong magnetic field and radio waves to produce an image of the breast. Before a woman undergoes a breast MRI examination, a gadolinium contrast fluid is injected into the body, which improves the quality of the MRI images. Currently in most Western countries, MRI screening is only used in women with the highest breast cancer risk. Not only the high breast cancer risk plays a role in offering these women MRI, but also the fact that their risk is already high at a relatively young age, when breast density is generally high.[42] The main advantages of MRI relative to mammography are that it generally results in a high detection rate and that it performs well in women with dense breasts.[42,43] Unfortunately using MRI for breast cancer screening has also some downsides. Due to the higher detection rate of MRI, it is also associated with more overdiagnosis compared

to mammography screening. Another downside is the fact that MRI screening is often associated with more false positive results.[44] Furthermore, MRI is about three to four times more expensive than mammography.

6. Evaluating outcomes of breast cancer screening

6.1 Clinical trials

Performing a randomized controlled trial (RCT) is thought the be the gold standard in measuring the effectiveness of an intervention.[45] In an RCT, participants are randomly assigned to either the intervention group or the control group. Randomization balances characteristics of the participants between the groups, so a difference in outcomes between the groups can be attributed to the intervention.

Recently in the Netherlands, two large RCTs have been performed evaluating mammography plus MRI (intervention group), relative to mammography alone (control group) in two different risk groups. These trials were the Familial MRI Screening Study (FaMRIsc) trial and the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial. In the FaMRIsc trial, women with a breast cancer family history but without a known pathogenic variant in BRCA1/2 or TP53 genes were included.[46] In the DENSE trial, women with extremely dense breast tissue were invited, who participated in the Dutch national screening program and who had a negative mammography result.[47] Outcomes of both trials will be discussed in this thesis.

Since RCTs mostly cover a limited follow-up period, it is difficult to evaluate the effect of screening on cancer mortality. As mentioned previously, cancer screening has the goal to detect cancer in an early stage to improve the chance for successful treatment and thereby improve survival. To be able to measure survival – or mortality – a lengthy follow-up period is needed, which can be difficult to reach because of time constraints, high costs and loss to follow-up of participants. At the same time, due to rapid technological developments, screening tools can be improved or replaced over the follow-up time which may make the screening test of interest not relevant anymore. Another limitation of RCTs in cancer screening are the limited amount of screening strategies that can be investigated. The more strategies one wants to investigate, the more participants are needed.

6.2 Microsimulation modeling

Modeling can overcome the shortcoming of RCTs by extrapolating evidence from RCTs. In this thesis, the Microsimulation Screening ANalysis (MISCAN) model was used to predict the incidence, mortality and cost-effectiveness of several screening strategies. The model simulates individual natural life histories from birth to death, and the natural history of breast cancer. The model consists of three parts: the demographic part, the natural history part and the screening part (**figure 3**). First, a large number of women is

simulated one by one (based on demography data). Subsequently, some of the simulated women develop breast cancer, based on the natural history part of the model. Those breast cancers can become clinically detected, screen-detected or they may progress into a larger tumor size. Finally, some women die from breast cancer, while others die from other causes. Several screening strategies can be modelled which can influence the tumor stage (TNM) at detection, and thereby possibly also influence survival.[48]

In this thesis, two versions of the MISCAN model were used: MISCAN-breast and MISCAN-Fadia (also referred to as Model E). MISCAN-breast, developed in the eighties and adjusted since, is mainly based on Dutch and European data and uses a semi-Markov process to model the development and growth of breast cancer.[48] MISCAN-Fadia, developed in the CISNET consortium since 2000, is based on US data only, and it models the disease using a continuous growth principle and applies a fatal diameter threshold. The fatal diameter is applied as the minimal tumor size from which the disease cannot be cured. If a tumor is detected at a smaller size than the fatal diameter, the tumor can be cured.[49] Besides the two MISCAN models, we also presented results of a microsimulation model from the University of Wisconsin-Madison (Madison, Wisconsin) and Harvard Medical School (Boston, Massachusetts): Model W-H. Model W-H is a discrete-event model, based on US data, also developed in the CISNET consortium.

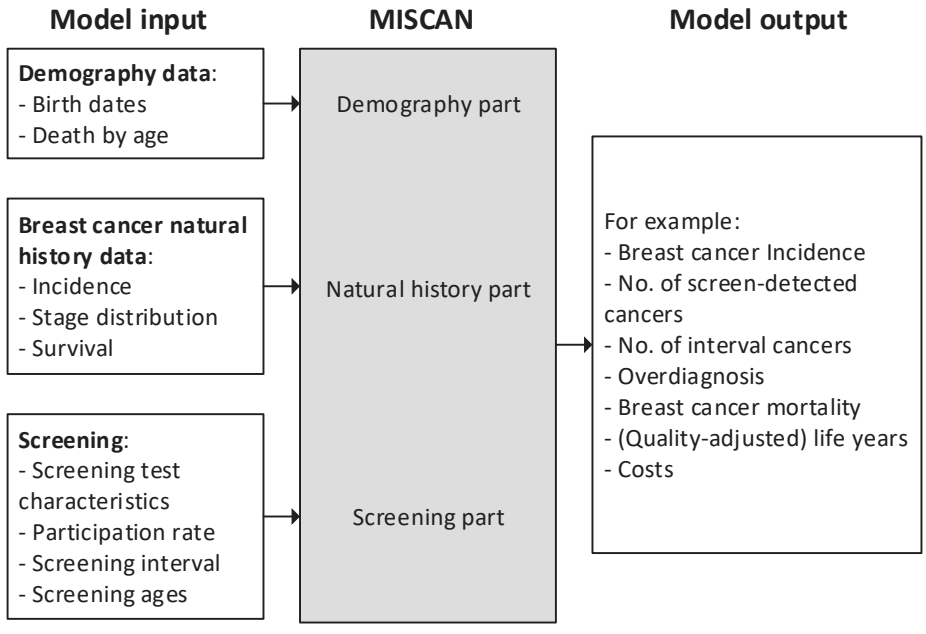


Figure 3. Overview of the MISCAN model

6.3 Cost-effectiveness analysis

Not only the performance of screening strategies should be evaluated, but also its cost-effectiveness. Since the Dutch government wants to decelerate the rising health care expenditure, cost-effectiveness analyses are needed to be able to achieve the highest value for money.

In a cost-effectiveness analysis, the costs and health effects of multiple interventions - in this case screening strategies - are compared to find the optimal intervention. Cost-effectiveness analyses are mainly performed using a modelling approach such as described in the previous paragraph, to include all medical costs and health effects associated with the interventions on a lifetime horizon. Health effects are often expressed in quality-adjusted life years (QALYs), which is a measure including years of life and the quality of life during those years. Results of cost-effectiveness are presented as a ratio of the difference in costs divided by the difference in health effects between an intervention compared to the next best intervention. Finally, a willingness to pay threshold needs to be applied to conclude whether a cost-effectiveness ratio is acceptable.

OUTLINE OF THIS THESIS

The aim of this thesis is to evaluate whether MRI screening would be an acceptable screening tool for women at increased breast cancer risk. Risk groups include women with a breast cancer family history; women with a pathogenic variant in ATM, PALB2 or CHEK2 genes, and women with extremely dense breast tissue. This thesis consists of three parts. In Part I, the effectiveness of MRI screening is evaluated, using data of the FaMRIsc trial, the DENSE trial and long-term estimations using microsimulation modelling. In Part II, the cost-effectiveness of multiple MRI screening strategies is evaluated, with the use of microsimulation modelling. Part III focusses on women's opinions regarding MRI screening relative to mammography.

Part I: effectiveness of MRI screening versus mammography

In **chapter 2**, the performance of the two screening strategies in the FaMRIsc trial are evaluated. In the FaMRIsc trial women aged 30-55 years with a cumulative lifetime risk for breast cancer of ≥20% due to a family history, were randomly assigned into two arms. Women in the intervention arm received yearly MRI and clinical breast examination, and mammography biennially. Women in the control arm received yearly mammography and clinical breast examination. We evaluate the sensitivity, specificity, positive predictive value and tumor characteristics of both screening strategies. Furthermore, we stratify these outcomes by breast density. In **chapter 3** outcomes of several screening strategies consisting of mammography and MRI with varying starting ages are evaluated

in women with an ATM, CHEK2 or PALB pathogenic variant, using two microsimulation models (MISCAN-Fadia and model W-H). Outcomes consist of life years gained (LYG), breast cancer deaths averted, total screens, false-positive screens, and benign biopsies, all on a lifetime horizon. In **chapter 4**, we focus on breast density: screening outcomes of mammography and supplemental MRI screening of the DENSE trial are stratified by women whose breast density remained extremely high and women whose breast density decreased. Hereby, we evaluate whether women with extremely dense breast tissue who are on an MRI screening scheme, should still be screened with MRI when their breast density decreases.

Part II: cost-effectiveness of MRI screening versus mammography

In part II, the cost-effectiveness of several screening strategies containing MRI and mammography are evaluated using microsimulation modelling. **Chapter 5** focusses on women with a cumulative lifetime risk for breast cancer of ≥20% due to their family history: the same population as in chapter 2. **Chapter 6** focusses on women with extremely dense breast tissue who are currently screened within the Dutch national screening program.

Part III: what do women want?

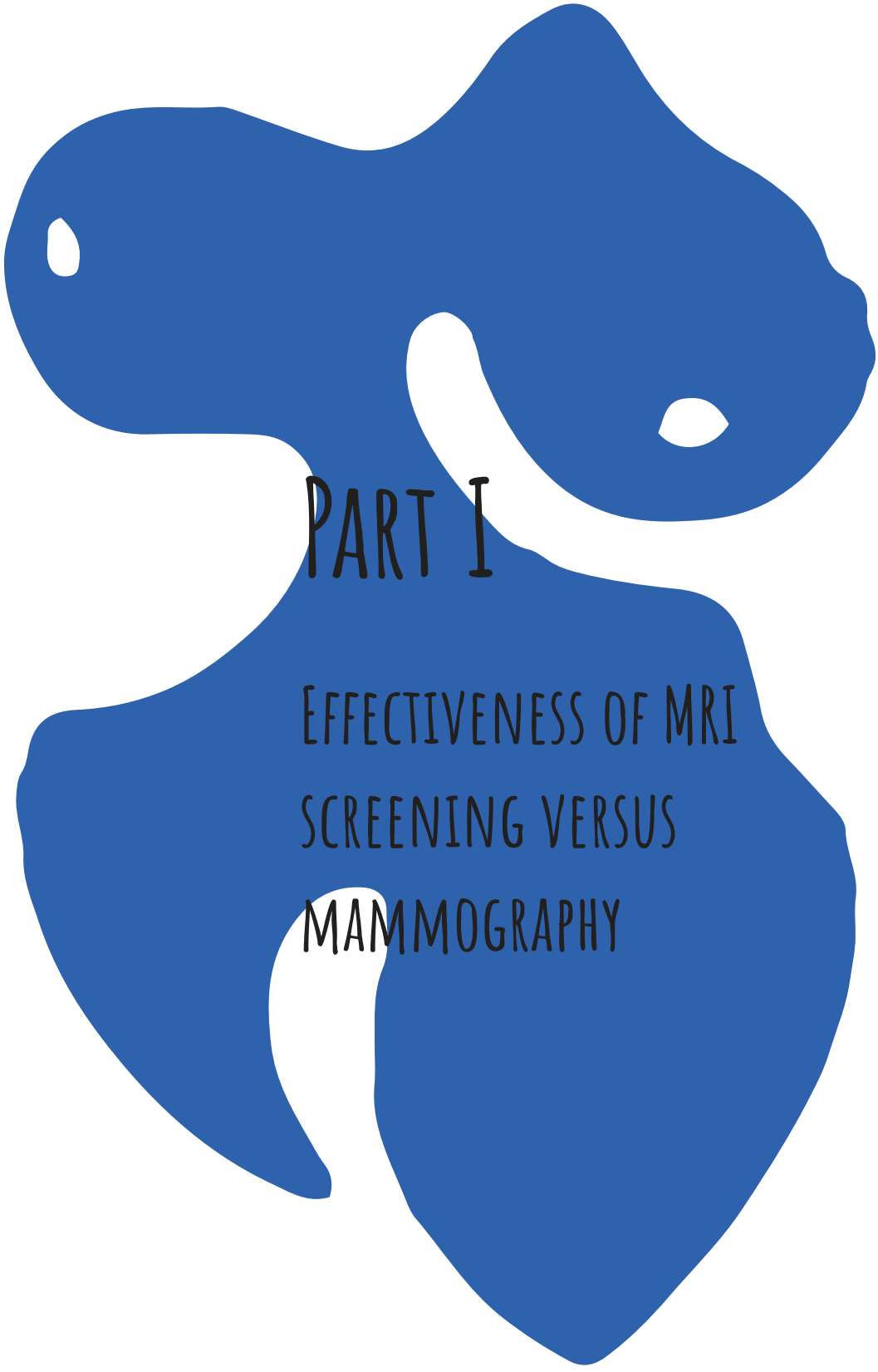
This part consists of **chapter 7**, in which the expectations, experiences and preferences of women themselves regarding MRI and mammography screening are evaluated. Outcomes of a questionnaire, which was sent to participants of the FaMRIsc trial, are compared between the two trial groups to evaluate whether women who are on an MRI screening scheme have a different opinion on MRI than women who are not screened with MRI.

REFERENCES

1. D'Orsi C, Sickles E, Mendelson E, Morris E. *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System.* Reston, VA: American College of Radiology; 2013.
2. Burton A, Maskarinec G, Perez-Gomez B, et al. Mammographic density and ageing: A collaborative pooled analysis of cross-sectional data from 22 countries worldwide. *PLoS Med.* 2017;14(6):e1002335.
3. Wanders JOP, Holland K, Karssemeijer N, et al. Changes in volumetric breast density and the association with breast cancer risk. *Submitted.*
4. American Joint Committee on Cancer (AJCC). *AJCC Cancer Staging Manual.* 8th ed. New York, NY: Springer; 2017.
5. Netherlands Comprehensive Cancer Organisation (IKNL). NKR cijfers. 2020; www.iknl.nl/nkr-cijfers. Accessed February 18, 2021.
6. Altobelli E, Rapacchietta L, Angeletti PM, Barbante L, Profeta FV, Fagnano R. Breast Cancer Screening Programmes across the WHO European Region: Differences among Countries Based on National Income Level. *Int J Environ Res Public Health.* 2017;14(4).
7. van der Waal D, Verbeek AL, den Heeten GJ, Ripping TM, Tjan-Heijnen VC, Broeders MJ. Breast cancer diagnosis and death in the Netherlands: a changing burden. *Eur J Public Health.* 2015;25(2):320-324.
8. Coughlin SS. Epidemiology of Breast Cancer in Women. *Adv Exp Med Biol.* 2019;1152:9-29.
9. Melchor L, Benitez J. The complex genetic landscape of familial breast cancer. *Hum Genet.* 2013;132(8):845-863.
10. Eccles SA, Aboagye EO, Ali S, et al. Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Res.* 2013;15(5):R92.
11. Hu C, Hart SN, Gnanaolivu R, et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med.* 2021;384(5):440-451.
12. Ahmed M, Rahman N. ATM and breast cancer susceptibility. *Oncogene.* 2006;25(43):5906-5911.
13. Zannini L, Delia D, Buscemi G. CHK2 kinase in the DNA damage response and beyond. *J Mol Cell Biol.* 2014;6(6):442-457.
14. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet.* 2007;39(2):165-167.
15. Evans DG, Brentnall A, Byers H, et al. The impact of a panel of 18 SNPs on breast cancer risk in women attending a UK familial screening clinic: a case-control study. *J Med Genet.* 2017;54(2):111-113.
16. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2006;15(6):1159-1169.
17. Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer. *Breast Cancer.* 2018;25(3):259-267.
18. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment.* 2017;162:95-103.
19. Collaborative Group on Hormonal Factors in Breast C. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol.* 2012;13(11):1141-1151.
20. MacMahon B, Cole P, Lin TM, et al. Age at first birth and breast cancer risk. *Bull World Health Organ.* 1970;43(2):209-221.

21. Collaborative Group on Hormonal Factors in Breast Cancer. 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.* 2002;360(9328):187-195.
22. Evans DG, Howell A. Can the breast screening appointment be used to provide risk assessment and prevention advice? *Breast Cancer Res.* 2015;17:84.
23. de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. *Int J Cancer.* 2015;137(1):165-172.
24. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer Journal for Clinicians.* 2007;57(3):75-89.
25. National Collaborating Centre for Cancer. Clinical Guideline. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. 2013; https://www.nice.org.uk/guidance/cg164. Accessed March 2, 2021.
26. NABON. Richtlijn Borstkanker. 2018; https://richtlijnendatabase.nl/richtlijn/borstkanker/screening.html. Accessed March 2, 2021.
27. Independent U.K. Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet.* 2012;380(9855):1778-1786.
28. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev.* 2009(4):CD001877.
29. van Luijt PA, Rozemeijer K, Naber SK, et al. The role of pre-invasive disease in overdiagnosis: A microsimulation study comparing mass screening for breast cancer and cervical cancer. *J Med Screen.* 2016;23(4):210-216.
30. Ripping TM, Verbeek AL, Fracheboud J, de Koning HJ, van Ravesteyn NT, Broeders MJ. Overdiagnosis by mammographic screening for breast cancer studied in birth cohorts in The Netherlands. *Int J Cancer.* 2015;137(4):921-929.
31. Roses RE, Arun BK, Lari SA, et al. Ductal carcinoma-in-situ of the breast with subsequent distant metastasis and death. *Ann Surg Oncol.* 2011;18(10):2873-2878.
32. Groen EJ, Elshof LE, Visser LL, et al. Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). *Breast.* 2017;31:274-283.
33. Bond M, Pavey T, Welch K, et al. Psychological consequences of false-positive screening mammograms in the UK. *Evid Based Med.* 2013;18(2):54-61.
34. Netherlands Comprehensive Cancer Organisation (IKNL). Monitor bevolkingsonderzoek borstkanker 2018/2019. 2020; https://iknl.nl/getmedia/cc1b7da5-28aa-43e6-b70a-b3dadf08ab89/monitor-bevolkingsonderzoek-borstkanker-2018-2019.pdf. Accessed February 25, 2021.
35. Hofvind S, Holen A, Roman M, Sebuodegard S, Puig-Vives M, Akslen L. Mode of detection: an independent prognostic factor for women with breast cancer. *J Med Screen.* 2016;23(2):89-97.
36. Sankatsing VDV, Fracheboud J, de Munck L, et al. Detection and interval cancer rates during the transition from screen-film to digital mammography in population-based screening. *BMC Cancer.* 2018;18(1):256.
37. de Munck L, de Bock GH, Otter R, et al. Digital vs screen-film mammography in population-based breast cancer screening: performance indicators and tumour characteristics of screen-detected and interval cancers. *Br J Cancer.* 2016;115(5):517-524.

38. National Institute for Public Health and the Environment (RIVM). Tijdelijke verlenging uitnodigingsinterval naar maximaal 3 jaar. 2020; https://www.rivm.nl/bevolkingsonderzoek-borstkanker/mammografie/later-uitgenodigd. Accessed March 2, 2021.
39. Siu AL, Force USPST. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2016;164(4):279-296.
40. Oeffinger KC, Fontham ET, Etzioni R, et al. Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society. *JAMA.* 2015;314(15):1599-1614.
41. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology. Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic. Version 2.2021. https://www.nccn.org/professionals/physician_gls/pdf/genetics_bop.pdf. Accessed April 2, 2021.
42. Lehman CD, Smith RA. The role of MRI in breast cancer screening. *J Natl Compr Canc Netw.* 2009;7(10):1109-1115.
43. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.
44. Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. *Ann Intern Med.* 2008;148(9):671-679.
45. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG.* 2018;125(13):1716.
46. Saadatmand S, Rutgers EJ, Tollenaar RA, et al. Breast density as indicator for the use of mammography or MRI to screen women with familial risk for breast cancer (FaMRIsc): a multicentre randomized controlled trial. *BMC Cancer.* 2012;12:440.
47. Emaus MJ, Bakker MF, Peeters PH, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. *Radiology.* 2015;277(2):527-537.
48. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Comput Methods Programs Biomed.* 1985;20(1):79-93.
49. van den Broek JJ, van Ravesteyn NT, Heijnsdijk EA, de Koning HJ. Simulating the Impact of Risk-Based Screening and Treatment on Breast Cancer Outcomes with MISCAN-Fadia. *Med Decis Making.* 2018;38(1_suppl):54S-65S.



Part I

Effectiveness of MRI screening versus mammography

Chapter 2

MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial

Saadatmand S*, Geuzinge HA*, Rutgers EJT, Mann RM, de Roy van Zuidewijn DBW, Zonderland HM, Tollenaar RAEM, Lobbes MBI, Ausems MGEM, van 't Riet M, Hooning MJ, Mares-Engelberts I, Luiten EJT, Heijnsdijk EAM, Verhoef C, Karssemeijer N, Oosterwijk JC, Obdeijn IM, de Koning HJ, Tilanus-Linthorst MMA, on behalf of the FaMRIsc study group
*These authors contributed equally to this study

Lancet Oncology 2019;20(8):1136-1147

ABSTRACT

Background: Approximately 15% of all breast cancers occur in women with a family history of breast cancer, but for whom no causative hereditary gene mutation has been found. Screening guidelines for women with familial risk of breast cancer differ between countries. We did a randomised controlled trial (FaMRIsc) to compare MRI screening with mammography in women with familial risk.

Methods: In this multicentre, randomised, controlled trial done in 12 hospitals in the Netherlands, women were eligible to participate if they were aged 30–55 years and had a cumulative lifetime breast cancer risk of at least 20% because of a familial predisposition, but were BRCA1, BRCA2, and TP53 wild-type. Participants who were breastfeeding, pregnant, had a previous breast cancer screen, or had a previous a diagnosis of ductal carcinoma in situ were eligible, but those with a previously diagnosed invasive carcinoma were excluded. Participants were randomly allocated (1:1) to receive either annual MRI and clinical breast examination plus biennial mammography (MRI group) or annual mammography and clinical breast examination (mammography group). Randomisation was done via a web-based system and stratified by centre. Women who did not provide consent for randomisation could give consent for registration if they followed either the mammography group protocol or the MRI group protocol in a joint decision with their physician. Results from the registration group were only used in the analyses stratified by breast density. Primary outcomes were number, size, and nodal status of detected breast cancers. Analyses were done by intention to treat. This trial is registered with the Netherlands Trial Register, number NL2661.

Findings: Between Jan 1, 2011, and Dec 31, 2017, 1355 women provided consent for randomisation and 231 for registration. 675 of 1355 women were randomly allocated to the MRI group and 680 to the mammography group. 218 of 231 women opting to be in a registration group were in the mammography registration group and 13 were in the MRI registration group. The mean number of screening rounds per woman was 4.3 (SD 1.76). More breast cancers were detected in the MRI group than in the mammography group (40 vs 15; p=0.0017). Invasive cancers (24 in the MRI group and eight in the mammography group) were smaller in the MRI group than in the mammography group (median size 9 mm [5–14] vs 17 mm [13–22]; p=0.010) and less frequently node positive (four [17%] of 24 vs five [63%] of eight; p=0.023). Tumour stages of the cancers detected at incident rounds were significantly earlier in the MRI group (12 [48%] of 25 in the MRI group vs one [7%] of 15 in the mammography group were stage T1a and T1b cancers; one (4%) of 25 in the MRI group and two (13%) of 15 in the mammography group were stage T2 or higher; p=0.035) and node-positive tumours were less frequent (two [11%]

of 18 in the MRI group vs five [63%] of eight in the mammography group; p=0.014). All seven tumours stage T2 or higher were in the two highest breast density categories (breast imaging reporting and data system categories C and D; p=0.0077) One patient died from breast cancer during follow-up (mammography registration group).

Interpretation: MRI screening detected cancers at an earlier stage than mammography. The lower number of late-stage cancers identified in incident rounds might reduce the use of adjuvant chemotherapy and decrease breast cancer-related mortality. However, the advantages of the MRI screening approach might be at the cost of more false-positive results, especially at high breast density.

INTRODUCTION

Approximately 15% of all breast cancers occur in women with a family history of breast cancer (familial risk) in whom no causative hereditary gene mutation has been found.[1] These women are at greater risk for breast cancer at a relatively young age.[2] In women with breast cancer, overall survival decreases considerably with increasing tumour size at detection and number of axillary lymph nodes involved, even with optimal adjuvant systemic therapy.[3,4] Screening aims to improve survival by detecting breast cancer at an early stage. However, it can also result in false-positive results.

Between 2004–18, several screening trials comparing MRI and mammography in women at high risk of developing breast cancer concluded that adding MRI to mammography screening improves early breast cancer detection in women with a familial or genetic predisposition.[5–7] As a result, guidelines for breast cancer screening were modified globally.[8–10] Unfortunately, these trials were all non-randomised studies with a paired design in which MRI and mammography were done simultaneously.[5–7,11] Therefore, it is unknown when an MRI-only detected tumour would have been detected by mammography, and whether this would have identified a difference in tumour stage that was clinically relevant. With this limited evidence, screening guidelines for women with familial risk differ between countries. American guidelines advise annual mammography, clinical breast examination, and MRI for women aged 30 years or older with a cumulative lifetime risk of at least 20%.[8] Dutch and UK guidelines omit MRI for women with familial risk without a BRCA1/2 mutation.[9,10]

Furthermore, breast density has not been considered in these studies.[5,6] Higher breast density, caused by more glandular and connective breast tissue in relation to fat, indicates a higher cancer risk overall and in women with familial risk.[12] High-density breast tissue impairs sensitivity of mammography,[12] but has less of an effect on MRI[13] and might cause a different amount of false-positive results for mammography than for MRI. Breast density is high in about 74% of women between 40–49 years of age, and in 45% of women in their 60s.[14] MRI might not be necessary for all women with familial breast cancer risk,[15] but breast density might be a parameter to identify subgroups of women for whom MRI screening could be useful.

The Familial MRI Screening study (FaMRIsc) was done to compare annual MRI and clinical breast examination plus biennial mammography versus screening with annual mammography and clinical breast examination in women with a familial breast cancer risk but without a known BRCA1/2 or TP53 mutation.

METHODS

Study design and participants

The FaMRIsc study was a multicentre, randomised, controlled trial. Women were eligible to participate if they were aged 30–55 years and had a cumulative lifetime breast cancer risk of at least 20% because of a familial predisposition according to the modified tables of Claus,[5,16] or as assessed at a clinical genetics centre. Exclusion criteria were previous invasive cancer or BRCA1, BRCA2, or TP53 mutations (proven or 50% risk of mutation), since MRI screening is already advised for these groups,[8–10] and a contraindication for contrast enhanced MRI. Participants were removed from the study after randomisation if they met one of the exclusion criteria, or no longer met the inclusion criteria. Previous screening, a ductal carcinoma in situ diagnosis, pregnancy, and breastfeeding were permitted.

Participants were recruited from 12 outpatient breast cancer clinics or family cancer clinics at seven academic medical centres in the Netherlands and five of the larger hospitals (appendix). The physician of the outpatient clinic or family cancer clinic enrolled participants after they provided written informed consent. The study follows the Declaration of Helsinki and was approved by the Institutional Review Board of the Erasmus University Medical Center (Rotterdam, Netherlands; reference number MEC-2010-292). The study protocol was published previously.[17]

Randomisation

Participants were randomly allocated (1:1) to receive either annual MRI and clinical breast examination plus biennial mammography (MRI group) or annual mammography and clinical breast examination (mammography group). Randomisation was done via a web-based system and stratified by centre. Allocation was based on a general number between 1–100 that was randomly generated by the computer. An algorithm decreased the possibility of the computer generating a number that led to allocation in an overrepresented study group by a factor of 5 minus 1. However, it remained impossible to predict what allocation would follow for the physician, participant, or researcher overseeing the randomisation.

Procedures

The mammography group received annual mammography according to Dutch guidelines[9] plus clinical breast examination. Dutch guidelines recommend clinical breast examination in women with a lifetime risk of breast cancer of at least 30%. The MRI group was screened with annual MRI and clinical breast examination, and mammography biennially. Leaving out mammography every other year was considered safe in the MRI group and might prevent overdiagnosis of low-grade ductal carcinoma in situ.[18] Women who did not provide consent for randomisation were given the option

to be in a registration group, if screened via one of the same methods as either group in the study. Women in the registration group were screened according to the MRI group protocol in a joint decision with their physician.

Mammographic examination was done with full-field digital mammography. All mammography examinations were assessed according to the 4th edition of the American College of Radiology (ACR) breast imaging reporting and data system (BI-RADS) and all MRI examinations were assessed according to the 1st edition of the ACR BI-RADS.[19] MRI and mammography were preferably scheduled on the same day for participants receiving both.

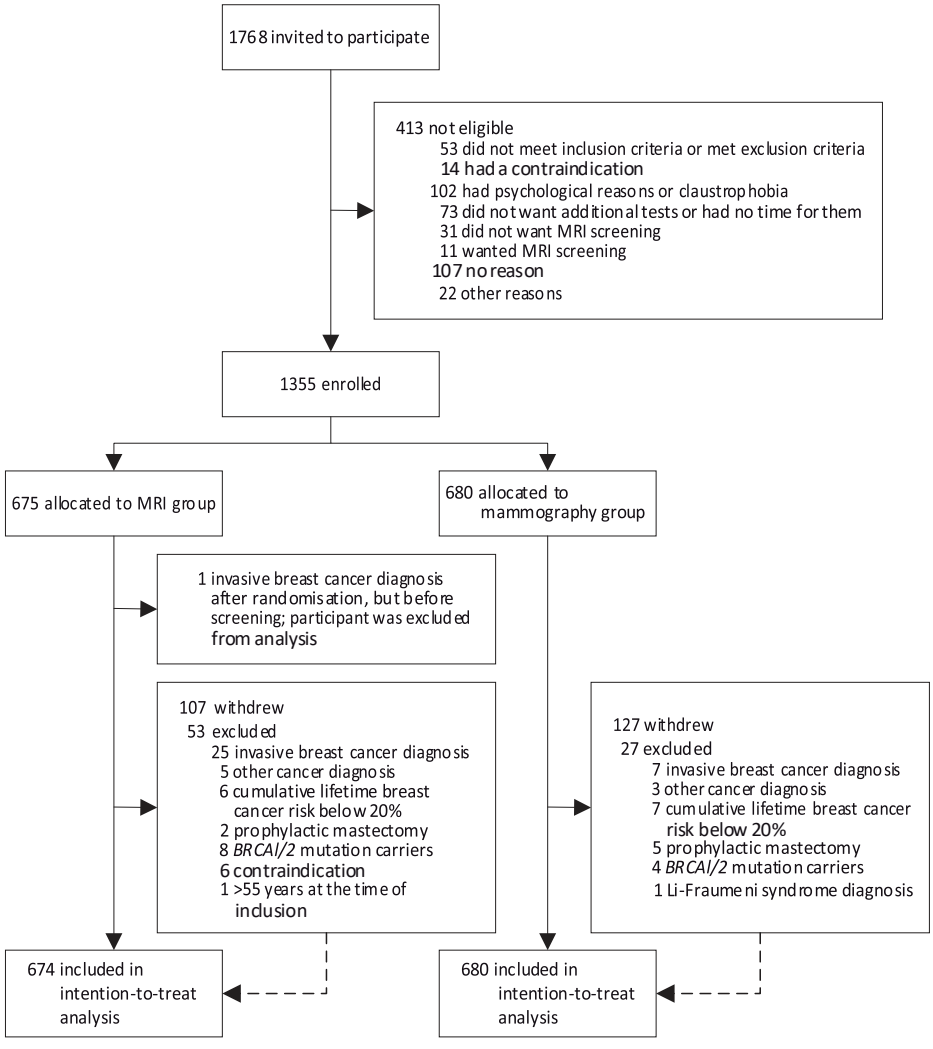


Figure 1. Trial profile

A positive screening test was defined as a mammographic or MRI examination with a BI-RADS score of 3, for which additional investigation or a repeat examination at 6 months per radiological judgement followed; a mammographic or MRI examination with a BI-RADS score of 4–5, indicating histology; or a clinical breast examination with an abnormality, for which additional diagnostic testing was recommended. In the MRI group, MRI and mammography were not independently read; RMM, HMZ, MBIL, I-MO, CdM, CEL, HBWG, MNJMW, WBV, ET assessed the images. To determine mammographic density, an automated breast density measurement with Volpara Density (version 1.3.0, Volpara Solutions, New Zealand)[20] was done on raw data of the first digital mammogram of all participants, and estimated from the mammograms by radiologists according to the BI-RADS breast composition categories: A=fatty, B=scattered fibroglandular, C=heterogeneously dense, D=extremely dense.[19] Dynamic contrast-enhanced breast MRI examinations were done according to our study protocol.[17]

Outcomes

Primary outcomes of this study were numbers (of both ductal carcinoma in situ and invasive cancers), size, and nodal status of detected breast cancers. Secondary outcomes were false-positive results, sensitivity and specificity of each screening method, positive predictive value of a BI-RADS score of 3 or above, and positive predictive value of biopsies. Cost-effectiveness and breast cancer mortality will be reported in future analyses.

Statistical analysis

The sample size was calculated on the basis of the breast cancer incidence of 7 per 1000 person-years at risk among women with familial risk screened in the Dutch MRI screening study.[5] We expected a sensitivity of 70% for MRI and 40% for mammography on the basis of previous studies.[5] After 4000 woman-years at risk in both study groups—eg, 1000 women in each group for 4 years—we expected the detection of 32 tumours in the MRI group and 18 tumours in the mammography group. With these 50 cancers, a difference in tumour size of 8 mm (SD 9) was expected to be significant (with a two-sided α-level of 0.05 and a power of 80%). A difference of 8 mm in tumour size was also considered to be clinically relevant. Accrual and the number of detected cancers were evaluated after 2, 4, and 6 years. Fewer women were enrolled and randomly assigned than expected and 50 breast cancers were not reported after 4 years; therefore, the study was continued for 3 additional years to reach this threshold. The study ended Dec 30, 2017.

All women who provided consent for randomisation and were screened at least once were included in all analyses (done by intention to treat). Data from participants who were excluded (eg, after they no longer met inclusion criteria) were included in the analyses up until exclusion. Data from participants who withdrew were included in the analyses up until withdrawal.

Tumour type (invasive or ductal carcinoma in situ), tumour stage, lymph node status, Bloom-Richardson grade, oestrogen receptor status, progesterone status, HER2 status, and ductal carcinoma in situ grade were compared between the screening groups with two-sided Fisher's exact tests. Age at detection, tumour size, and ductal carcinoma in situ size were compared with Mann-Whitney U tests.

Numbers of cancers were calculated per 1000 screening rounds or woman-years at risk and compared with exact rate ratio tests, assuming Poisson counts. Corresponding 95% CIs were calculated with a Poisson distribution. Woman-years at risk were calculated from the date of first screening examination to the date of discontinuation from the study, bilateral prophylactic mastectomy, detection of invasive cancer, death, or to the date at which the participant reached 60 years of age. To account for interval cancers, woman-years at risk were also calculated from the date of first screening examination to 1 year after the last screening visit; this calculation included women lost to follow-up after a screening visit. We defined interval cancers as cancers diagnosed between two screening rounds because of symptoms when the result of the previous screening round was negative. We retrieved data on interval cancers from the Dutch national pathology registry (PALGA) between Jan 17, 2017, and Jan 1, 2019.

We used the exact rate ratio test to compare biopsy frequency and false positive frequency between screening groups. Sensitivity was defined as the number of screen-detected breast cancers divided by the total number of breast cancers. Specificity was defined as the number of negative screens divided by all screens in women without breast cancer. Positive predictive value was calculated by dividing the number of screen-detected cancers by the number of positive screening tests (BI-RADS ≥3). Positive predictive value for biopsy was calculated by dividing the number of breast cancers by the number of biopsies. To compare sensitivity, specificity, and positive predictive value between the screening groups, we used Fisher's exact test, and CIs were calculated with the Clopper-Pearson interval. We repeated these analyses including incident screens only (all screens after the first screening round).

Analyses were also stratified by mammographic density (with BI-RADS breast composition categories A–D). These analyses included both randomly allocated and registration participants; participants from the MRI registration group were combined with the MRI group and participants from the mammography registration group were combined with the mammography group.

To test for linear trends in numbers of breast cancers, interval cancers, or false-positive results, or in tumour stage, sensitivity, or specificity when stratified by both BI-RADS breast density and automated breast (Volpara) density, we used linear-by-linear association tests.

To determine the level of agreement between the automated density measures and BI-RADS density estimates by the radiologists, Cohen's Kappa coefficient (κ) was

calculated. A post-hoc analysis of tumour stage, lymph node status, and specificity stratified by age (<50 years vs ≥50 years) was done per screening group.

Statistical analyses were done with IBM SPSS Statistics (version 24) and RStudio (version 1.0.44). P-values less than or equal to 0.05 were deemed significant. No independent data monitoring committee oversaw the study. This trial is registered with the Netherlands Trial Register, number NL2661.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

Inclusion and follow-up of participants took place between Jan 1, 2011, and Dec 30, 2017. 1355 women provided consent for randomisation and 231 for registration (**figure 1**). 675 of 1355 women who provided consent for randomisation were allocated to the MRI group, and 680 to the mammography group. One woman in the MRI group had a breast cancer diagnosis after randomisation, but before the first MRI screening, and was excluded from the analysis; therefore, 1354 women were included in the intention-to-treat population. She was excluded from the analysis because an invasive cancer before the first screening round was an exclusion criterion; furthermore, the cancer was neither screen-detected nor an interval cancer, as the first screening round had not been completed. 13 women were excluded after randomisation because they ultimately proved to have a cumulative lifetime risk of breast cancer below 20%, twelve were excluded because they were found to carry a BRCA1/2 mutation, and one was excluded because she was found to carry a TP53 mutation (**figure 1**); these participants were included in the analysis. 218 of 231 women opting to be in a registration group were in the mammography registration group and 13 were in the MRI registration group. The mean number of screening rounds per woman was 4.3 (SD 1.76), and the median follow-up after inclusion was 5.2 years for both screening groups (IQR 3.4–6.2 in the MRI group and 3.6–6.3 in the mammography group). Of the women who were randomly allocated, 57 requested the screening protocol of the other group during follow-up (45 [7%] of 675 in the MRI group and 13 [2%] of 680 in the mammography group); these participants were still analysed by intention to treat. 13 of the 45 requests in the MRI group were because of gadolinium retention information we sent to all participants in the MRI group in 2016. Before the end of follow-up, 234 women withdrew from the study (107

Table 1. Baseline characteristics of study population

	MRI-group (n=674)	Mammography-group (n=680)
Mean age (years ± SD)	44.7 ± 6.3	44.7 ± 6.3
Menopausal status		
Premenopausal	512 (76%)	505 (74%)
Postmenopausal	109 (16%)	116 (17%)
Unknown	53 (8%)	59 (9%)
Hormonal contraceptive use		
Now	103 (15%)	111 (16%)
In the past	462 (69%)	442 (65%)
Never	55 (8%)	50 (7%)
Unknown	54 (8%)	77 (11%)
Hormone replacement therapy use		
Now	7 (1%)	10 (2%)
In the past	14 (2%)	12 (2%)
Never	593 (88%)	577 (85%)
Unknown	60 (9%)	81 (12%)
Previous screening		
No screening	58 (9%)	53 (8%)
Unknown	13 (2%)	21 (3%)
Mammography		
Up to 2 years ago	535 (79%)	542 (80%)
More than 2 years ago	23 (3%)	29 (4%)
Unknown	14 (2%)	7 (1%)
MRI		
Up to 2 years ago	62 (9%)	81 (12%)
More than 2 years ago	90 (13%)	89 (13%)
Unknown	1 (0%)	1 (0%)
BI-RADS density category*		
A (entirely fat)	88 (13%)	92 (14%)
B (scattered densities)	248 (37%)	229 (34%)
C (heterogeneously dense)	237 (35%)	243 (36%)
D (extremely dense)	98 (15%)	102 (15%)
Unknown	3 (0%)	14 (2%)
No. of first-degree relatives with a history of breast cancer below the age of 50		
1	362 (54%)	397 (58%)
2	44 (7%)	37 (5%)
3 or more	2 (0%)	2 (0%)

Data are mean (SD) or n (%). BI-RADS=breast imaging reporting and data system. *As determined by the radiologist

[16%] of 675 in the MRI group; 127 [19%] of 680 in the mammography group). **Table 1** shows the characteristics of the participants by screening group (see appendix for baseline characteristics of participants in registration groups). MRIs and mammograms were mostly done on the same day, in incident rounds with a median 1 day (IQR 0–14) and a mean 12.8 days (SD 26.6) between the MRI and the mammogram.

55 cancers were detected in the randomly allocated participants (32 invasive cancers, 23 ductal carcinomas in situ; **table 2**). No bilateral breast cancers were detected and none had metastasised. Two triple-negative cancers were detected in the mammography group and one was detected in the MRI group. Invasive cancers in the MRI group were smaller than those in the mammography-group (median size 9 mm [5–14] in the MRI group vs 17 mm [13–22] in the mammography group; p=0.010; **table 2**). 14 (58%) of 24 invasive cancers were up to 10 mm in size in the MRI group, compared with only one (13%) of eight in the mammography group. Fewer invasive breast cancers were node positive in the MRI group than in the mammography group (four [17%] of 24 vs five [63%] of eight; p=0.023; **table 2**). Tumour stage of all incident cancers was significantly different between groups (p=0.035; **table 2**). In incident rounds, MRI screening resulted in lower numbers of late-stage cancers (one [6%] of 18 were ≥T2 in the MRI group vs two [25%] of eight in the mammography group; p=0.035 for stage difference) and node-positive cancers (two [11%] of 18 vs five [63%] of eight; p=0.014). Bloom-Richardson grade, oestrogen receptor status, progesterone receptor status, and HER2 status, ductal carcinoma in situ grade, and ductal carcinoma in situ size were not significantly different between screening groups (**table 2**). Descriptions of detected cancers including participants in the registration groups are shown in the appendix.

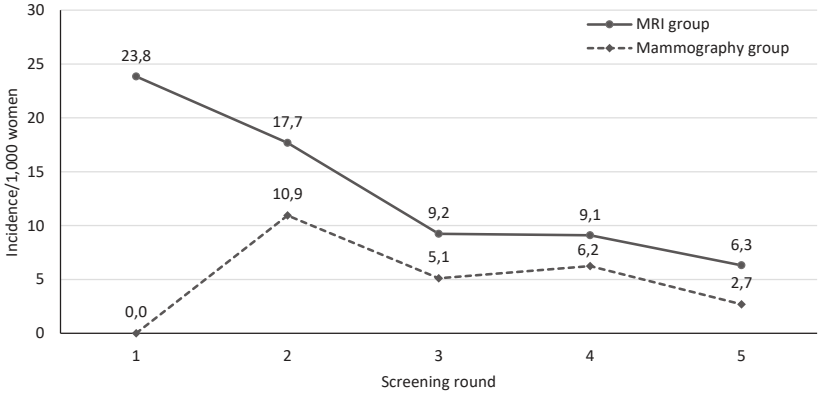


Figure 2. Breast cancer incidence per screening round by study group

Table 2. Characteristics of detected breast cancers, according to study group

	MRI group (n=674)	**Mammography group (n=680)**	***P* value**
Mean age at detection (years)	49.4 (7.1)	50.0 (4.6)	0.88
Cancers diagnosed during study			0.0017
No cancer	634 (94%)	665 (98%)	
Invasive breast cancer	24 (4%)	8 (1%)	
Ductal carcinoma in situ	16 (2%)	7* (1%)	
Size of invasive cancer (mm)	Mean: 11.9 (12.3) Median: 9 (5-14)	Mean: 18.0 (8.1) Median: 17 (13-22)	0.010
T stage			0.065[†]
T1a	7/24 (29%)	0	
T1b	7/24 (29%)	1/8 (13%)	
T1c	7/24 (29%)	5/8 (63%)	
T2	2/24 (8%)	2/8 (25%)	
T3	1/24 (4%)	0	
T4	0	0	
Node status			0.023
Positive	4/24 (17%)	5/8 (63%)	
Negative	20/24 (83%)	3/8 (38%)	
Bloom-Richardson grade			0.50
1	10/24 (42%)	2/8 (25%)	
2	9/24 (38%)	3/8 (38%)	
3	4/24 (17%)	3/8 (38%)	
Missing	1/24 (4%)	0	
Oestrogen-receptor status			0.25
Positive	22/24 (92%)	6/8 (75%)	
Negative	2/24 (8%)	2/8 (38%)	
Progesterone-receptor status			0.65
Positive	18/24 (75%)	5/8 (63%)	
Negative	6/24 (25%)	3/8 (38%)	
HER2 positive			1.00
Positive	2/24 (8%)	0	
Negative	22/24 (92%)	8/8 (100%)	
Ductal carcinoma in situ grade			1.00
1	5/16 (31%)	2/7 (29%)	
2	8/16 (50%)	4/7 (57%)	
3	3/16 (19%)	1/7 (14%)	
Ductal carcinoma in situ size (mm)	Mean: 34.18 (43.8) Median: 14 (9-35)[‡]	Mean: 30.29 (26.9) Median: 20 (7-60)	1.00
T stage incident rounds			0.035
Tis	7/25 (28%)	7/15 (47%)	
T1a+T1b	12/25 (48%)	1/15 (7%)	
T1c	5/25 (20%)	5/15 (33%)	

	MRI group (n=674)	Mammography group (n=680)	*P* value
T2 or higher	1/25 (4%)	2/15 (13%)	
Node status incident rounds			0.014
Positive	2/18 (11%)	5/8 (63%)	
Negative	16/18 (89%)	3/8 (38%)	
T stage <50 years			0.13
Tis	7/18 (39%)	5/8 (63%)	
T1a+T1b	6/18 (33%)	0	
T1c	4/18 (22%)	1/8 (13%)	
T2 or higher	1/18 (6%)	2/8 (25%)	
T stage ≥50 years			0.18
Tis	9/22 (41%)	2/7 (29%)	
T1a+b	8/22 (36%)	1/7 (14%)	
T1c	3/22 (14%)	4/7 (57%)	
T2 or higher	2/22 (9%)	0	
Node status, <50 years			0.011
Positive	1/11 (9%)	3/3 (100%)	
Negative	10/11 (91%)	0	
Node status ≥50 years			0.58
Positive	3/13 (23%)	2/5 (40%)	
Negative	10/13 (77%)	3/5 (60%)	

Data are mean (SD), n (%), median (IQR), or n/N (%). Tis=tumour in situ. *One ductal carcinoma in situ was detected after the woman requested screening with MRI and mammography. The ductal carcinoma in situ was detected by both MRI and mammography. †Based only on categories T1a +T1b, T1c, and T2 or higher. ‡Contains missing values. Tumour stage in MRI group (<50 years vs ≥50 years): p=0.93. Tumour stage in mammography group (<50 years vs ≥50 years): p=0.092. Node status in MRI group (<50 years vs ≥50 years): p=0.60. Node status in mammography group (<50 years vs ≥50 years): p=0.20.

Table 3 shows the performance of the two screening strategies within the randomisation groups. The number of breast cancers per 1000 screening rounds was significantly higher in the MRI group than in the mammography group (14.2 [95% CI 10.0–18.8] vs 4.9 [2.6–7.5]; p<0.00030). This difference in breast cancer incidence decreased and was no longer significant in the incident screening rounds (p=0.722). **Figure 2** shows the incidence per group per screening round. One (3%) of the 40 cancers in the MRI group and two (13%) of the 15 cancers in the mammography group were interval cancers (**table 3**). The interval cancer in the MRI group occurred 10 months after screening (T2, node positive, BI-RADS density D). One interval cancer in the mammography group occurred 9 months after screening (T2, node positive, BI-RADS density C), and the second occurred in the year after closure of the study (T1c, node negative, BI-RADS density B).

Sensitivity was higher in the MRI group than in the mammography group, but this difference was not significant (**table 3**). Specificity was significantly lower in the MRI

Table 3. Performance of the two screening strategies within the study groups

	MRI group (N=674)	Mammography group (N=680)	*P* value
Screening rounds	2812	3075	
Person-years at risk	3220	3326	
Screen-detected cancers	39*/40 (98%)	13/15 (87%)†	0.18
Interval breast cancers	1/40 (3%)	2/15 (13%)	
Breast cancers (per 1000 screening rounds)			
All breast cancers	14.2 (10.0-18.8)	4.9 (2.6-7.5)	0.00030
Screen-detected cancers	13.9 (9.6-18.5)*	4.2 (2.0-6.8)†	0.00012
Invasive screen-detected cancers	8.2 (5.0-11.7)	2.0 (0.7-3.6)	0.0010
Ductal carcinoma in situ	5.7 (3.2-8.5)*	2.3 (0.7-4.3)†	0.058
Interval cancers (per 1000 woman-years at risk)	0.3 (0.0-0.9)	0.6 (0.0-1.5)	1.000
Invasive cancers detection technique			
Mammography	3/23 (13%)	6/6 (100%)	
MRI	14/23 (61%)	NA	
Mammography and MRI	5/23 (22%)	NA	
Clinical breast examination only	1/23 (4%)	0	
Biopsies			<0.0001
N	149	54	
Incidence (per 1000 screening rounds)	53.0	17.6	
False positives (≥BI-RADS 3)			<0.001
N	449	276	
Incidence (per 1000 screening rounds)	159.7	89.8	
By mammography	98/449 (22%)	157/276 (57%)	
By MRI	275/449 (61%)	9/276 (3%)‡	
By both mammography and MRI	19/449 (4%)	0	
By clinical breast examination only	57/449 (13%)	110/276 (40%)	
Sensitivity (95% CI)	97.5% (86.8-99.9)	86.7% (59.5-98.3)	0.18
Specificity (95% CI)	83.8% (82.4-85.2)	91.0% (89.9-92.0)	<0.0001
Positive predictive value BI-RADS ≥ 3	8.0% (5.7-10.7)	4.5% (2.4-7.6)	0.074
Positive predictive value for biopsy	26.8% (20.0-34.7)	27.8% (16.5-41.6)	1.000
Incident screening rounds	2141	2407	
Breast cancers in incident rounds (per 1000 screening rounds)	10.0 (6.4-14.0)	5.9 (3.2-9.1)	0.72
Screen-detected in incident rounds	25/25 (100%)*	13/15 (87%)	
Interval cancers in incident rounds	0	2/15 (13%)	0.14
Biopsies in incident rounds			<0.0001
N	82	38	
Per 1000 screening rounds	38.3	15.8	
False positives in incident rounds			<0.0001
N	266	176	

	MRI group (N=674)	Mammography group (N=680)	*P* value
Per 1000 screening rounds	124.2	73.1	
Sensitivity in incident rounds	100% (86.3-100.0)	86.7% (59.51-98.3)	0.14
Specificity in incident rounds	87.4% (85.9-88.8)	92.6% (91.5-93.7)	<0.0001
Positive predictive value BI-RADS≥3 in incident rounds	8.6% (5.6-12.4)	6.9% (3.7-11.5)	0.61
Positive predictive value (biopsy) in incident rounds	30.5% (20.8-41.6)	39.5% (24.0-56.6)	0.54
Specificity <50 years	81.9% (80.1-83.6)	89.6% (88.2-90.9)	<0.0001
Specificity ≥50 years	87.7% (85.4-89.8)	93.5% (91.9-94.9)	<0.0001

Data are n, n/N (%), n (95% CI), or n% (95% CI). BI-RADS=breast imaging reporting and data system.

*One ductal carcinoma in situ was detected after the woman discontinued the trial protocol and went to the national breast cancer screening programme. Within the trial, this lesion was given a BI-RADS score of 3 and was considered stable over time. During the first screening at the national screening programme, this lesion was given a BI-RADS score of 4, and ultimately was diagnosed as ductal carcinoma in situ.

†One ductal carcinoma in situ was detected after the woman requested screening with MRI and mammography. The ductal carcinoma in situ was detected by both MRI and mammography.

‡These false positives occurred in women who requested the MRI protocol while being assigned to the mammography group.

Specificity in MRI group (<50 years vs ≥50 years): p=0.00010. Specificity in mammography group (<50 years vs ≥50 years): p=0.00026. Positive predictive value in MRI group (<50 years vs ≥50 years): p<0.0001 . Positive predictive value in mammography group (<50 years vs ≥50 years): p=0.107.

group than in the mammography group, although specificity improved for both groups in the incident screening rounds (**table 3**). 14 (61%) of the 23 invasive screen-detected cancers in the MRI group were detected by MRI only: eight were T1a/T1b, five were T1c (two of which were node positive), and one was T2. Three (13%) were detected by mammography only: one was T1a, two were T1b. Five (31%) of the 16 ductal carcinomas in situ detected in the MRI group were grade 1: one was detected by MRI only, one by mammography only, and three by both MRI and mammography.

Of the 449 false-positive results in the MRI group, 98 (22%) resulted from a positive mammogram while MRI was negative. Positive predictive value (for BI-RADS ≥3) was higher in the MRI group than in the mammography group, but this difference was not significant, whereas positive predictive value for biopsy was similar between screening groups (**table 3**), both for prevalent and incidence screens.

When we combined registration group data and randomisation group data (appendix), numbers of detected breast cancers increased significantly with increasing breast density for the mammography protocol (p=0.018) but not for the MRI protocol (p=0.92; **table 4**). All seven tumours stage T2 or higher and three of five interval cancers were in the two highest density categories (**table 4**). Automated breast density measurements were available for 537 (78%) of 687 participants in the MRI registration and MRI groups and for 733 (82%) of 898 participants in the mammography registration and mammography groups and were in slight agreement with the density assessments by radiologists, with a κ of 0.205.[21] However, results stratified by automated density grades

(appendix) were in accordance with those of BI-RADS breast density stratification.[19] According to estimates made by the radiologists, MRI detected more early-stage cancers, and more cancers with negative nodes than mammography in the three lowest breast density categories (A–C), in which MRI performs best (**table 4**). However, with automated (Volpara) breast density measurements, MRI detected more early-stage cancers in all density categories. Sensitivity did not differ significantly with increasing breast density in either protocol (**table 4**). Specificity decreased with increasing breast density for both screening protocols, as false positives increased (**table 4**).

In a post-hoc analysis, stratifying our results by age (<50 years vs ≥50 years), we observed no difference in tumour or nodal stage in either screening group (**table 2**), but higher specificity in both screening groups for women aged 50 years or older than for women younger than 50 years (**table 3**).

Median follow-up of patients after a breast cancer diagnosis was 4.3 years (IQR 3–6) and none of the patients in the randomisation groups died during follow-up, but one patient in the mammography registration group died due to breast cancer.

Table 4. All breast cancers, tumour staging, and false positives stratified by BI-RADS breast density category

	BI-RADS breast density				
	A	**B**	**C**	**D**	***P* value**
All participants in screening and registration groups*, N	206	549	562	239	
Screening rounds	993	2412	2413	1022	
All breast cancers, N	5	22	27[†]	11	
Incidence (per 1000 screening rounds)	5.0	9.1	11.2	10.8	0.13
Interval cancers, N	0	2	2	1	
Incidence (per 1000 screening rounds)	0	0.8	0.8	1.0	0.47
T stage					
Tis	1 (20%)	8 (36%)	11 (41%)	5 (50%)	0.11
T1a + T1b	2 (40%)	5 (23%)	8 (30%)	1 (9%)	0.98
T1c	2 (40%)	9 (41%)	4 (15%)	2 (18%)	0.49
T2+	0	0	4 (15%)	3 (27%)	0.0077
Node status					
Positive	1 (25%)	3 (21%)	6 (38%)	3 (50%)	
Negative	3 (75%)	11 (79%)	10 (63%)	3 (50%)	
Participants in the MRI group and MRI registration group, N	86	249	238	105	
Screening rounds	403	1033	973	440	
All breast cancers, N	5	15	17	5	
Incidence (per 1000 screening rounds)	12.4	14.5	17.5	11.4	0.92
Interval cancers, N	0	0	0	1	

	BI-RADS breast density				
	A	**B**	**C**	**D**	***P* value**
Incidence (per 1000 screening rounds)	0	0	0	2.3	0.10
Sensitivity	100.0% (47.8-10.0)	100.0% (78.2-100.0)	100.0% (80.5-100.0)	80.0% (28.4-99.5)	0.080
Specificity	90.5% (87.1-93.2)	85.3% (82.9-87.4)	82.8% (80.3-85.2)	77.0% (72.8-80.9)	<0.0001
False positives, N	38	150	164	100	
Incidence (per 1000 screening rounds)	94.3	145.2	168.6	227.3	<0.0001
T stage					
Tis	1 (20%)	7 (47%)	7 (41%)	1 (20%)	0.97
T1a + T1b	2 (40%)	5 (33%)	7 (41%)	0	0.54
T1c	2 (40%)	3 (20%)	2 (12%)	1 (20%)	0.42
T2+	0	0	1 (6%)	3 (60%)	0.0068
Node status					
Positive	1 (25%)	0	2 (20%)	2 (50%)	
Negative	3 (75%)	8 (100.0%)	8 (80%)	2 (50%)	
Participants in the mammography group and mammography registration group, N	120	300	324	134	
Screening rounds	590	1379	1440	582	
All breast cancers, N	0	7	10†	6	
Incidence (per 1000 screening rounds)	0	5.1	6.9	10.3	0.018
Interval cancers, N	0	2	2	0	
Incidence (per 1000 screening rounds)	0	1.5	1.4	0	0.99
Sensitivity	NA	71.4% (29.0-96.3)	80.0% (44.4-97.5)	100.0% (54.1-100.0)	0.18
Specificity	93.7% (91.5-95.5)	93.0% (92.3-93.6)	89.0% (87.3-90.6)	86.3% (83.2-90.0)	<0.00015
False positives, N	37	96	157	79	
Incidence (per 1000 screening rounds)	62.7	69.6	109.0	135.7	<0.0001
T stage					
Tis	0	1 (14%)	4 (40%)	4 (67%)	0.0065
T1a + T1b	0	0	1 (10%)	1 (17%)	0.12
T1c	0	6 (86%)	2 (20%)	1 (17%)	0.84
T2+	0	0	3 (30%)	0	0.35
Node status					
Positive	0	3 (50%)	4 (67%)	1 (50%)	
Negative	0	3 (50%)	2 (33%)	1 (50%)	

Data are n, n (%), or n% (95% CI). BI-RADS breast density was estimated at baseline. Tis=tumour in situ. BI-RADS=breast imaging reporting and data system.

*30 participants did not have density measurements.

†One ductal carcinoma in situ was detected after the woman requested screening with MRI and mammography. The ductal carcinoma in situ was detected by both MRI and mammography

DISCUSSION

To our knowledge, this randomised controlled trial is the first to compare MRI screening with mammography in women with familial risk of breast cancer. The number of cancers detected was significantly higher in the MRI group than in the mammography group. However, in incident rounds, this difference in breast cancer incidence was no longer significant. Median tumour size of the invasive breast cancers was smaller, and fewer invasive cancers were node positive in the MRI group than in the mammography group. To assess the effectiveness of the screening protocols in the long run, the results of the incident rounds are important: MRI screening resulted in lower numbers of late-stage cancers and node-positive cancers than mammography screening, thus both tumour stage and nodal status were significantly more favourable in the MRI group.

MRI screening might lead to a substantial reduction in mortality compared with mammography screening. In a study[4] of 93 569 patients diagnosed with primary breast cancer in the Netherlands in 2006–2012, 5-year relative survival for patients with T1c tumours was 98%, and for patients with T2 tumours it was 92%; for patients with N0 tumours 5-year relative survival was 98%, but for patients with N2 tumours, it was 86% despite up-to-date adjuvant therapy. Furthermore, with substantially fewer patients with node-positive disease, less adjuvant chemotherapy will be needed, sparing many women the early and late side-effects and cost. As prespecified, we intend to publish 10-year mortality results after linkage with our national database, since mortality reduction is the aim of screening. The current follow-up is too short, but tumour stage is a reliable proxy.[3,4]

Participants screened using the MRI protocol had a favourable shift in tumour stage in detected tumours compared with participants screened with the mammography protocol; tumour stages detected in the mammography group were very similar to stages detected by MRI screening in older multicentre studies in women with familial risk.[5–7] This result shows how both mammography and MRI have improved over the past decade, and suggests that we should no longer use the results of those older studies[6] for screening guidelines, as already shown by Obdeijn and colleagues.[22]

The MRI protocol had a clear disadvantage of more false-positive results and lower specificity, as was expected. Despite improvements in incident rounds for both screening groups, the difference in specificity between the groups remained significant and substantial. These results are in accordance with other high-risk MRI screening studies.[6] The false-positive frequency is potentially explained by the relatively young age of our participants, as the frequency of false positives clearly decreased in women aged 50 years or older and increased with increasing breast density, and might furthermore be the consequence of the very early stage at detection. The ACR expects a positive predictive value of 24% for biopsies, and the positive predictive value for biopsies we

found in this study was just above this.[19] Our positive predictive values for BI-RADS scores of 3 or higher and for biopsies are also similar to the positive predictive values found in two recent cohort studies of MRI screening.[11,23]

Another drawback of screening is overdiagnosis. The incidence of all cancers detected was higher in the MRI group than in the mammography group. The difference in cancer incidence decreased after the first screening round and was no longer significant, although incidence remained higher in the MRI group until the fourth round of screening. However, with a mean age of 49 years at detection (the average Dutch life expectancy is 84 years), hardly any of the invasive cancers are expected to be a result of overdiagnosis, as even early stage oestrogen-positive breast cancers might metastasise within 20 years.[3] Nevertheless, the detection of ductal carcinoma in situ is expected to be partly overdiagnosis; especially ductal carcinoma in situ grade 1, for which trials are ongoing to investigate whether active surveillance is safe.[24,25] In incident rounds, we detected an equal number of ductal carcinoma in situ with MRI and mammography.

A possible unwanted side-effect of MRI screening is retention of minute amounts of gadolinium in the brain and other tissues, although no harmful effect has been identified with the macrocyclic gadolinium products used in our 12 hospitals so far. A letter we sent to all participants in the MRI group in 2016 with evidence[26] that gadolinium could be retained in tissues, did not lead to a substantial withdrawal of participants.

Whether breast density measurements were estimated by radiologists or fully automated with Volpara, all tumours stage T2 or higher and most of the interval cancers were only detected in the two highest breast density categories. MRI proved however, to be capable of detecting relevantly more early-stage cancers especially at categories A–C, and according to our automated density assessment, also in the highest-density category D, resulting in fewer late-stage cancers in incident rounds. With increasing breast density, specificity decreased both in the MRI and mammography groups, consistent with the results of Kerlikowske and colleagues.[12] Based on our results, density seems to be more important than age when choosing a screening strategy.

Previous studies have concluded that mammography is of limited additional value to MRI screening in women with familial risk.[27] Of the false-positive results in the MRI group, 98 (22%) of 449 were caused by mammography only. On the other hand, three (12%) minimal cancers (≤1 cm) were only detected by our low-frequency mammography. We do not know at which stage these minimal tumours would have been detected with MRI, but either an even lower frequency of mammography screening should be considered, or mammography should be omitted entirely. Clinical breast examination generated a substantial amount of false-positive results in both screening groups and detected only one cancer, making the additional value of clinical breast examination negligible.

Studies have shown that digital breast tomosynthesis has the potential to increase screening sensitivity and specificity in comparison with digital mammography.[28] If digital

breast tomosynthesis would have replaced mammography in this study, we would expect a gain in diagnostic accuracy in the mammography group. However, the average additional cancer yield for digital breast tomography is 1.2 per 1000 cases, compared with 3.5–4.4 per 1000 cases for ultrasound (with a considerable accompanying increase in false-positive frequency; thus, ultrasound is not recommended in Dutch screening guidelines)[9,29] but 15.5 per 1000 cases for MRI.[28,30,31]

One limitation of our study was that the numbers of detected cancers were small when stratified according to density or age categories, because the study was powered to show a difference in tumour size between the two screening groups. Therefore, we did not see a significant decrease in sensitivity with increasing breast density, which has been shown in previous studies.[12] Importantly, in the MRI group, the number of late-stage cancers detected decreased in incident rounds, but we also have to evaluate long-term survival and cost-effectiveness.

Another limitation was that previous screening might have affected cancer incidence; however, a nearly equal amount of previous MRI screening was done in both screening groups. Previous screening might have affected cancer incidence more in the mammography group because there was more previous mammography screening in the study population, although it was also distributed equally in both screening groups. Fortunately, the study continued for 7 years, with an average of 4.3 screening rounds per person. The highest cancer incidence in the mammography group was in the second year, and a nearly equally steep decline in cancer incidence can be seen in both groups thereafter. This suggests that the effect of previous screening on the complete study must have been limited and will not have affected our primary endpoints, tumour size and nodal status.

The biggest strength of our study (aside from the randomised design) is that the results are representative of daily, real-life practice, as patients were not only included at university hospitals with specialised high-risk breast screening units, but also at five larger, general hospitals throughout the Netherlands. However, better specificity might be achievable if MRIs are done in expert clinics. Further improvements might come from abbreviated MRI and, for specificity, artificial intelligence-based assistance.

We conclude that in real-life practice, MRI screening can result in an important and favourable shift in tumour stage at time of breast cancer detection compared with mammography screening, reducing the incidence of late-stage cancers and thus reducing the need for adjuvant chemotherapy and the risk of mortality. Especially in breast density category D, MRI screening would come at the cost of lower specificity. Clinical breast examinations could be omitted, and the frequency of combined MRI and mammography screenings could be further reduced.

Acknowledgments

This trial was funded by the Dutch Government ZonMw (200320002), the Dutch Cancer Society (DDHK 2009-4491), A Sister's Hope, Pink Ribbon, Stichting Coolsingel, and J&T Rijke Stichting. We thank all women who participated in this study. Furthermore, we thank Ada van Eekelen, Lydia Ruiter, Lenny Polman, Suzanne Gerretsen, Ypie Bruining, Aukje Postma, and Christel Haekens especially for their dedicated work.

REFERENCES

1. Margolin S, Johansson H, Rutqvist LE, Lindblom A, Fornander T. Family history, and impact on clinical presentation and prognosis, in a population-based breast cancer cohort from the Stockholm County. *Fam Cancer*. 2006;5:309–21.
2. Brandt A, Bermejo JL, Sundquist J, Hemminki K. Age of onset in familial breast cancer as background data for medical surveillance. *BR J Cancer*. 2010;102:42–47.
3. Pan H, Gray R, Braybrooke J, et al. 20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years. *N Engl J Med*. 2017;377:1836–46.
4. Saadatmand S, Bretveld R, Siesling S, Tilanus-Linthorst MMA. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients. *BMJ*. 2015;351:h4901.
5. Kriege M, Brekelmans CTM, Boetes C, et al. Efficacy of MRI and mammography for breast-cancer screening in women with familial or genetic predisposition. *N Engl J Med*. 2004;351:427–37.
6. Phi XA, Houssami N, Hooning MJ, et al. Accuracy of screening in women at familial risk of breast cancer without a known gene mutation: individual patient data meta-analysis. *Eur J Cancer*. 2017;85:31–38.
7. Leach MO, Boggis CR, Dixon AK, et al. Screening with magnetic resonance imaging and mammography of a UK population at high risk of breast cancer: a prospective multicenter cohort study (MARIBS). *Lancet*. 2005;365: 1769–78.
8. Smith RA, Andrews KS, Brooks D, et al. Cancer screening in the United States, 2017: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin*. 2017;67:100–21.
9. Nationaal Borstkanker Overleg Nederland (NABON). Richtlijn borstkanker: versie 2.0. 2017; https://www.oncoline.nl/borstkanker. Accessed Sept 15, 2018.
10. National Institute for Health and Care Excellence (NICE). Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. 2017; https://www.nice.org.uk/guidance/cg164/chapter/Recommendations#surveillance-and-strategies-for-early-detection-of-breast-cancer. Accessed April 15, 2018.
11. Bick U, Engel C, Krug B, et al. High-risk breast cancer surveillance with MRI: 10-year experience from the German consortium for hereditary breast and ovarian cancer. *Breast Cancer Res Treat*. 2019;175:217-228
12. Kerlikowske K, Scott CG, Mahmoudzadeh AP, et al. Automated and clinical breast imaging reporting and data system density measures predict risk for screen-detected and interval cancers: a case-control study. *Ann Intern Med*. 2018;168:757–65.
13. Lehman CD, Gatsonis C, Kuhl CK, et al. MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer. *N Engl J Med*. 2007;356:1295–303.
14. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. *AJR Am J Roentgenol*. 2012;198:W292–95.
15. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, et al. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst*. 2013;105:1314–21.
16. Claus EB, Risch NR, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer*. 1994;73:643–51.

17. Saadatmand S, Rutgers EJ, Tollenaar RA, et al. Breast density as indicator for the use of mammography or MRI to screen women with familial risk for breast cancer (FaMRIsc): a multicenter randomized controlled trial. *BMC Cancer*. 2012;12:440.
18. Kuhl CK, Schrading S, Bieling HB, et al. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *Lancet*. 2007;370:485–92.
19. American College of Radiology. Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas), 4th edition. Reston, VA; 2003.
20. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. *IEEE Trans Med Imaging*. 2006;25:273–82.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
22. Obdeijn IM, Winter-Warnars GA, Mann RM, Hooning MJ, Hunink MG, Tilanus-Linthorst MMA. Should we screen BRCA1 mutation carriers only with MRI? A multicenter study. *Breast Cancer Res Treat*. 2014;144:577–82.
23. Lee JM, Ichikawa L, Valencia E, et al. Performance Benchmarks for screening breast MR imaging in community practice. *Radiology*. 2017;285:44–52.
24. Elshof LE, Tryfonidis K, Slaets L, et al. Feasibility of a prospective, randomized, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ – the LORD study. *Eur J Cancer*. 2015;51:1497–510.
25. Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen-detected DCIS; the LORIS trial. *Eur J Cancer*. 2015;51:2296–303.
26. Huckle JE, Altun E, Jay M, Semelka RC. Gadolinium deposition in humans: when did we learn that gadolinium was deposited in vivo? *Invest Radiol*. 2016;51:236–40.
27. Vreemann S, van Zelst JCM, Schlooz-Vries M, et al. The added value of mammography in different age-groups of women with and without BRCA mutation screened with breast MRI. *Breast Cancer Res*. 2018;20:84.
28. Zackrisson S, Lang K, Rosso A, et al. One-view breast tomosynthesis versus two-view mammography in the Malmo Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. *Lancet Oncol*. 2018;19:1493–503.
29. Berg WA, Blume JD, Cormack JB, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA*. 2008;299:2151–63.
30. Gilbert FJ, Tucker L, Young KC. Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool. *Clin Radiol*. 2016;71:141–50.
31. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental breast MR imaging screening of women with average risk of breast cancer. *Radiology*. 2017;283:361–70.

SUPPLEMENTARY APPENDIX

List of investigators

Writing committee

S Saadatmand (Erasmus University Medical Centre), HA Geuzinge (Erasmus University Medical Centre), IM Obdeijn (Erasmus University Medical Centre), MJ Hooning (Erasmus University Medical Centre), MMA Tilanus-Linthorst (Erasmus University Medical Centre), HJ de Koning (Erasmus University Medical Centre), C Verhoef (Erasmus University Medical Centre), EAM Heijnsdijk (Erasmus University Medical Centre), EJT Rutgers (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), R Mann MD (Radboud University Hospital), N Karssemeijer (Radboud University Hospital), DBW de Roy van Zuidewijn (Medical Centre Leeuwarden), HM Zonderland (Academic Medical Centre), RAEM Tollenaar (Leiden University Medical Centre), MBI Lobbes (Maastricht University Medical Center), MGEM Ausems (University Medical Centre, Utrecht), M van 't Riet (Reinier de Graaf Gasthuis), I Mares- Engelberts (Vlietland ziekenhuis), EJT Luiten (Amphia ziekenhuis), JC Oosterwijk (University Medical Centre Groningen)

Other co-authors of the FaMRIsc study group

CHM van Deurzen (Erasmus University Medical Centre), LB Koppert (Erasmus University Medical Centre), E Madsen (Erasmus University Medical Centre), J Rothbarth (Erasmus University Medical Centre), C de Monye (Erasmus University Medical Centre), MM van Rosmalen (Erasmus University Medical Centre), CE Loo (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), J Wesseling (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), J Remmelzwaal (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), M Schlooz-Vries (Radboud University Hospital), HBW Gort (Medical Centre Leeuwarden), R Roi-Antonides (Medical Centre Leeuwarden), S van der Meij (Amsterdam University Medical Centre), WE Mesker (Leiden University Medical Centre), MNJM Wasser (Leiden University Medical Centre), K Keymeulen (Academic Hospital, Maastricht), WB Veldhuis (University Medical Centre Utrecht), AJ Witkamp (University Medical Centre Utrecht), E van Druten (Reinier de Graaf Gasthuis), E Tetteroo (Amphia Ziekenhuis), C Contant (Maasstad ziekenhuis)

List of participating hospitals

Erasmus University Medical Center, Antoni van Leeuwenhoek – the Netherlands Cancer Institute, Maastricht University Medical Center, Leiden University Medical Center, Radboud university medical center, Academic Medical Center and University Medical Center Utrecht, Medical Centre Leeuwarden, Reinier de Graaf Hospital Delft, Amphia Hospital Breda, and Vlietland Hospital Vlaardingen/Rotterdam, Maasstad Hospital Rotterdam

	Randomised women	Women in the registration group	PI
Erasmus University Medical Center	535	150	Dr. M. Tilanus- Linthorst
Antoni van Leeuwenhoek - the Netherlands Cancer Institute	299	11	Prof E Rutgers
Medical Center Leeuwarden	97	17	Dr. D de Roy van Zuijdewijn
Leiden University Medical Center	89	15	Prof R Tollenaar
Reinier de Graaf hospital Delft	60	5	Dr. M van 't Riet
Radboud University Medical Center	55	13	Dr. M Schlooz-Vries
Maastricht University Medical Center	54	1	Dr. K Keymeulen
University Medical Center Utrecht	54	8	Dr. M Aussems
Amsterdam University Medical Center (AMC)	41	0	Dr. H Zonderland
Amphia Hospital Breda	31	0	Dr. E Luiten
Vlietland Hospital Vlaardingen/Rotterdam	30	11	Dr. I Mares-Engelberts
Maasstad Hospital Rotterdam	9	0	Dr. C Contant
Total	1354	231	

Footnote with Figure 1

The 22 other reasons not to be randomized were:

- 7 women did not receive info about the trial
- 3 women because of costs
- 4 women continued screening in the national breast cancer screening programme
- 8 women moved to a different city and/or continued screening in different hospital

Table S1. Characteristics of all participating women at baseline, according to study group

Characteristic	MRI group (N=674)	Mammography group (N=680)	Registration group (N=231)
Mean age (years ± SD)	44.7 ± 6.3	44.7 ± 6.3	45.2 ± 6.6
Menopausal status			
Premenopausal	512 (76%)	505 (74%)	148 (64%)
Postmenopausal	109 (16%)	116 (17%)	44 (19%)
Unknown	53 (8%)	59 (9%)	39 (17%)
Hormonal contraceptive use			
Now	103 (15%)	111 (16%)	32 (14%)
In the past	462 (69%)	442 (65%)	126 (55%)
Never	55 (8%)	50 (7%)	21 (9%)
Unknown	54 (8%)	77 (11%)	52 (23%)
Hormone replacement therapy use			
Now	7 (1%)	10 (2%)	3 (1%)
In the past	14 (2%)	12 (2%)	8 (4%)
Never	593 (88%)	577 (85%)	167 (72%)
Unknown	60 (9%)	81 (12%)	53 (23%)
Previous screening			
No screening	58 (9%)	53 (8%)	14 (6%)
Unknown	13 (2%)	21 (3%)	15 (7%)
Mammography			
≤ 2 years ago	535 (79%)	542 (80%)	172 (74%)
>2 years ago	23 (4%)	29 (4%)	19 (8%)
Unknown	14 (2%)	7 (1%)	4 (2%)
MRI			
≤ 2 years ago	62 (9%)	81 (12%)	22 (10%)
>2 years ago	90 (13%)	89 (13%)	29 (13%)
Unknown	1 (0%)	1 (0%)	51 (22%)
BI-RADS density category†			
A (entirely fat)	88 (13%)	92 (14%)	28 (12%)
B (scattered densities)	248 (37%)	229 (34%)	71 (31%)
C (heterogeneously dense)	237 (35%)	243 (36%)	84 (36%)
D (extremely dense)	98 (15%)	102 (15%)	41 (18%)
Unknown	3 (0%)	14 (2%)	8 (3%)
No. of first-degree relatives with a history of breast cancer below the age of 50			
1	362 (54%)	397 (58%)	113 (49%)
2	44 (7%)	37 (5%)	14 (6%)
≥3	2 (0%)	2 (0%)	1 (0%)

† Determined by radiologists, according to the fourth ACR BI-RADS (4th edition)

Table S2. Characteristics of detected breast cancers, according to the MRI protocol; and the Mammography protocol

	MRI group and MRI registration group (N=687)	**Mammography group and mammography registration group (N=898)**	***P* value**
Mean age at detection (years ± SD)	49.4 ± 7.0	49.3 ± 5.5	0.701
No cancer	645 (94%)	875 (97%)	
Invasive breast cancers	26 (4%)	14 (2%)	
Ductal carcinoma in situ	16 (2%)	9† (1%)	0.002
Size of invasive cancers (mm± SD) [mm IQR]	Mean: 12.9 ± 13.0 Median: 9 [5-14]	Mean: 17.9 ± 8.6 Median: [14-19]	0.004
T stage			
T1a	7 (27%)	0	
T1b	7 (27%)	2 (14%)	
T1c	8 (31%)	9 (64%)	
T2	3 (12%)	3 (23%)	
T3	1 (4%)	0	
T4	0	0	0.039‡
Node status			
Positive	5 (19%)	8 (57%)	
Negative	21 (81%)	6 (43%)	0.031
BR grade			
1	10 (39%)	4 (29%)	
2	10 (39%)	6 (43%)	
3	4 (15%)	4 (29%)	
Missing	2 (8%)	0	0.566
ER positive	24 (92%)	12 (86%)	0.602
PR positive	20 (77%)	11 (79%)	1.000
HER2 positive	2 (8%)	0	0.533
Ductal carcinoma in situ grade			
1	5 (31%)	2 (22%)	
2	8 (50%)	5 (56%)	
3	3 (19%)	2 (22%)	1.000
Ductal carcinoma in situ size (mm ± SD) [mm IQR]	Mean: 34.2 ± 43.8 Median: 14 [9-35] §	Mean: 42.2 ± 50.5 Median: 20 [8-60]	0.879
T stage in incident rounds			
Tis	7 (26%)	9 (45%)	
T1a+T1b	12 (44%)	2 (10%)	
T1c	6 (22%)	7 (35%)	
T2+	2 (7%)	2 (10%)	0.076
Node status in incident rounds			
Positive	3 (15%)	5 (46%)	
Negative	17 (85%)	6 (55%)	0.095

† One ductal carcinoma in situ was detected after the woman demanded screening with MRI and mammography. The ductal carcinoma in situ was detected by both MRI and mammography

‡ Based on categories 'T1a+T1b'; 'T1c' and 'T2+'

§ Contains missing values

Table S3. Screen-detected cancers, interval cancers, detection technique, biopsies, false positives, sensitivity, specificity, PPV and detection rates, according to study protocol

	MRI group and MRI registration group (N=687)	Mammography group and mammography registration group (N=898)	*P* value
Screening rounds	2861	3995	
Woman-years at risk	3271	4331	
Screen-detected cancers	41† (98%)	19‡ (83%)	
Interval cancers	1 (2%)	4 (2%)	0.049
No. of breast cancers per 1000 screening rounds (95% CI)			
All breast cancers	14.7† (10.5-19.2)	5.8‡ (3.5-8.3)	0.002
Screen-detected cancers	14.3† (10.1-18.9)	4.8‡ (2.8-7.0)	<0.001
Invasive	8.7 (5.6-12.2)	2.5 (1.0-4.3)	<0.001
Ductal carcinoma in situ	5.6† (3.1-8.4)	2.3‡ (1.0-3.8)	0.041
No. of interval cancers per 1000 women- years at risk (95% CI)	0.3 (0.0-0.9)	0.9 (0.2-1.8)	0.573
Detection technique of invasive cancers			
Mammography§	3 (12%)	10 (100%)	
MRI§	16 (64%)	n.a.	
Both mammography and MRI§	5 (20%)	n.a.	
Clinical breast examination only	1 (4%)	0	
Biopsies (rate¶)	153 (53.5)	72 (18.0)	<0.001
False positives (rate ¶)	458 (162.6)	369 (92.4)	<0.001
Sensitivity (95% CI)	97.6% (87.4-99.9)	82.6% (61.2-95.0)	0.049
Specificity (95% CI)	83.8% (82.3-85.1)	90.7% (89.8-91.6)	<0.001
Positive predictive value BI-RADS ≥ 3 (95% CI)	8.2% (6.0-11.0)	4.9% (3.0-7.5)	0.059
Positive predictive value for biopsy (95% CI)	28.1% (21.1-35.9)	31.9% (21.4-44.0)	0.638
No. of incident screening rounds	2177	3113	
No. of breast cancers per 1000 incident rounds (95% CI)	12.4 (7.8-17.5)	6.4 (3.9-9.3)	0.035
Screen-detected cancers in incident rounds	27† (100%)	18 (90%)	
Interval cancers in incident rounds	0	2 (10%)	0.176
Biopsies in incident rounds (rate ¶)	85 (39.0)	54 (17.3)	<0.001
False positives in incident rounds (rate¶)	268 (123.1)	249 (80.0)	<0.001
Sensitivity in incident rounds (95% CI)	100.0% (87.2-100.0)	90.0% (68.3-98.8)	0.176
Specificity in incident rounds (95% CI)	87.5% (86.1-88.9)	91.9% (90.9-92.9)	<0.001
Positive predictive value BI-RADS ≥ 3 in incident rounds (95% CI)	9.2% (6.1-13.0)	6.7% (4.0-10.4)	0.351
Positive predictive value for biopsy in incident rounds (95% CI)	31.8% (22.1-42.8)	37.0% (24.3-51.3)	0.716

† One ductal carcinoma in situ was detected after the woman discontinued the trial protocol and went to the national breast cancer screening program. Within the trial, this lesion was given a BI-RADS score 3, and was considered stable over time. At the moment the woman underwent her first screening at the national screening program, this lesion was given a BI-RADS score 4, and ultimately appeared to be ductal carcinoma in situ

‡ One ductal carcinoma in situ was detected after the woman demanded screening with MRI and mammography. The ductal carcinoma in situ was detected by both MRI and mammography

§ Possibly in combination with a positive clinical breast examination

¶ Rate per 1000 screening rounds

Table S4. All breast cancers, interval cancers, sensitivity, specificity, T staging, node status and false positives by Volpara Density Grade categories

	Volpara Density Grade†					
	VDG 1	**VDG 2**	**VDG 3**	**VDG 4**	**VDG unknown**	***p* trend**
All screening and registration groups	N=32	N=257	N=579	N=402	N=315	
Screening rounds	140	1240	2642	1787	1047	
All breast cancers (rate‡)	2 (14.3)	7 (5.6)	27§ (10.2)	18 (10.1)	11 (11.5)	0.431¶
Interval cancers (rate‡)	0	2 (1.6)	1 (0.4)	2 (1.1)	0	0.896¶
T stage						
Tis	0	3 (43%)	12§ (44%)	5 (28%)	5 (46%)	0.758¶
T1a + T1b	1 (50%)	1 (14%)	6 (22%)	6 (33%)	2 (18%)	0.422¶
T1c	1 (50%)	3 (43%)	8 (30%)	2 (11%)	3 (27%)	0.214¶
T2+	0	0	1 (4%)	5 (28%)	1 (9%)	0.014¶
Node status						
Positive	1 (50%)	1 (25%)	4 (27%)	5 (39%)	2 (33%)	
Negative	1 (50%)	3 (75%)	11 (73%)	8 (62%)	4 (67%)	
MRI group and MRI registration group	N=16	N=99	N=250	N=172	N=150	
Screening rounds	63	459	1086	748	505	
All breast cancers (rate‡)	2 (31.7)	4 (8.7)	18 (16.6)	12 (16.0)	6 (11.9)	0.746¶
Interval cancers (rate‡)	0	0	0	1 (1.3)	0	0.235¶
Sensitivity (95% CI)	100.0% (15.8-100.0)	100.0% (47.8-100.0)	100.0% (80.5-100.0)	91.7% (61.5-99.8)		0.265
Specificity (95% CI)	96.7% (88.7-99.6)	87.5% (84.1-90.4)	83.0% (80.6-85.2)	81.1% (78.1-83.9)		<0.001
False positives (rate‡)	2 (31.7)	57 (124.2)	182 (167.6)	139 (185.8)		<0.001
T stage						
Tis	0	3 (75%)	9 (50%)	3 (25%)	1 (17%)	0.732¶
T1a + T1b	1 (50%)	1 (25%)	6 (33%)	4 (33%)	2 (33%)	0.950¶
T1c	1 (50%)	0	3 (17%)	2 (17%)	2 (33%)	0.829¶
T2+	0	0	0	3 (25%)	1 (17%)	0.040¶
Node status						
Positive	1 (50%)	0	1 (11%)	2 (22%)	1 (20%)	
Negative	1 (50%)	1 (100%)	8 (89%)	7 (78%)	4 (80%)	
Mammography group and mammography registration group	N=16	N=158	N=329	N=230	N=165	
Screening rounds	77	781	1556	1039	542	
All breast cancers (rate‡)	0	3 (3.8)	9§ (5.8)	6 (5.8)	5 (9.2)	0.459¶
Interval cancers (rate‡)	0	2 (2.6)	1 (0.6)	1 (1.0)	0	0.475¶

	Volpara Density Grade†					
	VDG 1	**VDG 2**	**VDG 3**	**VDG 4**	**VDG unknown**	***p* trend**
Sensitivity (95% CI)	n.a.	33.3% (0.8-90.6)	88.9% (51.8-99.7)	83.3% (35.9-99.6)		0.169
Specificity (95% CI)	94.8% (87.2-98.6)	92.2% (90.0-94.5)	91.7% (90.2-93.1)	88.2% (86.1-90.1)		<0.001
False positives (rate‡)	4 (51.9)	61 (78.1)	128 (82.3)	122 (117.4)		0.001
T stage						
Tis	0	0	3§ (33%)	2 (33%)	4 (80%)	0.291¶
T1a + T1b	0	0	0	2 (33%)	0	0.080¶
T1c	0	3 (100%)	5 (56%)	0	1 (20%)	0.144¶
T2+	0	0	1 (11%)	2 (33%)	0	0.160¶
Node status						
Positive	0	1 (33%)	3 (50%)	3 (75%)	1 (100%)	
Negative	0	2 (67%)	3 (50%)	1 (25%)	0	

† Volpara Density measured at baseline; ‡ Rate per 1000 screening rounds; § One ductal carcinoma in situ was detected after the woman demanded screening with MRI and mammography. The ductal carcinoma in situ was detected by both MRI and mammography

¶ p-value of trend analysis on categories VDG1-VDG4

Chapter 3

Breast cancer screening strategies for women with ATM, CHEK2, and PALB2 pathogenic variants: a comparative modeling analysis

Lowry KP, Geuzinge HA, Stout NK, Alagoz O, Hampton J, Kerlikowske K, de Koning HJ, Miglioretti DL, van Ravesteyn NT, Schechter C, Sprague BL, Tosteson ANA, Trentham-Dietz A, Weaver D, Yaffe MJ, Yeh JM, Couch FJ, Hu C, Kraft P, Polley EC, Mandelblatt JS, Kurian AW, Robson ME, on behalf of the Breast Working Group of the Cancer Intervention and Surveillance Modeling Network (CISNET), Breast Cancer Surveillance Consortium (BCSC), and the Cancer Risk Estimates Related to Susceptibility consortium (CARRIERS).

JAMA Oncology, 2022, online ahead of print

ABSTRACT

Importance: Screening mammography and magnetic resonance imaging (MRI) are recommended for women with *ATM*, *CHEK2* and *PALB2* pathogenic variants (PVs). However, there are few data to guide screening regimens for these women.

Objective: To estimate the benefits and harms of screening strategies using mammography and MRI at various start ages.

Design: We used two established breast cancer microsimulation models from the Cancer Intervention and Surveillance Modeling Network (CISNET) to evaluate different screening strategies. Age-specific breast cancer risks were estimated using Cancer Risk Estimates Related to Susceptibility (CARRIERS) consortium data. Mammography and MRI screening performance were estimated from published literature.

Participants: U.S. women with *ATM, CHEK2* or *PALB2* PVs born in 1985 and followed from age 25 for their lifetimes.

Interventions: Screening strategies with combinations of annual mammography alone and with MRI starting at age 25, 30, 35, or 40 until age 74.

Main Outcomes and Measures: Lifetime breast cancer mortality reduction, life years gained (LYG), breast cancer deaths averted, total screens, false-positive screens, and benign biopsies per 1,000 women screened.

Results: Average model-estimated lifetime breast cancer risk was 21%, 28%, and 38% in *ATM*, *CHEK2* and *PALB2* PVs, respectively. Across PVs, annual mammography from ages 40-74 reduced breast cancer mortality by 36-39% compared to no screening. Screening with annual MRI starting at age 35 followed by annual mammography and MRI at 40 reduced breast cancer mortality by 54.4-57.6% with 4,661-5,001 false-positive screens and 1,280-1,368 benign biopsies per 1000 women. Annual MRI starting at age 30 followed by mammography and MRI at 40 reduced mortality by 55.4-59.5% with 5,075-5,415 false-positive screens and 1,439-1,528 benign biopsies per 1000 women. When starting MRI at age 30, initiating annual mammography starting at age 30 versus 40 did not meaningfully impact mortality (≤0.3%) but added 649-650 false-positive screens and 58-59 benign biopsies per 1000 women.

Conclusions and Relevance: In women with *ATM, CHEK2,* and *PALB2* PVs, annual MRI screening starting at age 30-35 followed by annual MRI and mammography at age 40

reduces breast cancer mortality by more than 50%. In the setting of MRI screening, mammography prior to age 40 offers little additional benefit.

INTRODUCTION

Genetic testing for breast cancer susceptibility has been an important aspect of cancer prevention since *BRCA1* and *BRCA2* (*BRCA1/2*) were identified in 1994-1995.[1,2] This discovery facilitated development of breast cancer screening and risk reduction guidelines for women with *BRCA1/2* pathogenic variants (PVs) and their relatives.[3,4] More recently a group of non-*BRCA1/2* PVs conferring moderate-to-high risk of breast cancer has been recognized, the most common of which are *ATM*, *CHEK2,* and *PALB2*. Each of these PVs increases breast cancer risk by at least two-fold and collectively they are identified in 2-3% of women diagnosed with breast cancer and approximately 1% of the population.[5,6]

Due to the increasing availability and affordability of multi-gene panel testing,[7,8] a growing number of women are learning they are carriers of these moderate-to-high risk PVs. The optimal approach to breast cancer screening in these women has not been established. Based on expert opinion and experience with MRI screening in women with *BRCA1/2* PVs, [9-11] the National Comprehensive Cancer Network (NCCN) recommends consideration of annual MRI in addition to mammography in *ATM* and *CHEK2* PV carriers starting at age 40 and in *PALB2* PV carriers at age 30.[12] Clinical trials comparing multiple approaches to breast cancer screening among women with each PV are not feasible given the prohibitively large sample sizes and follow-up time required.

In the absence of clinical trials, simulation modeling can be used to synthesize available data and compare screening strategies based on the projected impact on screening outcomes. Simulation models from the Cancer Intervention and Surveillance Modeling Network (CISNET) have previously informed cancer screening guidelines for the United States Preventive Services Task Force[13-15] and the American Cancer Society.[16,17] For this analysis, we adapted two CISNET breast cancer simulation models for women with *ATM*, *CHEK2*, and *PALB2* PVs using risk estimates from the Cancer Risk Estimates Related to Susceptibility (CARRIERS) consortium, the largest United States (US) consortium of familial- and population-based studies of breast cancer risk due to cancer susceptibility genes.[5] Using these models, we evaluated the benefits and harms of screening strategies using MRI and mammography to inform guideline recommendations for carriers of these moderate-to-high risk PVs.

METHODS

Model overview

The CISNET models used in this analysis were Model E (Erasmus Medical Center, Rotterdam, the Netherlands) and Model W-H (University of Wisconsin-Madison, Madison, Wisconsin; Harvard Medical School, Boston, Massachusetts). Full details regarding the development

and validation of these models have been described previously[18-21] and can be found at https://resources.cisnet.cancer.gov/registry. These models were independently developed using different structures, assumptions and methods to implement unobservable parameters for breast cancer natural history.[22] The use of two separate models therefore provides a plausible range of results given the inherent uncertainty of unobservable parameters related to breast cancer natural history. The models also share some common data elements including non-breast cancer mortality risk, screening performance and treatment effectiveness.[22] Both models have been previously validated and reproduce age-specific Surveillance Epidemiology and End Results (SEER) incidence and mortality in the US population.[23,24] This modeling analysis was determined not to be human participants research by the institutional review boards of Erasmus Medical Center, University of Wisconsin–Madison, and Harvard Medical School; therefore, this study was exempt from institutional review board approval and did not require informed consent.

The models simulate lifetime horizons of individual women and background US breast cancer incidence (including ductal carcinoma in situ (DCIS) and invasive breast cancer) in the absence of screening and treatment based on age-period-cohort models.[25] Breast cancer survival is dependent on age and tumor size and/or stage at diagnosis, estrogen receptor (ER) status and human epidermal growth factor 2 (HER2) status, and treatment effectiveness. When screening is applied, cancers can be diagnosed at earlier size/stage than with clinical detection, potentially reducing mortality. Women can die of breast cancer or non-cancer causes.

Population

We modeled US women with *ATM, CHEK2* or *PALB2* PVs born in 1985 (the youngest birth cohort for whom intensive breast cancer screening could potentially be recommended in 2010-2020) and followed from age 25 for their lifetimes.

Model input parameters

Input parameters for cancer risk and incidence, subtype, screening performance, and treatment effectiveness are summarized in **Table 1**.

Breast cancer risk

Parameters for breast cancer incidence and subtype for each PV were derived from data provided by CARRIERS.[5] We used aggregated data for the 32,247 cases and 32,544 controls in the 12 population-based studies of participants.[26-37] Population-based studies of breast cancer risk are more generalizable than studies of women accrued after clinical genetic testing, which is often performed due to strong family cancer history or early age at diagnosis. Our use of the population-based subset of CARRIERS data thus ensures that the models' results are more broadly relevant across the population. Overall breast

cancer risk estimates from CARRIERS for 28 cancer-predisposition genes have been previously published.[5] For this analysis, age-specific odds ratios of breast cancer for *ATM*, *CHEK2*, and *PALB2* were separately estimated using logistic regression models adjusted for study, first-degree family history of breast cancer, race and ethnicity, age, and an interaction of age and PV. These odds ratios were then applied to background age-, cohort-, and period-specific derived breast cancer incidence for the 1985 birth cohort of the US population (**Table S1**).[5] ER and HER2 breast cancer subtype distributions for each PV were also calculated and incorporated into the simulation models (**Table 1**).

Screening performance and breast cancer stage

Screening performance (i.e., sensitivity, specificity, and benign biopsy rates) for mammography and MRI were derived from published estimates from the High Risk Ontario Breast Screening Program (OBSP). The OBSP is an organized screening program for women at high-risk for breast cancer receiving annual mammography and MRI screening due to various risk factors (including genetic risk, family history, and history of prior radiation therapy to the chest).[9] Because mammography performance in the OBSP is calculated for mammography performed in conjunction with MRI (with cancers detected only by MRI counting as false negatives), we calibrated screening performance by simulating joint mammography and MRI screening and calculating the sensitivity of each modality in the same fashion. We also adjusted OBSP estimates of mammography sensitivity by age using data from the Breast Cancer Surveillance Consortium to account for the higher prevalence of dense breasts among young women.[43] Breast cancer stage distributions by mode of detection were estimated based on OBSP data,[9] with adjustments for missing pathologic stage information in women treated with neoadjuvant chemotherapy based on published estimates.[39] We assumed equal screening performance across PVs.

Treatment and mortality

To isolate the effects of screening on mortality, we assumed all women diagnosed with breast cancer received guideline-concordant age, stage, and subtype-specific therapy.[12] Treatment efficacy was based on meta-analyses of clinical trials.[41] Risk of non-breast cancer mortality was based on age and birth cohort-specific all-cause mortality rates.[42]

Analyses

We evaluated five primary screening strategies for each PV: annual mammography alone starting at age 40, and annual mammography starting at age 40 with annual MRI starting at ages 40, 35, 30, and 25. Annual mammography starting at age 40 was chosen because it is the least intensive screening mammography strategy recommended in the setting of elevated risk for breast cancer.[4,40] In a secondary analysis, we examined

Table 1. Model input parameters. Values shown in parentheses were used in sensitivity analyses.

Breast cancer risk and subtype				
	ATM	***PALB2***	***CHEK2***	**Reference**
Odds ratio of breast cancer	1.82	3.67	2.36	CARRIERS[5] (Age-specific odds ratios in Table S1)
Subtype distributions				CARRIERS[5]
ER+/HER2-	70%	47%	67%	
ER+/HER2+	22%	13%	22%	
ER-/HER2+	4%	1%	5%	
ER-/HER2-	4%	39%	7%	
Screening performance				
	MMG	**MRI**	**MMG + MRI**	
Sensitivity	Overall: 40.8% Age 30-39: 40.0% Age 40-49: 40.4% Age 50-69: 41.9%	90.8% (84.7%)	96.0% (92.2%)	Chiarelli et al.,[9]* with age-specific adjustments for MMG†
Specificity				
Initial Screen	88.0%	79.7% (78.8-80.6%)	72.2% (71.2-73.1%)	Chiarelli et al.[9]
With DBT	89.6%	---	73.8%	Conant et al.[38]‡
Second or later screen	92.5%	90.5% (89.9-91.0%)	84.5% (83.8-85.2%)	Chiarelli et al.[9]
With DBT	94.1%	---	85.5%	Conant et al.[38]‡
False-positives with biopsy performed (%)				Chiarelli et al.[9]
Initial Screen	19%	36%	28%	
Second or later screen	13%	38%	26%	
AJCC stage (screen-detected cancers)				
DCIS	22%	22%	23%	Chiarelli et al,[9] adjusted for missing stage of cancers treated with NAC[39]
I	48%	58%	57%	
II	24%	15%	15%	
III	6%	4%	4%	
Treatment/Mortality				
Treatment receipt	Guideline treatment by age, stage, and receptor status			NCCN[40]
Treatment efficacy	Estimated from meta-analyses of randomized trials			Peto et al.[41]
Non-breast cancer mortality	Age and birth cohort-specific all-cause mortality			Gangnon et al.[42]

ER=Estrogen receptor; HER2=Human epidermal growth factor receptor 2; BCSC=Breast Cancer Surveillance Consortium; AJCC=American Joint Commission Committee; NAC=Neoadjuvant chemotherapy; NCCN=National Comprehensive Cancer Network; MMG=mammogram; MRI=magnetic resonance imaging.

*In Chiarelli et al., all women received mammography and MRI performed concurrently, and sensitivity calculations for each modality include cancers detected by the other modality as false negatives. The models were therefore calibrated with mammography and MRI performed concurrently, adjusting the individual performance of each modality until the model output matched the observed data.

†Age-specific mammography sensitivity was derived from the overall sensitivity reported in Chiarelli et al. by adjusting for differences in density by age based on data from the BCSC.

‡Specificity of MMG and MMG+MRI were adjusted by decreasing false-positives due to MMG by 15%.

two additional strategies testing the impact of earlier start ages of mammography (35 and 30) with MRI screening at age 30. For all strategies we assumed screening was continued until age 74. To project the efficacy of screening, we assumed 100% screening participation. Simulations were continued until all women died of either breast cancer or non-breast cancer causes, and individual events were tracked and aggregated as lifetime population metrics.

Outcomes included lifetime screening benefits with screening versus without screening per 1,000 women screened, including breast cancer mortality reduction (expressed as the percent relative reduction in total breast cancer deaths), absolute breast cancer deaths averted and life years gained (LYG). We assessed cumulative lifetime screening resources and harms (total screens, false-positive screens, and benign biopsies) per 1,000 women screened. We also calculated ratios of screening harms per LYG (relative to no screening) to reflect the trade-offs between screening harms and benefits. Finally, we calculated incremental ratios of false-positive screens and benign biopsies per LYG for each strategy relative to the next least intensive screening strategy. Outcomes were reported as means and ranges across models.

Sensitivity analyses

We varied parameters related to breast cancer risk and screening performance to evaluate the impact of parameter uncertainty on the robustness of results. We assumed higher and lower breast cancer risk for each PV by adding and subtracting the standard error (SE) from the age-specific risk estimates provided by CARRIERS (**Table S1**). We varied the sensitivity of MRI alone and mammography with MRI using the lower bounds of the confidence intervals (CIs)[9] and MRI specificity across the upper and lower CIs in the OBSP (**Table 1**)[9]. To account for potential differences in screening specificity by age, we evaluated outcomes using alternative screening specificity estimates stratified by age and screening round provided by the Breast Cancer Surveillance Consortium (**Table S2**). Finally, we considered a scenario with digital breast tomosynthesis (DBT) used for mammography screening; for this scenario, we assumed a 15% reduction in mammography false-positives (**Table 1**) with no change in overall sensitivity based on prior work.[38]

RESULTS

Breast cancer incidence and mortality

Among women with *ATM*, *CHEK2*, and *PALB2* PVs, average model projections and ranges for cumulative lifetime breast cancer risk in the absence of screening were 21% (range across models 18-24%), 28% (23-33%), and 38% (36-40%) respectively; cumulative

lifetime risks of breast cancer death in the absence of screening were 3.4% (2.4-4.5%), 4.6% (3.1-6.1%) and 7.7% (6.4-9.1%) for *ATM, CHEK* and *PALB2,* respectively.

Screening benefits and harms by screening strategy

Across PVs, lifetime mortality benefits and screening harms increased with multimodality screening and with younger screening ages compared to mammography alone starting at age 40. Per 1,000 women screened, annual mammography alone starting at age 40 compared to no screening resulted in a 36-39% reduction in breast cancer mortality, 13.3-29.7 breast cancer deaths averted, 2,092-2,224 false-positive screens, and 279-296 benign biopsies across PVs (**Table 2** and **Table 3**). Annual mammography and MRI starting at age 40 reduced breast cancer mortality by 52-54% and resulted in 18.4-42.4 breast cancer deaths averted, 4,233-4,569 false-positive screens, and 1,109-1,196 benign biopsies compared to no screening. The most intensive strategy (annual MRI alone from ages 25-39 and annual mammography and MRI from 40-74) reduced breast cancer mortality by 56-60%, averted 20.5-40.0 breast cancer deaths and resulted in 5,592-5,932 false-positive screens and 1,637-1,725 benign biopsies per 1,000 women.

Screening efficiency

Compared to mammography alone from 40-74, annual MRI screening from 35-39 followed by mammography and MRI from 40-74 was more efficient (fewer false-positive screens and biopsies per LYG) than annual mammography and MRI from 40-74 for all PVs (**Figure 1** and **Table S3**). This strategy resulted in 7.0-15.3 additional false-positives and 2.7-5.9 benign biopsies per LYG relative to mammography alone from 40-74. Starting MRI at age 30 versus 35 similarly resulted in 12.8-15.2 additional false-positives and 4.9-5.9 benign biopsies per LYG, making it another efficient strategy. However, starting MRI at 25 was considerably less efficient than starting at 30, resulting in 47.0-57.9 additional false-positive screens and 18.0-22.2 benign biopsies per LYG. Results were similar when comparing false-positives and benign biopsies per breast cancer death averted (**Figure S2**).

Strategies with earlier mammography

With annual MRI screening from ages 30-39, starting mammography earlier than age 40 increased false- positive screens and benign biopsies but had little impact on mortality reduction or LYG (**Table 4**). For example, with annual MRI starting at 30, adding mammography at age 30 versus 40 decreased mortality by only 0.1-0.3% and increased LYG by 3-5 per 1,000 women screened but resulted in 649-650 additional false-positive exams and 58-59 additional biopsies per 1000 women screened **(Table 4)**.

Table 2. Lifetime benefits of MRI screening strategies with annual mammography from ages 40-74 alone and with annual magnetic resonance imaging (MRI) at varying start ages for women with pathogenic variants in ATM, CHEK2, and PALB2.

	Breast Cancer Mortality Reduction (%)			Life Years Gained			Breast Cancer Deaths Averted		
	ATM	CHEK2	PALB2	ATM	CHEK2	PALB2	ATM	CHEK2	PALB2
Annual MMG at 40	38.5 (37.8-39.2)	38.4 (38.0-38.8)	36.4 (34.6-38.2)	291 (263-319)	370 (330-409)	621 (559-684)	13 (9-18)	17 (12-23)	30 (22-37)
+MRI at 40	53.6 (52.9-54.3)	53.6 (53.3-53.9)	52.3 (51.4-53.1)	420 (388-452)	533 (489-577)	921 (876-967)	18 (13-24)	24 (16-32)	42 (33-52)
+MRI at 35	57.6 (57.2-58.0)	57.0 (56.3-57.7)	54.4 (54.2-54.7)	473 (447-498)	591 (555-627)	992 (959-1,025)	20 (14-26)	26 (18-33)	44 (34-54)
+MRI at 30	59.5 (58.5-60.4)	58.4 (57.2-59.6)	55.4 (55.3-55.4)	501 (478-523)	620 (587-652)	1,025 (998-1,051)	20 (14-26)	26 (18-34)	45 (35-55)
+MRI at 25	60.2 (58.9-61.2)	58.9 (57.5-60.3)	55.7 (55.5-55.8)	510 (489-531)	630 (599-661)	1,037 (1,013-,061)	20 (14-26)	26 (18-34)	45 (35-55)

Results are shown as model averages (ranges) of cumulative lifetime outcomes per 1,000 women screened across Model E and Model W-H.

Table 3. Total screens, false-positive screens, and benign biopsies for screening strategies with annual mammography ages 40-74 alone and with annual magnetic resonance imaging (MRI) for women with pathogenic variants in ATM, CHEK2, and PALB2.

	Total Screens			False-Positive Screens			Benign Biopsies		
	ATM	CHEK2	PALB2	ATM	CHEK2	PALB2	ATM	CHEK2	PALB2
Annual MMG at 40	29,182 (29,148-29,215)	28,505 (28,582-28,527)	27,412 (27,321-27,503)	2,224 (2,222-2,227)	2,174 (2,172-2,175)	2,092 (2,085-2,099)	296 (296-297)	290 (290-290)	279 (278-280)
+MRI at 40	57,173 (57,050-57,296)	55,511 (55,463-55,559)	52,814 (52,664-52,964)	4,569 (4,555-4,583)	4,441 (4,438-4,443)	4,233 (4,213-4,252)	1,196 (1,193-1,200)	1,163 (1,162-1,164)	1,109 (1,104-1,114)
+MRI at 35	61,789 (61,568-62,010)	60,104 (60,058-60,150)	57,392 (57,149-57,636)	5,001 (4,979-5,023)	4,871 (4,861-4,880)	4,661 (4,635-4,688)	1,368 (1,362-1,374)	1,334 (1,331-1,337)	1,280 (1,272-1,287)
+MRI at 30	66,100 (65,867-66,333)	64,403 (63,988-64,818)	61,694 (61,474-61,913)	5,415 (5,393-5,437)	5,284 (5,249-5,319)	5,075 (5,057-5,093)	1,528 (1,517-1,538)	1,493 (1,479-1,508)	1,439 (1,429-1,449)
+MRI at 25	71,507 (71,247-71,767)	69,819 (69,731-69,906)	67,100 (66,827-67,373)	5,932 (5,907-5,957)	5,802 (5,789-5,815)	5,592 (5,563-5,621)	1,725 (1,718-1,732)	1,691 (1,687-1,695)	1,637 (1,629-1,645)

Results are shown as model averages (ranges) of cumulative lifetime outcomes per 1,000 women screened across Model E and Model W-H.

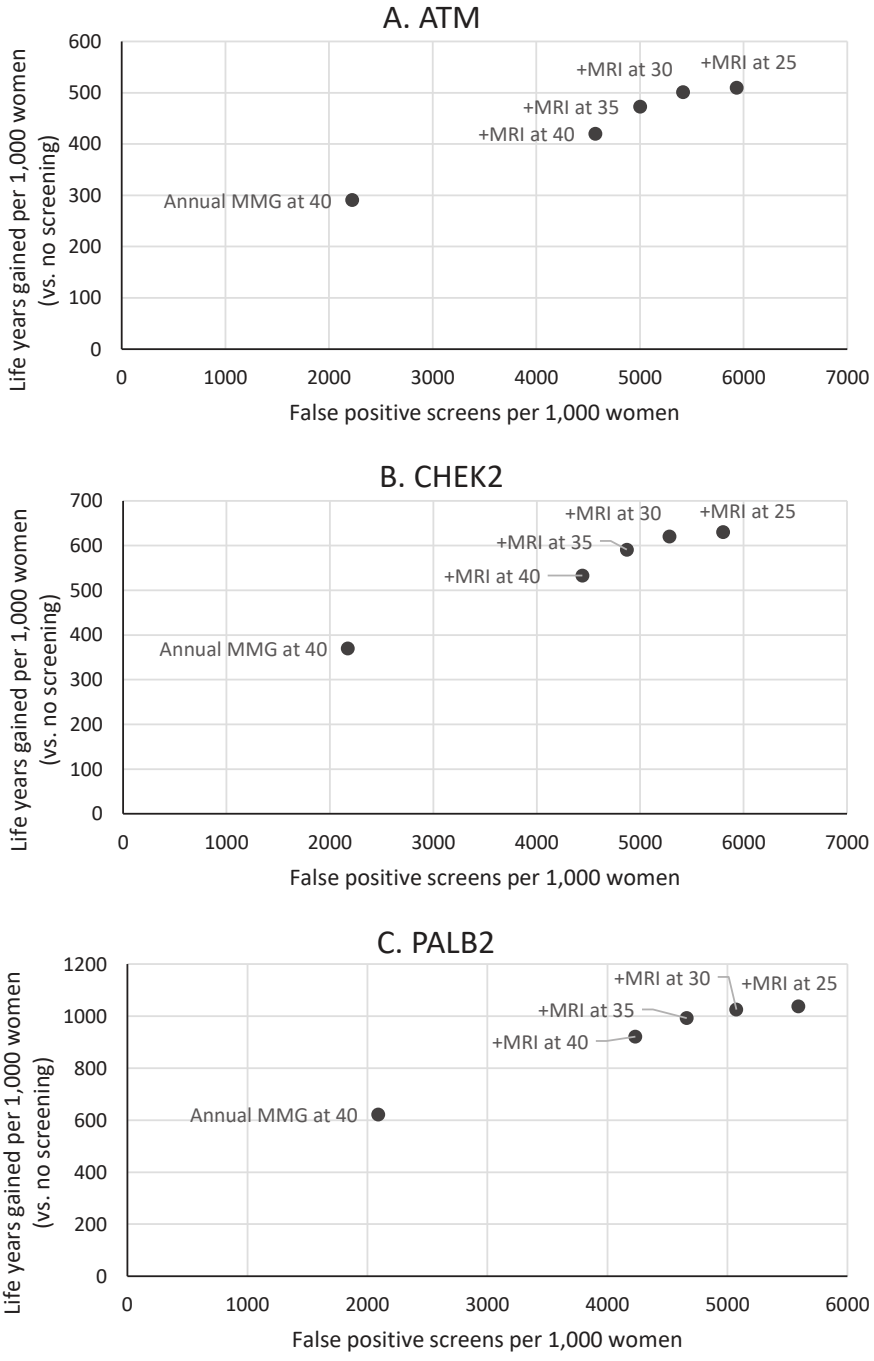


Figure 1. False-positive exams and life years gained for screening strategies for women with pathogenic variants in *ATM* (panel A), *CHEK2* (panel B), and *PALB2* (panel C).

Results are mean model projections across Model E and Model W-H. MMG=Mammography; MRI=Magnetic resonance imaging. In all strategies, MMG is performed annually from ages 40-74; MRI varies in start age by strategy.

Table 4. Impact of starting mammography (MMG) at earlier ages for strategies with magnetic resonance imaging (MRI) screening starting at age 30 for women with pathogenic variants in ATM, CHEK2, and PALB2.

	Breast Cancer Mortality Reduction (%)			LYG		
	ATM	**CHEK2**	**PALB2**	**ATM**	**CHEK2**	**PALB2**
MRI at 30, annual MMG at 40 vs no screening	59.5 (58.5-60.4)	58.4 (57.2-59.6)	55.4 (55.3-55.4)	501 (478-523)	620 (587-652)	1,025 (998-1,051)
MRI at 30, annual MMGat 35 vs 40	0.2 (0.1-0.2)	0.1 (0.1-0.2)	0.1 (0.1-0.1)	2 (2-2)	3 (2-3)	3 (3-3)
MRI at 30, annual MMG at 30 vs 40	0.3 (0.2-0.3)	0.2 (0.1-0.2)	0.1 (0.1-0.2)	3 (3-4)	4 (4-5)	5 (5-5)

Table 4 continued	**False positive Screens**			**Benign Biopsies**		
	ATM	**CHEK2**	**PALB2**	**ATM**	**CHEK2**	**PALB2**
MRI at 30, annual MMG at 40 vs no screening	5,415 (5,437-5,393)	5,284 (5,249-5,319)	5,075 (5,057-5,093)	1,528 (1,517-1,538)	1,493 (1,479-1,508)	1,439 (1,429-1,449)
MRI at 30, annual MMG at 35 vs 40	338 (291-386)	339 (291-387)	338 (291-385)	37 (20-55)	38 (20-55)	37 (20-55)
MRI at 30, annual MMG at 30 vs 40	650 (603-696)	650 (603-696)	649 (602-695)	59 (41-76)	59 (41-76)	58 (41-76)

Results are shown as model averages (ranges) of cumulative lifetime outcomes per 1,000 women screened across Model E and Model W-H.

Sensitivity analyses

Sensitivity analyses varying breast cancer risk (**Table S5**, **Table S6, and Figure S2**), MRI sensitivity (**Table S7**), and screening specificity (**Table S8**, **Table S9**, **Table S10**, **Table S11**, and **Figures S3A-S3C**) did not meaningfully change conclusions regarding the relative efficiency of the screening strategies, although absolute benefits and harms varied by scenario. Using lower and higher breast cancer risk estimates, average lifetime breast cancer risk ranged from 18-25% for *ATM*, 25-31% for *CHEK2*, and 33-47% for PALB2. When screening with annual MRI from 30-39 and mammography and MRI from 40-74, breast cancer mortality reduction ranged from 57-60% across the range of breast cancer risk considered and only decreased to 52-55% with lower MRI sensitivity. When varying MRI specificity, false-positive screens per 1,000 women ranged from 4,841-5,165 with the best specificity to 5,318-5,673 with the lowest specificity based on OBSP data (**Table S8 and Table S9**) and 5,106-5,375 false-positives using age-specific specificity from the Breast Cancer Surveillance Consortium (**Table S10**). When mammography specificity was adjusted for DBT screening, false-positives decreased to 4,800-5,117 per 1000 women (**Table S11**).

DISCUSSION

This is the first study to use comparative modeling to evaluate breast cancer screening strategies for women with *ATM*, *CHEK*2, and *PALB2* PVs. We found that for women with these PVs, annual MRI and mammography screening reduced breast cancer mortality by more than 50% for all strategies considered. Based on our results, annual MRI screening starting at age 30-35 followed by combined annual MRI and mammography at age 40 offers the best balance of screening benefits and harms. Starting mammography earlier than age 40 increases false-positive screens and benign biopsies but adds little benefit for women receiving MRI.

Our study adds to the growing body of literature supporting the use of MRI in women at elevated risk of breast cancer. Prior modeling analyses using CISNET models have estimated that MRI screening reduces breast cancer mortality by 38-62% in women with *BRCA1/2* PVs [44,45] and 56-71% in women with a history of radiation therapy to the chest.[46] We estimated a 52-60% reduction in breast cancer mortality with MRI among women with PVs in *ATM*, *CHEK*2, or *PALB2*, suggesting that MRI has important benefits even in the setting of moderate risk due to genetic susceptibility. Of note, women with *PALB2* PVs benefited most from MRI screening with the most breast cancer deaths averted and greatest life expectancy gains for all strategies, although the incremental benefits of starting at age 30 versus 35 were similar to *CHEK2* and *ATM*. This larger benefit is expected given their higher risk of breast cancer overall and for ER-negative breast cancers, which have poorer prognosis.[5]

We examined incremental benefits and harms for MRI screening starting at ages 25-40 to quantify relative benefits of earlier screening. Although there is no established benefit/risk threshold for women with moderate or high-risk of breast cancer, we found that MRI screening prior to age 30 considerably increases false-positive screens and biopsies with little impact on mortality or life expectancy. In other settings, efficiency ratios of procedures per LYG have been used to guide cancer screening policies, such as 40 colonoscopies per LYG for colorectal cancer screening.[16] Based on our results, starting MRI at 35 and starting MRI at 30 resulted in approximately 3-6 and 5-6 additional benign biopsies per LYG relative to the next least intensive strategy, respectively, suggesting the trade-offs of starting MRI screening in this age range are likely acceptable.

Another finding of our analysis is that mammography screening prior to age 40 offered little benefit when women were receiving annual screening MRI, but increased screening harms. The value of screening mammography in women younger than 40 receiving MRI has been questioned,[9,47,48] as it is uncommon for mammography to detect a cancer missed on MRI in young women[48-50] and the few cancers not detected on MRI are typically more indolent, including low-grade DCIS.[51] Our models suggest that earlier mammography in this setting also has considerable disadvantages, substantially

increasing false-positive screens and benign biopsies. Young women are also potentially more susceptibility to risks of radiation-induced malignancy from mammography screening,[52] potentially further reducing the benefit and increasing harms of early mammography. For this reason, the use of MRI screening starting at 25 followed by combined mammography/MRI screening starting at 30 is recommended for women with *BRCA1/2* PVs.[40] Our results suggest that when screening with MRI, mammography could be delayed until age 40 in women with *ATM, CHEK2, and PALB2* PVs.

This study has many strengths, including the use of two well-established CISNET models, consistent results across models and the use of population-based breast cancer risk estimates from the largest US study of genetic breast cancer risk. Several limitations are also worth noting. First, our model outcomes are population-level metrics that do not account for all factors that should be considered when selecting a screening strategy for an individual woman, including family history. Second, we only considered annual MRI screening intervals because longer intervals in women with genetic susceptibility to breast cancer have not been evaluated in clinical trials or used in clinical practice. Outcomes for biennial and triennial MRI intervals, age-specific MRI intervals and novel combinations of alternating MRI and mammography should be evaluated in future analyses intended to guide clinical trial design. Third, we did not consider costs or quality of life in this analysis, and future analyses are warranted to evaluate cost-effectiveness. Fourth, we did not estimate screening-related overdiagnosis given the paucity of data on overdiagnosis due to MRI screening. While MRI screening detects more breast cancers than mammography, MRI may preferentially detect more biologically significant cancers[51] and the proportion of cancers due to overdiagnosis may be lower than mammography. Finally, we did not evaluate alternative screening modalities such as whole breast ultrasound or contrast-enhanced mammography. To date, however, MRI is the only advanced imaging modality shown to decrease advanced breast cancers[11] and interval cancers.[10] Additional data on cancer outcomes with other modalities for women with genetic susceptibility to breast cancer are needed to inform guidelines for their use.

In summary, this comparative modeling analysis supports the use of MRI screening in women with moderate-to-high risk of breast cancer due to *ATM, CHEK2,* and *PALB2* PVs. Annual MRI screening starting at age 30-35 followed by annual MRI and mammography starting at age 40 reduces breast cancer mortality by more than 50% in these women, while additional mammography prior to age 40 is likely of little benefit.

Acknowledgements

The authors would like to thank Yu-Ru Su, PhD (Kaiser Permanente Washington Health Research Institute, Seattle WA) for analyzing the screening performance data from the Breast Cancer Surveillance Consortium used in sensitivity analyses.

REFERENCES

1. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science.* 1994;266(5182):66-71.
2. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature.* 1995;378(6559):789-792.
3. Lynch HT, Snyder C, Lynch J. Hereditary breast cancer: practical pursuit for clinical translation. *Ann Surg Oncol.* 2012;19(6):1723-1731.
4. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin.* 2007;57(2):75-89.
5. Hu C, Hart SN, Gnanaolivu R, et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med.* 2021;384(5):440-451.
6. Breast Cancer Association C, Dorling L, Carvalho S, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med.* 2021;384(5):428-439.
7. Kurian AW, Ward KC, Hamilton AS, et al. Uptake, Results, and Outcomes of Germline Multiple-Gene Sequencing After Diagnosis of Breast Cancer. *JAMA Oncol.* 2018;4(8):1066-1072.
8. Stadler ZK, Schrader KA, Vijai J, Robson ME, Offit K. Cancer genomics and inherited risk. *J Clin Oncol.* 2014;32(7):687-698.
9. Chiarelli AM, Blackmore KM, Muradali D, et al. Performance Measures of Magnetic Resonance Imaging Plus Mammography in the High Risk Ontario Breast Screening Program. *J Natl Cancer Inst.* 2020;112(2):136-144.
10. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.
11. Warner E, Hill K, Causer P, et al. Prospective study of breast cancer incidence in women with a BRCA1 or BRCA2 mutation under surveillance with and without magnetic resonance imaging. *J Clin Oncol.* 2011;29(13):1664-1669.
12. Fracheboud J, Otto SJ, van Dijck JAAM, Broeders MJM, Verbeek ALM, de Koning HJ. Decreased rates of advanced breast cancer due to mammography screening in The Netherlands. *British Journal of Cancer.* 2004;91:861-867.
13. Mandelblatt JS, Stout NK, Schechter CB, et al. Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies. *Ann Intern Med.* 2016;164(4):215-225.
14. Meza R, Jeon J, Toumazis I, et al. Evaluation of the Benefits and Harms of Lung Cancer Screening With Low-Dose Computed Tomography: Modeling Study for the US Preventive Services Task Force. *JAMA.* 2021;325(10):988-997.
15. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of Benefits, Burden, and Harms of Colorectal Cancer Screening Strategies: Modeling Study for the US Preventive Services Task Force. *JAMA.* 2016;315(23):2595-2609.
16. Peterse EFP, Meester RGS, Siegel RL, et al. The impact of the rising colorectal cancer incidence in young adults on the optimal age to start screening: Microsimulation analysis I to inform the American Cancer Society colorectal cancer screening guideline. *Cancer.* 2018;124(14):2964-2973.
17. Meester RGS, Peterse EFP, Knudsen AB, et al. Optimizing colorectal cancer screening by race and sex: Microsimulation analysis II to inform the American Cancer Society colorectal cancer screening guideline. *Cancer.* 2018;124(14):2974-2985.

18. van den Broek JJ, van Ravesteyn NT, Heijnsdijk EA, de Koning HJ. Simulating the Impact of Risk-Based Screening and Treatment on Breast Cancer Outcomes with MISCAN-Fadia. *Med Decis Making.* 2018;38(1_suppl):54S-65S.
19. Alagoz O, Berry DA, de Koning HJ, et al. Introduction to the Cancer Intervention and Surveillance Modeling Network (CISNET) Breast Cancer Models. *Medical Decision Making.* 2018;38(1_suppl):3S-8S.
20. van den Broek JJ, van Ravesteyn NT, Mandelblatt JS, et al. Comparing CISNET Breast Cancer Incidence and Mortality Predictions to Observed Clinical Trial Results of Mammography Screening from Ages 40 to 49. *Med Decis Making.* 2018;38(1_suppl):140S-150S.
21. Alagoz O, Ergun MA, Cevik M, et al. The University of Wisconsin Breast Cancer Epidemiology Simulation Model: An Update. *Medical Decision Making.* 2018;38(1_suppl):99S-111S.
22. Mandelblatt JS, Near AM, Miglioretti DL, et al. Common Model Inputs Used in CISNET Collaborative Breast Cancer Modeling. *Medical Decision Making.* 2018;38(1_suppl):9S-23S.
23. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. *J Natl Cancer Inst Monogr.* 2006(36):37-47.
24. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr.* 2006(36):56-65.
25. Gangnon RE, Sprague BL, Stout NK, et al. The contribution of mammography screening to breast cancer incidence trends in the United States: an updated age-period-cohort model. *Cancer Epidemiol Biomarkers Prev.* 2015;24(6):905-912.
26. Palmer JR, Ruiz-Narvaez EA, Rotimi CN, et al. Genetic susceptibility loci for subtypes of breast cancer in an African American population. *Cancer Epidemiol Biomarkers Prev.* 2013;22(1):127-134.
27. Calle EE, Rodriguez C, Jacobs EJ, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer.* 2002;94(9):2490-2501.
28. Patel AV, Jacobs EJ, Dudas DM, et al. The American Cancer Society's Cancer Prevention Study 3 (CPS-3): Recruitment, study design, and baseline characteristics. *Cancer.* 2017;123(11):2014-2024.
29. Bernstein L, Allen M, Anton-Culver H, et al. High breast cancer incidence rates among California teachers: results from the California Teachers Study (United States). *Cancer Causes Control.* 2002;13(7):625-635.
30. Vachon CM, Li J, Scott CG, et al. No evidence for association of inherited variation in genes involved in mitosis and percent mammographic density. *Breast Cancer Res.* 2012;14(1):R7.
31. Kolonel LN, Henderson BE, Hankin JH, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol.* 2000;151(4):346-357.
32. Olson JE, Sellers TA, Scott CG, et al. The influence of mammogram acquisition on the mammographic density and breast cancer association in the Mayo Mammography Health Study cohort. *Breast Cancer Res.* 2012;14(6):R147.
33. Eckel N, Li Y, Kuxhaus O, Stefan N, Hu FB, Schulze MB. Transition from metabolic healthy to unhealthy phenotypes and association with cardiovascular disease risk across BMI categories in 90 257 women (the Nurses' Health Study): 30 year follow-up from a prospective cohort study. *Lancet Diabetes Endocrinol.* 2018;6(9):714-724.
34. Hirko KA, Chai B, Spiegelman D, et al. Erythrocyte membrane fatty acids and breast cancer risk: a prospective analysis in the nurses' health study II. *Int J Cancer.* 2018;142(6):1116-1129.

35. Ambrosone CB, Ciupak GL, Bandera EV, et al. Conducting Molecular Epidemiological Research in the Age of HIPAA: A Multi-Institutional Case-Control Study of Breast Cancer in African-American and European-American Women. *J Oncol.* 2009;2009:871250.
36. Prentice RL, Caan B, Chlebowski RT, et al. Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA.* 2006;295(6):629-642.
37. Trentham-Dietz A, Sprague BL, Hampton JM, et al. Modification of breast cancer risk according to age and menopausal status: a combined analysis of five population-based case-control studies. *Breast Cancer Res Treat.* 2014;145(1):165-175.
38. Conant EF, Barlow WE, Herschorn SD, et al. Association of Digital Breast Tomosynthesis vs Digital Mammography With Cancer Detection and Recall Rates by Age and Breast Density. *JAMA Oncol.* 2019.
39. Bergquist JR, Murphy BL, Storlie CB, Habermann EB, Boughey JC. Incorporation of Treatment Response, Tumor Grade and Receptor Status Improves Staging Quality in Breast Cancer Patients Treated with Neoadjuvant Chemotherapy. *Ann Surg Oncol.* 2017;24(12):3510-3517.
40. Daly MB, Pal T, Berry MP, et al. Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2021;19(1):77-102.
41. Early Breast Cancer Trialists' Collaborative G, Peto R, Davies C, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet.* 2012;379(9814):432-444.
42. Gangnon RE, Stout NK, Alagoz O, Hampton JM, Sprague BL, Trentham-Dietz A. Contribution of Breast Cancer to Overall Mortality for US Women. *Medical Decision Making.* 2018;38(1_suppl):24S-31S.
43. Sprague BL, Gangnon RE, Burt V, et al. Prevalence of mammographically dense breasts in the United States. *J Natl Cancer Inst.* 2014;106(10).
44. Plevritis SK, Kurian AW, Sigal BM, et al. Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. *JAMA.* 2006;295(20):2374-2384.
45. Heijnsdijk EA, Warner E, Gilbert FJ, et al. Differences in natural history between breast cancers in BRCA1 and BRCA2 mutation carriers and effects of MRI screening-MRISC, MARIBS, and Canadian studies combined. *Cancer Epidemiol Biomarkers Prev.* 2012;21(9):1458-1468.
46. Yeh JM, Lowry KP, Schechter CB, et al. Clinical Benefits, Harms, and Cost-Effectiveness of Breast Cancer Screening for Survivors of Childhood Cancer Treated With Chest Radiation : A Comparative Modeling Study. *Ann Intern Med.* 2020;173(5):331-341.
47. Obdeijn IM, Winter-Warnars GA, Mann RM, Hooning MJ, Hunink MG, Tilanus-Linthorst MM. Should we screen BRCA1 mutation carriers only with MRI? A multicenter study. *Breast Cancer Res Treat.* 2014;144(3):577-582.
48. Obdeijn IM, Mann RM, Loo CCE, et al. The supplemental value of mammographic screening over breast MRI alone in BRCA2 mutation carriers. *Breast Cancer Res Treat.* 2020;181(3):581-588.
49. Kuhl C, Weigel S, Schrading S, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *J Clin Oncol.* 2010;28(9):1450-1457.
50. Narayan AK, Visvanathan K, Harvey SC. Comparative effectiveness of breast MRI and mammography in screening young women with elevated risk of developing breast cancer: a retrospective cohort study. *Breast Cancer Res Treat.* 2016;158(3):583-589.

51. Sung JS, Stamler S, Brooks J, et al. Breast Cancers Detected at Screening MR Imaging and Mammography in Patients at High Risk: Method of Detection Reflects Tumor Histopathologic Results. *Radiology*. 2016;280(3):716-722.
52. Berrington de Gonzalez A, Berg CD, Visvanathan K, Robson M. Estimated Risk of Radiation-Induced Breast Cancer From Mammographic Screening for Young BRCA Mutation Carriers. *JNCI: Journal of the National Cancer Institute*. 2009;101(3):205-209.

SUPPLEMENTARY APPENDIX

Table S1. Age-specific odds ratios for breast cancer used in base case analysis and sensitivity analyses. Lower and upper bounds were estimated by adding and subtracting one standard error from the base case.

Age	*ATM*			CHEK2			PALB2		
	Lower	Base	Upper	Lower	Base	Upper	Lower	Base	Upper
35	2.29	3.01	3.97	2.44	3.15	4.06	2.15	3.22	4.84
40	2.16	2.72	3.43	2.40	2.98	3.69	2.36	3.31	4.65
45	2.04	2.46	2.97	2.36	2.82	3.36	2.57	3.40	4.49
50	1.91	2.23	2.59	2.31	2.66	3.07	2.78	3.49	4.38
55	1.78	2.01	2.28	2.25	2.52	2.83	2.96	3.59	4.35
60	1.63	1.82	2.04	2.15	2.38	2.64	3.06	3.69	4.44
65	1.46	1.64	1.86	2.02	2.25	2.51	3.07	3.79	4.67
70	1.28	1.49	1.73	1.87	2.13	2.43	3.02	3.89	5.02

Odds ratios were provided from the CARRIERS consortium and estimated using logistic regression adjusted for study, first degree family history of breast cancer, race/ethnicity, age, and an interaction of age and pathogenic variant.

Table S2. Specificity of screening with mammography alone, MRI alone, and mammography combined with MRI stratified by age group and screening round.

	Initial screen			Rescreen		
	MMG	MRI	MMG+MRI	MMG	MRI	MMG+MRI
Age <50	59%	79%	68%	82%	92%	87%
Age ≥50	70%	85%	77%	88%	95%	92%

Data provided from the Breast Cancer Surveillance Consortium. Specificity was calculated based on 7,424 MRI and 5,671 mammography screening examinations performed for high-risk screening in women without a personal history of breast cancer at BCSC facilities in 2005 -2020.

Table S3. Ratios of screening harms per life year gained for screening strategies with varying start age of magnetic resonance imaging, relative to no screening.

	False-Positive Screens per LYG Model Average (Range)			Benign Biopsies per LYG Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	7.7 (7.0-8.5)	5.9 (5.3-6.6)	3.4 (3.1-3.7)	1.0 (0.9-1.1)	0.8 (0.7-0.9)	0.5 (0.4-0.5)
+MRI at 40	10.9 (10.1-11.7)	8.4 (7.7-9.1)	4.6 (4.4-4.8)	2.9 (2.7-3.1)	2.2 (2.0-2.4)	1.2 (1.2-1.3)
+MRI at 35	10.6 (10.1-11.1)	8.3 (7.8-8.8)	4.7 (4.6-4.8)	2.9 (2.8-3.0)	2.3 (2.1-2.4)	1.3 (1.3-1.3)
+MRI at 30	10.8 (10.3-11.4)	8.6 (8.0-9.1)	5.0 (4.8-5.1)	3.1 (2.9-3.2)	2.4 (2.3-2.6)	1.4 (1.4-1.5)
+MRI at 25	11.6 (11.2-12.1)	9.2 (8.8-9.7)	5.4 (5.3-5.5)	3.4 (3.3-3.5)	2.7 (2.6-2.8)	1.6 (1.6-1.6)

MMG: mammography; MRI: magnetic resonance imaging.
Strategies ranked in order of increasing total number of screens, with ratios calculated for each strategy relative to no screening.

Table S4. Incremental screening harms per life year gained for screening strategies with varying start age of magnetic resonance imaging.

	False-Positive Screens per LYG Model Average (Range)			Benign Biopsies per LYG Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	7.7 (7.0-8.5)	5.9 (5.3-6.6)	3.4 (3.7-3.1)	1.0 (0.9-1.1)	0.8 (0.7-0.9)	0.5 (0.4-0.5)
+MRI at 40	18.1 (17.7-18.6)	13.8 (13.5-14.2)	7.2 (6.7-7.6)	7.0 (6.8-7.1)	5.3 (5.2-5.5)	2.8 (2.6-3.0)
+MRI at 35	15.3 (14.9-15.6[a])	12.2 (12.0-12.4[a])	7.0 (6.4-7.6)[a]	5.9 (5.8-6.0[a])	4.7 (4.6-4.8[a])	2.7 (2.5-3.0[a])
+MRI at 30	15.2 (14.9-15.5[b])	14.4 (14.0-14.7)	12.8 (11.6-14.0)	5.9 (5.8-6.0[b])	5.5 (5.4-5.7)	4.9 (4.5-5.4)
+MRI at 25	57.9 (43.5-72.3)	54.3 (41.2-67.3)	47.0 (32.6-61.3)	22.2 (16.7-27.7)	20.8 (15.8-25.7)	18.0 (12.5-23.4)

MMG: mammography; MRI: magnetic resonance imaging.
Incremental ratios are calculated for each strategy relative to the next least screening intensive strategy. Strategies ranked in order of increasing total number of screens, with incremental ratios calculated for each strategy relative to the next least screening intensive strategy.
[a]Starting MRI at 40 is less efficient than MRI at 35; incremental ratios for MRI at 35 are calculated relative to Mammography at 40 (without MRI)
[b]Starting MRI at 35 or 40 is less efficient than MRI at 30; incremental ratios for MRI at 30 are calculated relative to Mammography at 40 (without MRI)

Table S5. Sensitivity analysis of benefits of screening strategies assuming higher breast cancer risk. Age-specific risk estimates from CARRIERS were increased by one standard error (see Table S1).

	Breast Cancer Mortality Reduction (%) Model Average (Range)			*LYG* Model Average (Range)			*Deaths Averted* Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	38 (37-38)	38 (37-38)	38 (35-40)	248 (229-266)	325 (290-360)	521 (454-589)	11 (8-15)	15 (10-20)	25 (18-33)
+MRI at 40	53 (52-54)	53 (53-54)	54 (52-56)	355 (338-372)	470 (430-510)	775 (709-840)	16 (11-20)	22 (14-29)	36 (26-46)
+MRI at 35	57 (57-57)	57 (56-57)	56 (55-57)	397 (383-410)	516 (481-551)	824 (766-883)	17 (12-21)	23 (15-30)	37 (27-47)
+MRI at 30	59 (58-60)	59 (58-60)	57 (57-58)	419 (406-431)	539 (506-573)	846 (792-901)	17 (12-22)	23 (16-30)	38 (28-47)
+MRI at 25	60 (60-61)	59 (58-60)	58 (57-58)	426 (415-437)	547 (515-579)	854 (802-907)	17 (12-22)	23 (16-30)	38 (28-48)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

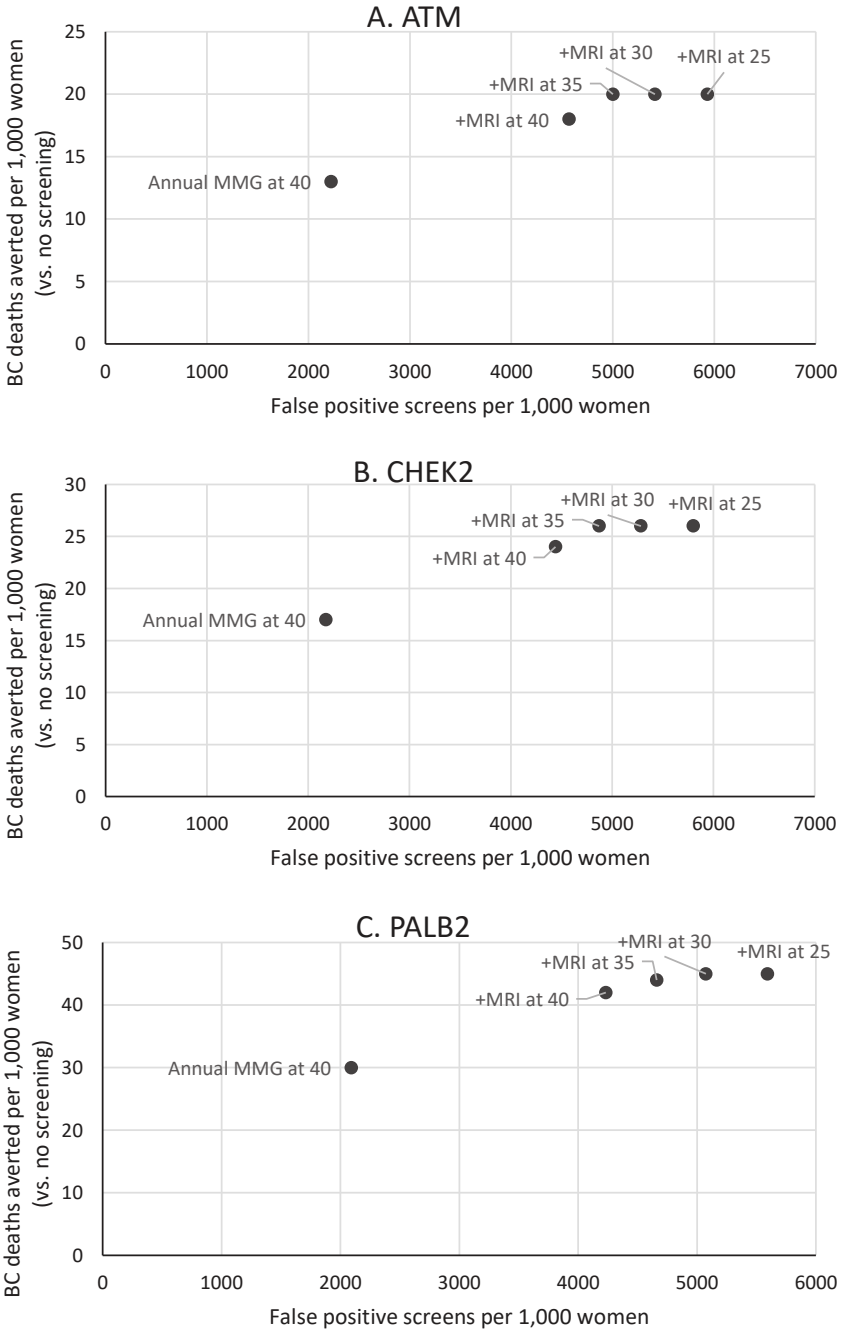


Figure S1. False-positive exams and breast cancer (BC) deaths averted for screening strategies for women with pathogenic variants in ATM (panel A), CHEK2 (panel B), and PALB2 (panel C).
Results are mean model projections across Model E and Model W-H. MMG=Mammography; MRI=Magnetic resonance imaging. In all strategies, MMG is performed annually from ages 40-74; MRI start age varies by strategy.

Table S6. Sensitivity analysis of benefits of screening strategies assuming lower breast cancer risk. Age-specific risk estimates from CARRIERS were decreased by one standard error (see Table S1). Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

	Breast Cancer Mortality Reduction (%) Model Average (Range)			LYG Model Average (Range)			Deaths Averted Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	39 (38-40)	38 (38-39)	36 (35-38)	339 (303-375)	418 (377-459)	765 (692-837)	15 (10-21)	19 (13-25)	36 (28-45)
+MRI at 40	54 (54-55)	54 (54-54)	52 (52-53)	490 (448-531)	604 (560-649)	1131 (1079-1183)	22 (14-29)	27 (19-35)	52 (41-63)
+MRI at 35	58 (57-58)	57 (56-58)	54 (54-54)	557 (524-592)	678 (641-714)	1224 (1197-1251)	23 (16-30)	29 (20-37)	54 (43-65)
+MRI at 30	60 (59-60)	58 (57-59)	55 (54-55)	593 (565-622)	717 (683-750)	1271 (1257-1285)	24 (17-31)	30 (21-38)	55 (44-65)
+MRI at 25	60 (59-61)	59 (57-60)	55 (54-55)	606 (579-633)	729 (698-760)	1287 (1279-1296)	24 (17-31)	30 (21-38)	55 (45-66)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

Table S7. Sensitivity analysis of benefits of screening strategies assuming the lower confidence limit of MRI sensitivity.

	Breast Cancer Mortality Reduction (%) Model Average (Range)			LYG Model Average (Range)			Deaths Averted Model Average (Range)		
	ATM	CHEK2	PALB2	ATM	CHEK2	PALB2	ATM	CHEK2	PALB2
MMG at 40	38 (37-38)	37 (36-38)	35 (35-56)	281 (263-299)	356 (330-381)	599 (559-639)	13 (9-17)	17 (12-22)	29 (22-35)
+MRI at 40	50 (49-52)	50 (49-52)	49 (47-51)	393 (351-434)	498 (442-553)	854 (781-927)	17 (12-23)	23 (15-31)	40 (30-50)
+MRI at 35	54 (53-55)	53 (53-54)	51 (49-53)	437 (400-474)	548 (497-599)	915 (849-981)	19 (13-25)	24 (16-22)	41 (31-52)
+MRI at 30	55 (55-56)	55 (54-55)	52 (50-53)	462 (425-498)	573 (523-622)	942 (880-1004)	19 (13-25)	25 (17-33)	42 (32-52)
+MRI at 25	56 (56-56)	55 (55-55)	52 (50-53)	470 (435-505)	581 (533-629)	953 (892-1013)	19 (13-25)	25 (17-33)	42 (32-52)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

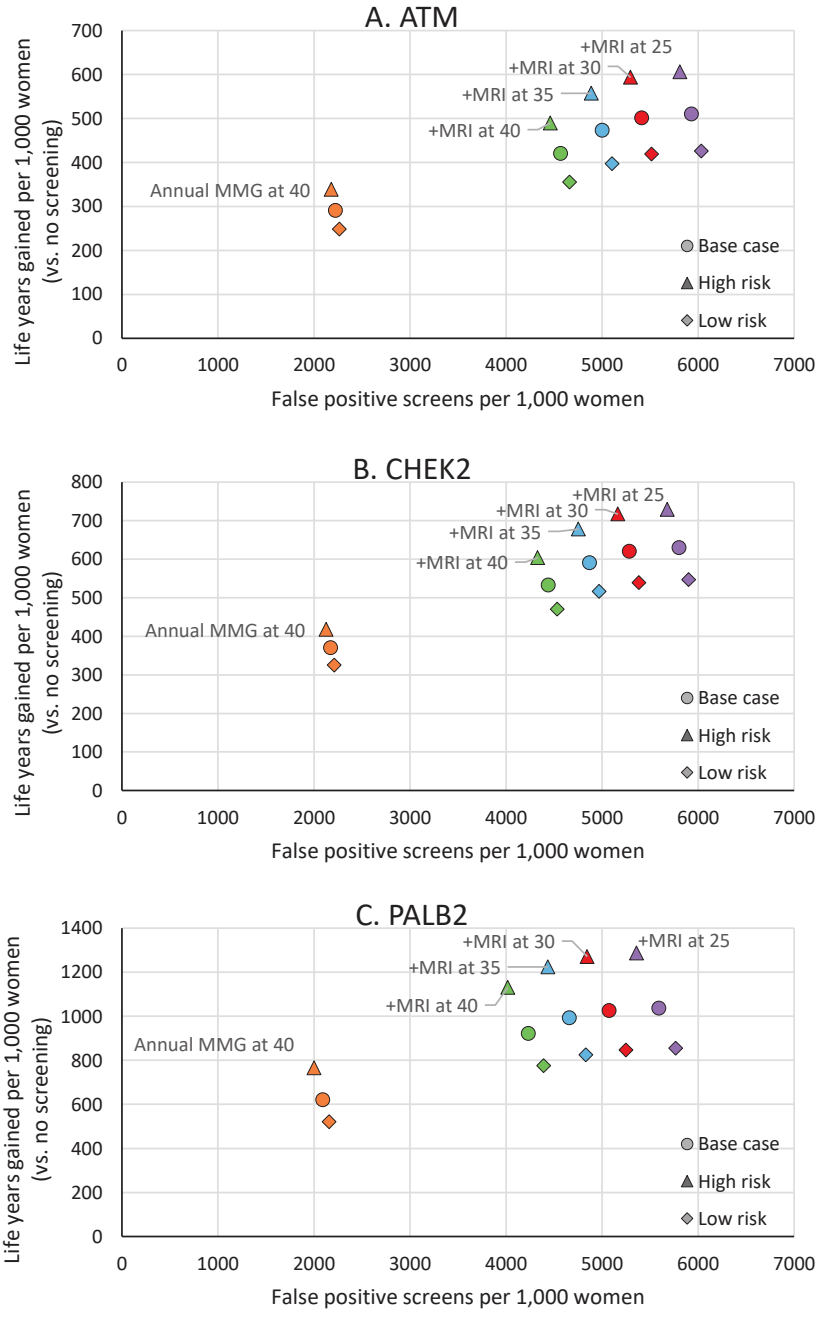


Figure S2. False-positive screens versus life years gained for screening strategies for women with pathogenic variants in ATM (panel A), CHEK2 (panel B), and PALB2 (panel C), varying breast cancer risk +/- one standard error based on CARRIERS data.

MMG=Mammography; MRI=Magnetic resonance imaging. Results are mean model projections across Model E and Model W-H. In all strategies, MMG is performed annually from ages 40-74; MRI varies in start age by strategy.

Table S8. Sensitivity analysis of false-positive screens and benign biopsies of screening strategies assuming the lower confidence limit of the MRI specificity.

	Breast cancer mortality reduction (%) Model Average (Range)			*Life years gained* Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	38.5 (37.8-39.2)	38.4 (38.0-38.8)	36.4 (34.6-38.2)	291 (263-319)	370 (330-409)	621 (559-684)
+MRI at 40	53.6 (52.9-54.3)	53.6 (53.3-53.9)	52.3 (51.4-53.1)	420 (388-452)	533 (489-577)	921 (876-967)
+MRI at 35	57.6 (57.2-58.0)	57.0 (56.3-57.7)	54.4 (54.2-54.7)	473 (447-498)	591 (555-627)	992 (959-1025)
+MRI at 30	59.5 (58.5-60.4)	58.4 (57.2-59.6)	55.4 (55.3-55.4)	501 (478-523)	620 (587-652)	1025 (998-1051)
+MRI at 25	60.2 (58.9-61.2)	58.9 (57.5-60.3)	55.7 (55.5-55.8)	510 (489-531)	630 (599-661)	1037 (1013-1061)

Table S8 continued

	False-positive screens Model Average (Range)			*Benign biopsies* Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	2224 (2222-2227)	2174 (2172-2175)	2092 (2085-2099)	296 (296-297)	290 (290-290)	279 (278-280)
+MRI at 40	4772 (4757-4787)	4638 (4636-4640)	4421 (4401-4441)	1249 (1245-1253)	1214 (1213-1215)	1157 (1152-1163)
+MRI at 35	5232 (5209-5255)	5096 (5086-5106)	4878 (4850-4905)	1432 (1425-1438)	1396 (1393-1399)	1339 (1332-1347)
+MRI at 30	5673 (5649-5696)	5536 (5499-5573)	5318 (5298-5337)	1601 (1589-1613)	1565 (1550-1581)	1508 (1498-1519)
+MRI at 25	6223 (6196-6249)	6086 (6073-6100)	5867 (5837-5898)	1811 (1819-1804)	1776 (1772-1780)	1719 (1710-1727)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

Table S9. Sensitivity analysis of false-positive screens and benign biopsies of screening strategies assuming the upper confidence limit of the MRI specificity.

	Breast cancer mortality reduction (%) Model Average (Range)			*Life years gained* Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	38.5 (37.8-39.2)	38.4 (38.0-38.8)	36.4 (34.6-38.2)	291 (263-319)	370 (330-409)	621 (559-684)
+MRI at 40	53.6 (52.9-54.3)	53.6 (53.3-53.9)	52.3 (51.4-53.1)	420 (388-452)	533 (489-577)	921 (876-967)
+MRI at 35	57.6 (57.2-58.0)	57.0 (56.3-57.7)	54.4 (54.2-54.7)	473 (447-498)	591 (555-627)	992 (959-1025)
+MRI at 30	59.5 (58.5-60.4)	58.4 (57.2-59.6)	55.4 (55.3-55.4)	501 (478-523)	620 (587-652)	1025 (998-1051)
+MRI at 25	60.2 (58.9-61.2)	58.9 (57.5-60.3)	55.7 (55.5-55.8)	510 (489-531)	630 (599-661)	1037 (1013-1061)

Table S9 continued

	False-positive screens Model Average (Range)			Benign biopsies Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	2224 (2222-2227)	2174 (2172-2175)	2092 (2085-2099)	296 (196-197)	290 (290-290)	279 (178-280)
+MRI at 40	4367 (4353-4380)	4244 (4242-4246)	4045 (4027-4064)	1141 (1138-1145)	1110 (1109-1110)	1058 (1053-1063)
+MRI at 35	4773 (4752-4794)	4649 (4640-4658)	4449 (4424-4474)	1303 (1298-1309)	1271 (1268-1274)	1219 (1212-1226)
+MRI at 30	5165 (5145-5186)	5040 (5008-5073)	4841 (4824-4858)	1454 (1444-1465)	1422 (1408-1435)	1370 (1361-1379)
+MRI at 25	5655 (5631-5680)	5531 (5519-5543)	5331 (5303-5358)	1642 (1635-1649)	1609 (1606-1613)	1557 (1550-1565)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

Table S10. Sensitivity analysis of false-positive screens and benign biopsies of screening strategies using age-specific specificity from the Breast Cancer Surveillance Consortium.

	Breast cancer mortality reduction (%) Model Average (Range)			*Life years gained* Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	38.5 (37.8-39.2)	38.4 (38.0-38.8)	36.4 (34.6-38.2)	291 (263-319)	370 (330-409)	621 (559-684)
+MRI at 40	53.6 (52.9-54.3)	53.6 (53.3-53.9)	52.3 (51.4-53.1)	420 (388-452)	533 (489-577)	921 (876-967)
+MRI at 35	57.6 (57.2-58.0)	57.0 (56.3-57.7)	54.4 (54.2-54.7)	473 (447-498)	591 (555-627)	992 (959-1025)
+MRI at 30	59.5 (58.5-60.4)	58.4 (57.2-59.6)	55.4 (55.3-55.4)	501 (478-523)	620 (587-652)	1025 (998-1051)
+MRI at 25	60.2 (58.9-61.2)	58.9 (57.5-60.3)	55.7 (55.5-55.8)	510 (489-531)	630 (599-661)	1037 (1013-1061)

Table S10 continued

	False-positive screens Model Average (Range)			Benign biopsies Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	1871 (1867-1875)	1837 (1835-1838)	1781 (1773-1788)	256 (256-257)	252 (251-252)	244 (243-245)
+MRI at 40	4225 (4209-4241)	4124 (411-4131)	3960 (3938-3981)	1108 (1104-1112)	1082 (1080-1084)	1039 (1033-1045)
+MRI at 35	4805 (4781-4828)	4702 (4687-4716)	4537 (4508-4565)	1344 (1337-1351)	1317 (1313-1321)	1274 (1266-1282)
+MRI at 30	5375 (5338-5412)	5271 (5224-5317)	5106 (5073-5138)	1563 (1546-1580)	1536 (1517-1555)	1493 (1477-1508)
+MRI at 25	6082 (6055-6109)	5978 (5960-5996)	5812 (5781-5844)	1834 (1826-1841)	1807 (1801-1812)	1764 (1755-1772)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

Table S11. Sensitivity analysis of false-positive screens and benign biopsies of screening strategies assuming the use of digital breast tomosynthesis for mammography screening.

	Breast cancer mortality reduction (%) Model Average (Range)			*Life years gained* Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	38.5 (37.8-39.2)	38.4 (38.0-38.8)	36.4 (34.6-38.2)	291 (263-319)	370 (330-409)	621 (559-684)
+MRI at 40	53.6 (52.9-54.3)	53.6 (53.3-53.9)	52.3 (51.4-53.1)	420 (388-452)	533 (489-577)	921 (876-967)
+MRI at 35	57.6 (57.2-58.0)	57.0 (56.3-57.7)	54.4 (54.2-54.7)	473 (447-498)	591 (555-627)	992 (959-1025)
+MRI at 30	59.5 (58.5-60.4)	58.4 (57.2-59.6)	55.4 (55.3-55.4)	501 (478-523)	620 (587-652)	1025 (998-1051)
+MRI at 25	60.2 (58.9-61.2)	58.9 (57.5-60.3)	55.7 (55.5-55.8)	510 (489-531)	630 (599-661)	1037 (1013-1061)

Table S11 continued

	False-positive screens Model Average (Range)			Benign biopsies Model Average (Range)		
	ATM	*CHEK2*	*PALB2*	*ATM*	*CHEK2*	*PALB2*
MMG at 40	1766 (1764-1768)	1726 (1725-1727)	1661 (1656-1667)	237 (236-237)	232 (231-232)	223 (222-224)
+MRI at 40	4263 (4250-4277)	4144 (4142-4146)	3950 (3932-3967)	1115 (1111-1118)	1083 (1083-1084)	1033 (1028-1038)
+MRI at 35	4827 (4681-4722)	4580 4571-4589)	4385 (4360-4410)	1288 (1283-1294)	1257 (1254-1259)	1206 (1199-1213)
+MRI at 30	5411 (5093-5140)	4995 (4959-5030)	4800 (4780-4819)	1448 (1437-1459)	1416 (1402-1431)	1365 (1355-1376)
+MRI at 25	5634 (5610-5658)	5513 (5525-5500)	5317 (5290-5344)	1646 (1639-1653)	1614 (1611-1688)	1563 (1556-1571)

MMG: mammography; MRI: magnetic resonance imaging.
Outcomes are shown as mean projections (ranges) per 1,000 women across Model E and Model W-H.

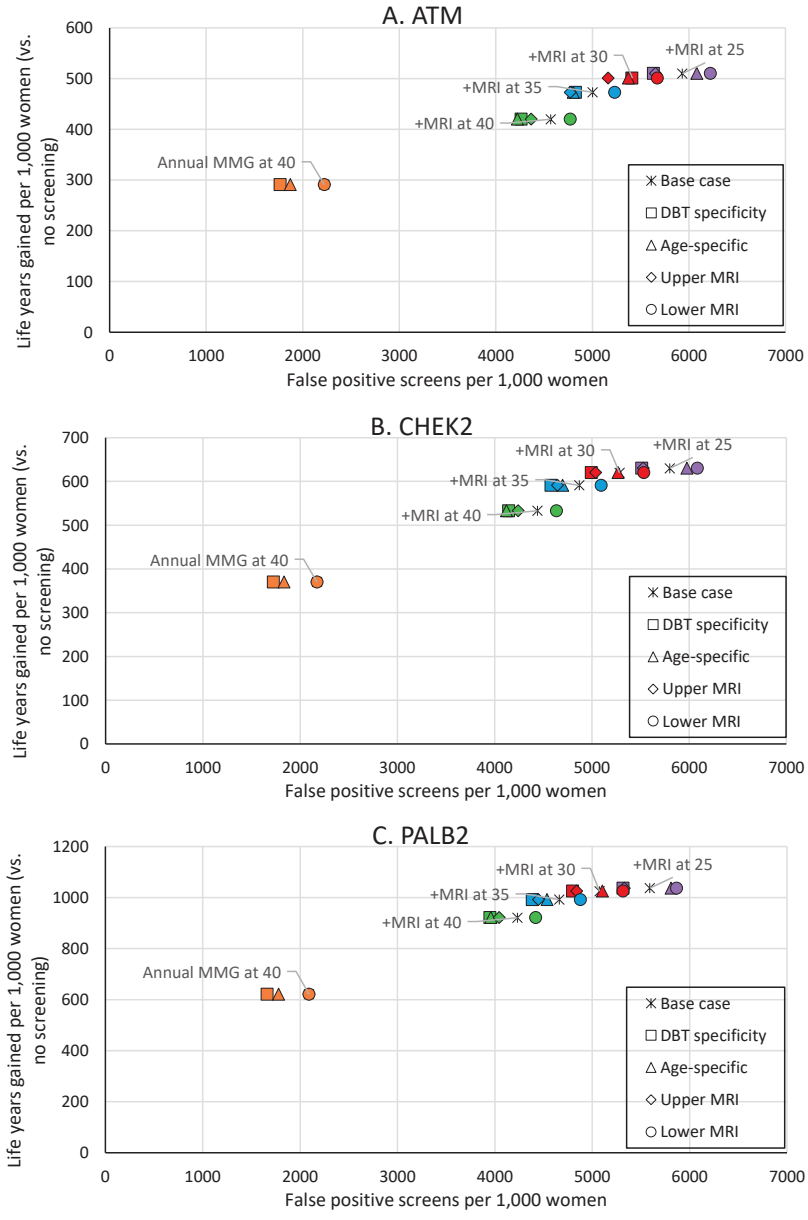


Figure S3. False-positive screens versus life years gained for screening strategies for women with pathogenic variants in ATM (panel A), CHEK2 (panel B), and PALB2 (panel C), under varying assumptions for screening specificity.
In the base case, screening specificity estimates were based on published data from the Ontario Breast Screening Program (OBSP). The upper and lower bounds of the 95% CI for MRI specificity were also considered. Specificity estimates from the Breast Cancer Surveillance Consortium (BCSC) stratified by age group and screening round were also considered. Improved mammography specificity due to digital breast tomosynthesis (DBT) was estimated based on published data from the PROSPR consortium. Results are mean model projections across Model E and Model W-H. MMG=Mammography; MRI=Magnetic resonance imaging. In all strategies, MMG is performed annually from ages 40-74; MRI varies in start age by strategy.

Additional funding and acknowledgements for CARRIERS

Funding/support

Additional support for the contributing studies was provided by NIH awards U01CA164974, R01CA098663, R01CA100598, R01CA185623, P01CA151135, R01 CA097396, P30CA16056, U01CA164973, U01CA164920, R01CA204819, K24CA194251 and K24CA194251-04S1, UL1TR002373, P30CA014520, U01CA82004, U01CA199277, P30CA033572, P30CA023100, UM1CA164917, R01CA077398, R01CA047147, R01CA067264, UM1CA186107, P01CA87969, R01CA49449, U01CA176726, R01CA58860, U01 CA58860, K07CA92044 and R01 CA67262; NHLBI contracts (HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C); NIEHS intramural awards (Z01-ES044005, Z01-ES049033 and Z01-ES102245); American Cancer Society; Susan G Komen for the Cure (JRP, SMD), Breast Cancer Research Foundation (FJC, CBA, JNW, SMD, KLN), Karin Grunebaum Cancer Research Foundation (JRP), the University of Wisconsin-Madison Office of the Vice Chancellor for Research and Graduate Education (ESB), California Breast Cancer Act of 1993, the California Breast Cancer Research Fund (contract 97-10500), California Department of Public Health, and the Lon V Smith Foundation [LVS39420]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Acknowledgements

California Teachers Study: The collection of cancer incidence data used in the California Teachers Study was supported by the California Department of Public Health pursuant to California Health and Safety Code Section 103885; Centers for Disease Control and Prevention's National Program of Cancer Registries, under cooperative agreement 5NU58DP006344; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201800032I awarded to the University of California, San Francisco, contract HHSN261201800015I awarded to the University of Southern California, and contract HHSN261201800009I awarded to the Public Health Institute. The opinions, findings, and conclusions expressed herein are those of the author(s) and do not necessarily reflect the official views of the State of California, Department of Public Health, the National Cancer Institute, the National Institutes of Health, the Centers for Disease Control and Prevention or their Contractors and Subcontractors, or the Regents of the University of California, or any of its programs.

The authors would like to thank the California Teachers Study Steering Committee that is responsible for the formation and maintenance of the Study within which this research was conducted. A full list of California Teachers Study team members is available at https://www.calteachersstudy.org/team.

WHI investigators: The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: https://www-whi-org.s3.us-west-2.amazonaws.com/wp-content/uploads/WHI-Investigator-Long-List.pdf

Program Office: (National Heart, Lung, and Blood Institute, Bethesda, Maryland) Jacques Rossouw, Shari Ludlam, Joan McGowan, Leslie Ford, and Nancy Geller

Clinical Coordinating Center: (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross Prentice, Andrea LaCroix, and Charles Kooperberg

Investigators and Academic Centers: (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Jennifer Robinson; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; (University of Nevada, Reno, NV) Robert Brunner

Women's Health Initiative Memory Study: (Wake Forest University School of Medicine, Winston-Salem, NC) Mark Espeland

NHS: The authors would like to thank the participants and staff of the NHS for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY.

Chapter 4

Decreasing breast density over time and the effect on performance of mammographic and supplemental MRI screening - Results of the DENSE trial

Geuzinge HA, Bakker MF, Mann RM, Veenhuizen SGA, Monninkhof EM, Loo CE, van Diest PJ, Lobbes MBI, Karssemeijer N, van der Zwaag J, de Koning HJ, Duvivier KM, Pijnappel RM, Veldhuis WB, van Gils CH, on behalf of the DENSE trial study group

To be submitted

ABSTRACT

Background: Women with extremely dense breasts and a negative screening mammogram were invited to participate in the DENSE trial to undergo supplemental magnetic resonance imaging (MRI) screening. After the first screening round, the same women were re-invited for two more MRI screening rounds, when having a negative mammogram. This was independent of their breast density.

Purpose: To investigate density changes in relation to participant characteristics, and the effect of density changes on performance of supplemental MRI screening.

Material and Methods: The DENSE trial is embedded within the Dutch biennial mammography screening program for women aged 50-74. Mammographic breast density was repeatedly measured using Volpara. We examined change in breast density in relation to baseline participant characteristics. Cancer detection, recall, and false-positives rates and positive predictive values (PPVs) of mammography and supplemental MRI, were compared between women who remained extremely dense and women who decreased to a lower density category.

Results: Of the 4,783 initial participants, 3,436 and 2,678 women underwent a second and third MRI, respectively. By the third round, 37% of the women had decreased to a lower density category. The strongest decreases in density were observed in the youngest age groups. In women whose breast density remained extremely dense versus women whose density decreased, mammography resulted in detection rates of 1.4 and 3.2 cancers per 1,000 screens (p=0.17) and a PPV of 8.1 and 21.7 (p=0.09), respectively. Supplemental MRI after a negative mammogram, resulted in an additional cancer detection of 6.8 and 4.9 per 1,000 screens (p=0.48) and a PPV of 22.7 and 17.1 (p=0.50), in the respective groups.

Conclusions: Despite an increase in cancer detection rate at mammography with decreasing density, supplemental MRI was of beneficial value in detecting additional cancer independent of breast density.

INTRODUCTION

In most Western countries, women at average breast cancer risk are offered mammography screening. Of those women, approximately 8% have extremely dense breast tissue,[1] which is an important risk factor for developing breast cancer.[2] At the same time, the sensitivity of mammographic screening is limited in women with dense breast tissue due to a masking effect, resulting in higher interval cancer rates compared to women with less dense glandular tissue.[1] Recently, results from the DENSE trial showed that women with extremely dense breast tissue benefit from additional MRI screening. In this trial, women with extremely dense breast tissue were offered additional MRI screening after a negative mammogram.[3] The results of the first screening round showed that undergoing supplemental MRI screening resulted in significantly less interval cancers.[3]

In the DENSE trial, all participants were invited for two additional biennial MRI screening rounds, regardless of changes in breast density. It is known that breast density decreases with age,[4,5] and that it is affected by menopausal status.[4] It remains unknown how changes in density between the first and subsequent rounds affect screening performance. In other studies that investigated supplemental MRI screening in women with dense breasts, the effect of density changes on the performance of supplemental imaging has not been described.[6,7] When considering the implementation of supplemental screening methods in a breast cancer screening program, it is important to study the effect of changes in density and to what extent this influences screening performance.

The aim of this study was twofold: 1) to investigate the extent of change in breast density over three screening rounds in DENSE trial participants, and whether this change was dependent on participant characteristics; and 2) to investigate the effect of decreasing breast density on the performance of incident mammography and MRI examinations in the DENSE trial.

METHODS

Design and study population

Longitudinal data of the first, second and the third screening round of the DENSE trial were used. The DENSE trial is a multicenter, randomized, controlled trial within the Dutch national screening program. In the Netherlands, women aged 50-74 years are invited for biennial digital mammography. Women with a negative mammogram and extremely dense breast tissue (defined as Volpara Density Grade 4) were eligible for participation in the trial. Women were randomized in a 1:4 ratio to the MRI invitation group (invited for supplemental MRI screening) and the control group (biennial mammography only).

Those who accepted the invitation and underwent MRI, are referred to as the MRI participants. Women who had a negative mammogram in the regular screening program two years after the first MRI, were re-invited for MRI screening, (round 2) regardless of their density at that point in time. [8] This process was repeated two years later (round 3).

Women with a Breast Imaging Reporting and Data System (BI-RADS) score 4 or 5 on MRI were recalled for additional diagnostic work-up. In case of a BI-RADS 3 score, independent double reading was performed. Women were re-invited for a follow-up MRI examination after six months if consensus was reached on a BI-RADS 3 score. This follow-up MRI had to be either negative (BI-RADS 1 or 2, with return to regular mammography screening) or positive (BI-RADS 4 or 5, after which women were recalled for additional work-up). MRI examinations were performed with 3.0-T MRI systems, using the macrocyclic gadolinium-based contrast agent gadobutrol (0.1 mmol per kilogram body weight) (Gadovist, Bayer). More details of the trial protocol have been published previously.[9]

Outcomes of mammography were obtained from screening units of the national screening program. To validate malignant findings by mammography, data were linked with the Netherlands Cancer Registry. Until now, data until February 2018 have been validated, so data on screening performance after February 2018, which is the last part of the third screening round, were not included in our analyses. Results of MRI screening were obtained from hospitals, and only results before February 2018 were included.

At baseline, MRI participants were asked to complete a questionnaire on lifestyle factors, reproductive factors and family history of breast cancer. Data on age were obtained along with screening data.

This study was approved by the Dutch Minister of Health, Welfare and Sport. Ethical approval was obtained on November 11, 2011. All MRI participants provided informed consent.

Breast density

As part of the DENSE trial, breast density was measured on mammography using Volpara imaging software, version 1.5 (Volpara Health Technologies).[10] This software provides volumetric percent breast density (VBD) and Volpara density grades (VDGs), which are based on the average VBD of both breasts (VDG1: 0%≤VBD<4.5%, VDG2: 4.5%≤VBD<7.5%, VDG3: 7.5%≤VBD<15.5%, VDG4: ≥15.5%) and correspond with the BI-RADS breast density categories (A-D).[11] Breast density was measured every screening round.

Statistical analyses

To describe changes in breast density, we provided distributions of VDG in each mammography screening round, using data of all women in the trial (n=40,373).

Furthermore, median VBD was calculated by screening round, by age groups, and by VDG. To evaluate whether changes in VBD over time differed by levels of VBD at baseline, we divided baseline VBD in quartiles and compared differences using the Jonckheere Terpstra test. All other analyses were restricted to the MRI participants (n=4,783).

To identify determinants of changing breast density, baseline characteristics were compared between women whose density remained high (VDG4) and women whose density decreased (<VDG4) going from the first to the third round. Baseline characteristics consisted of age, body mass index (BMI), menopausal status, hormone replacement therapy use, family history, parity, age at first live birth, and alcohol use. BMI and age were categorized using commonly used BMI categories (<18.5; 18.5-24.9; ≥25 kg/m2) and 5-year age groups, respectively. The associations between baseline characteristics and density (VDG4 versus <VDG4) were compared using univariable logistic regression, and bivariable logistic regression adjusting for age. A *p*-value <0.05 was considered statistically significant.

Screening performance measures consisted of cancer detection rates, recall rates, false positive rates and positive predictive value (PPV) for mammography and supplemental MRI separately. Both invasive breast cancers and ductal carcinoma in situ (DCIS) were included as cancer cases. Definitions of performance measures can be found in **Table 3**.[3] All screening performance measures were compared between women whose breast density remained extremely high (VDG4) and women whose breast density decreased to a lower category (<VDG4). Furthermore, proportions of cancers detected with mammography and cancers detected with MRI were compared between the density categories (VDG4 versus <VDG4) . Comparisons were made using chi-square test, or Fisher's exact test in case of low cell counts. Results of the second and third screening rounds were presented separately and combined, using generalized estimating equations (GEE). This was done to account for correlation between screening examinations in the same woman, using a working independence correlation structure.[12] In additional analyses, we included age in the GEE models to adjust for a possible confounding effect.

Analyses were performed using RStudio version 1.2.5001.

RESULTS

Study population

Between December 2011 and November 2015, 4,783 MRI participants were enrolled in the first screening round of the trial. In March 2020, 2,678 MRI participants had completed all three screening rounds consisting of mammography with supplemental MRI. **Figure S1** shows a flow chart of the DENSE trial. **Table 1** shows the numbers of

mammograms and MRIs by screening round and by VDG. In the second round, for 184 out of 4,381 mammograms (4%), VDG was not measured. In the third round, for 378 out of 3,251 mammograms (12%), VDG was not measured. This was caused by screening units who replaced their mammography machines after which Volpara software was no longer available, and by women who relocated and went to a screening unit without Volpara software. A total of 4,509 MRI participants completed the baseline questionnaire, equaling to a response rate of 94%. Their median age was 54.2 (IQR: 51.3-59.7), and 60% (n=2768/4578) of the MRI participants were postmenopausal at baseline (**Table S1**).

Breast density over time

Of all women in the DENSE trial (intervention and control arm, n=40,373), 73% (n=24,902) of the women was still extremely dense (VDG4) in the second round, and 62% (n=17,243) was still extremely dense in the third round. The remaining women decreased to VDG3 (27% in the second (n=9314) and 37% in the third round (n=10,263)), or VDG2 (0.07% in the second (n=23) round and 0.4% in the third round (n=110)). None of the women decreased to VDG1 (**Table 1**). Similar distributions were shown in the subgroup of women that underwent supplemental MRI screening (the MRI participants). Women who remained at VDG4 over the three screening rounds had a higher median VBD at baseline (**Figure 1**) compared to women who decreased to VDG3 or VDG2 (**Figure 1, Table S2**). Overall, the median VBD in women who decreased to VDG3/2, was close to the upper bound of VDG3 (VDG3: 7.5%≤ VBD <15.5%).

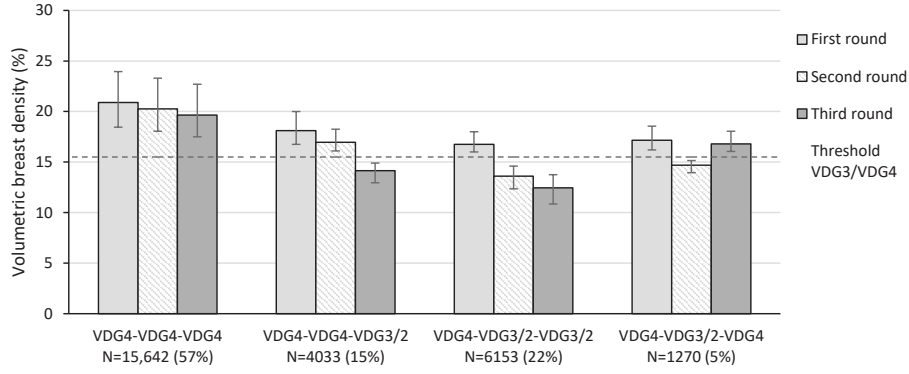


Figure 1. Median VBD (with IQR) by VDG-pattern over time

Median VBD with interquartile ranges. VBD: volumetric breast density, VDG: Volpara density grade, VDG3: 7.5% ≤ VBD < 15.5%, VDG4: ≥ 15.5%

I: women who remained VDG4 (VDG4-VDG4-VDG4)

II: women who decreased to VDG3/2 in the third screening round (VDG4-VDG4-VDG3/2)

III: women who decreased to VDG3/2 in the second screening round and who remained VDG3/2 in the third round (VDG4-VDG3/2-VDG3/2)

IV: women who decreased to VDG3/2 in the second screening round and who returned to VDG4 remained the third round (VDG4-VDG3/2-VDG4)

Table 1. Distribution of breast density categories at each screening round of the DENSE trial

	All women in the DENSE trial						MRI participants				
	Total	**VDG4**	**VDG3**	**VDG2**	**VDG missing**		**Total**	**VDG4**	**VDG3**	**VDG2**	**VDG missing**
First round						**First round**					
No. of women with mammograms	40,373	40,373 (100%)	n/a	n/a	n/a	No. of women with mammograms	4,783	4783	n/a	n/a	n/a
						No. of women with MRIs	4,783	4783	n/a	n/a	n/a
Second round						**Second round**					
No. of women with mammograms	35,723	24,902 (73%)	9,314 (27%)	23 (0.07%)	1484	No. of women with mammograms	4,381	3,030 (72%)	1,164 (28%)	3 (0.07%)	184
						No. of women with MRIs	3,436	2,391 (72%)	919 (28%)	2 (0.06%)	124
Third round*						**Third round***					
No. of women with mammograms	31,289	17,243 (62%)	10,263 (37%)	110 (0.4%)	3673	No. of women with mammograms	3,251	1,773 (62%)	1,089 (38%)	11 (0.4%)	378
						No. of women with MRIs	2,678	1,487 (62%)	885 (37%)	9 (0.4%)	297
Validated data of the third round†						**Validated data of the third round†**					
No. of women with mammograms	10,703	6,491 (64%)	3,690 (36%)	35 (0.3%)	487	No. of women with mammograms	1,105	667 (63%)	392 (37%)	4 (0.4%)	42
						No. of women with MRIs	875	541 (64%)	298 (35%)	3 (0.4%)	33

*Here we only included women who completed the second mammogram and MRI before attending the third round. 46 women who participated in the third round, but who skipped the second round were thus excluded.

† Of the third round, only data until February 2018 were validated for the occurrence of breast cancer. VDG missings were caused by women receiving mammography screening in units where no Volpara software was installed.

Figure 2 shows that women who were younger than 52 years of age in the first round showed the highest median baseline VBD and the largest decrease in breast density over time. The median baseline VBD was lower in groups of higher ages and the decrease was smaller. Women aged 65 years and older during the first round did not show a decrease in breast density over time. When dividing VBD at baseline in quartiles, higher baseline levels of breast density were associated with a slightly larger decrease in VBD over time (first quartile: -2.0 percentage point, fourth quartile: -3.3 percentage point, p<0.001, **Table S3**).

Besides age, also BMI, menopausal status and parity were statistically significantly associated with a decrease in breast density (**Table 2**). Women with a relatively high BMI (≥25) had a higher odds of a decrease in breast density (adjusted OR: 2.44 (95% CI: 1.53-3.97)) compared to women with a relatively low BMI (<18.5). Compared to women being premenopausal, women being postmenopausal had a somewhat lower odds of a decrease in breast density (adjusted OR: 0.81 (95% CI: 0.61-1.08), and women being perimenopausal had a somewhat higher odds of a decrease in breast density (adjusted OR: 1.30 (95% CI: 0.99-1.70). Having two or more children also resulted in a statistically significantly higher odds of a decrease in breast density, compared to women without children (adjusted OR 1.43 (1.18-1.73)). Women who were ≥30 years when they gave birth for the first time, had a statistically significantly lower odds of a decrease in breast density compared to women who were aged <25 year at first live birth (adjusted OR: 0.69 (95% CI 0.54-0.89).

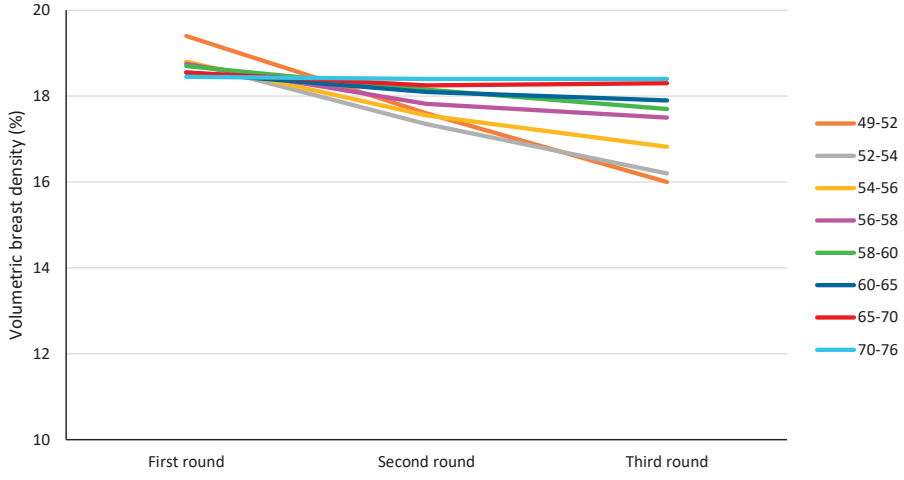


Figure 2. Median VBD over time, by age category in the first round
VBD: volumetric breast density

Table 2. Characteristics of MRI participants at baseline, compared between women whose breast density was still extremely dense (VDG4) in the third screening round and women whose VDG had decreased

MRI participants	VDG4*	VDG3/2*	OR for decrease in VDG	*P* value	Adjusted OR† for decrease in VDG	*P* value
Median age, yrs (IQR)	55 (52-60)	53 (51-57)				
49-55	902 (51%)	705 (64%)	1 (Ref)		n/a	n/a
55-59	409 (23%)	224 (20%)	0.70 (0.58-0.86)	<0.001	n/a	n/a
60-64	285 (16%)	98 (9%)	0.44 (0.34-0.56)	<0.001	n/a	n/a
65-76	177 (10%)	73 (7%)	0.53 (0.39-0.71)	<0.001	n/a	n/a
Quartiles of VBD						
Q1	217 (12%)	496 (28%)	1 (Ref)		1 (Ref)	
Q2	342 (20%)	347 (20%)	0.44	<0.001	0.41 (0.32-0.51)	<0.001
Q3	515 (29%)	182 (10%)	0.15	<0.001	0.13 (0.10-0.16)	<0.001
Q4	672 (38%)	71 (4%)	0.05	<0.001	0.04 (0.03-0.05)	<0.001
Missing	27	4				
Median BMI, kg/m2 (IQR)	22 (20-23)	22 (21-24)				
<18.5	82 (5%)	31 (3%)	1 (Ref)		1 (Ref)	
18.5-24.9	1,453 (86%)	842 (82%)	1.53 (1.02-2.37)	0.05	1.43 (0.94-2.22)	0.10
≥25	154 (9%)	158 (15%)	2.71 (1.71-4.39)	<0.001	2.44 (1.53-3.97)	<0.001
Missing	84	69				
Menopausal status‡						
Premenopausal	165 (10%)	120 (11%)	1 (Ref)		1 (Ref)	
Perimenopausal	446 (26%)	414 (39%)	1.28 (0.97-1.68)	0.08	1.30 (0.99-1.70)	0.06
Postmenopausal	1,115 (65%)	523 (50%)	0.64 (0.50-0.84)	<0.001	0.81 (0.61-1.08)	0.15
Missing	47	43				
HRT use						
Never	1,248 (73%)	771 (74%)	1 (Ref)		1 (Ref)	
In the past	219 (13%)	115 (11%)	0.85 (0.67-1.08)	0.20	1.02 (0.80-1.32)	0.86
Currently	238 (14%)	158 (15%)	1.07 (0.86-1.34)	0.52	0.94 (0.75-1.17)	0.58
Missing	68	56				
First degree family member bc history						
Yes	295 (27%)	172 (25%)	1 (Ref)		1 (Ref)	
No	779 (73%)	526 (75%)	1.16 (0.93-1.44)	0.19	1.13 (0.91-1.41)	0.28
Missing	699	402				
Parity						
Nulliparous	429 (25%)	201 (19%)	1 (Ref)		1 (Ref)	
1 birth	198 (12%)	116 (11%)	1.25 (0.94-1.66)	0.15	1.25 (0.94-1.66)	0.13
≥2 births	1,078 (63%)	727 (70%)	1.44 (1.19-1.75)	<0.001	1.43 (1.18-1.73)	<0.001
missing	68	56				

MRI participants	VDG4*	VDG3/2*	OR for decrease in VDG	*P* value	Adjusted OR† for decrease in VDG	*P* value
Age at first live birth, yrs§						
<25	266 (21%)	188 (22%)	1 (Ref)		1 (Ref)	
25-29	528 (42%)	367 (44%)	0.98 (0.78-1.23)	0.89	0.87 (0.69-1.10)	0.24
≥30	476 (37%)	284 (34%)	0.84 (0.67-1.07)	0.16	0.69 (0.54-0.89)	0.004
Missing	6	4				
Alcohol consumption						
No	204 (12%)	133 (13%)	1 (Ref)		1 (Ref)	
<1 glass per week	391 (23%)	303 (29%)	1.19 (0.91-1.55)	0.20	1.17 (0.90-1.54)	0.25
≥1 glass per week	1106 (65%)	608 (58%)	0.84 (0.66-1.07)	0.16	0.87 (0.69-1.12)	0.27
Missing	72	56				

VBD= volumetric breast density; Q1=minimal value – 25th percentile; Q2=25th percentile – median; Q3=median-75th percentile; Q4=75th percentile-maximum value; bc= breast cancer; HRT= hormone replacement therapy

* VDG in the third screening round; † Adjusted for age; ‡ Women aged ≥60 years, or reporting having had a hysterectomy or bilateral oophorectomy or women reporting 0 periods within last 12 months without use of hormonal contraceptives were categorized as postmenopausal. Women reporting regular periods (12-18 times in last 12 months) without use of hormonal contraceptives were categorized as premenopausal. All other women were categorized as perimenopausal; § if given live birth

Performance of mammography and MRI

During the second round, nine cancers were detected by mammography, and another 20 cancers were found by supplemental MRI after a negative mammogram. In the third round (data until February 2018), one cancer was detected by mammography and six by supplemental MRI. **Table 3** shows the outcomes of mammography and supplemental MRI screening, stratified by women who remained extremely dense (VDG4) and women who decreased to a lower density category (VDG3/2).

Comparing all performance measures of mammography and MRI between women whose breast density remained extremely high versus women whose density decreased, only small differences were found that were not statistically significant (**Table 3** and **Table S4**). The cancer detection rate of mammography was slightly lower in women whose density remained extremely high versus those whose density decreased (VDG4: 1.4, VDG3/2: 3.2, p=0.17), and the PPV in women whose density remained extremely high was lower compared to those whose density decreased (VDG4: 8.1, VDG3/2: 21.7, p=0.09). The cancer detection rate of MRI was slightly higher in women who remained extremely dense than in women whose density decreased (VDG: 6.8, VDG3/2: 4.9, p=0.48), which was also the case for the PPV (VDG4: 22.7, VDG3/2: 17.1, p=0.50). Adjusting the GEE models for age did not change the differences between the groups (**Table S4**).

Table 3. Performance of mammography and MRI by VDG and screening round

	VDG4		VDG3/2		
	No. /total no.	**Rate (95%CI)/ 1,000 screens**	**No. /total no.**	**Rate (95%CI)/ 1,000 screens**	***P* value**
Second round – Mammography					
Cancer detection	5/3,030	1.7 (0.7-3.9)	4/1,167	3.4 (1.3-8.8)	0.27
Recall[†]	52/3,030	17.2 (13.1-22.4)	18/1,167	15.4 (9.8-24.2)	0.80
False positives[‡]	47/3,025	15.5 (11.7-20.6)	14/1,163	12.0 (7.2-20.1)	0.48
PPV[§]	5/52	9.6 (4.2-20.6)	4/18	22.2 (9.0-45.2)	0.22
Second round – MRI					
Cancer detection	15/2,391	6.3 (3.8-10.3)	5/921	5.4 (2.3-12.6)	0.94
Recall[†]	78/2,391	32.6 (26.2-40.5)	29/921	31.5 (22.0-44.9)	0.96
False positives[‡]	63/2,376	26.4 (20.6-33.6)	24/916	26.1 (17.6-38.5)	1.00
PPV[§]	15/78	19.2 (12.0-29.3)	5/29	17.2 (7.6-34.5)	1.00
Third round – Mammography[‖]					
Cancer detection	0/667	0.0 (0.0-5.7)	1/396	2.5 (0.4-14.2)	0.37
Recall[†]	10/667	15.0 (8.2-27.4)	5/396	12.6 (5.4-29.2)	0.96
False positives[‡]	10/667	15.0 (8.2-27.4)	4/395	10.1 (3.9-25.7)	0.69
PPV[§]	0/10	0.0 (0.0-27.8)	1/5	20.0 (3.6-62.4)	0.33
Third round – MRI[‖]					
Cancer detection	5/541	9.2 (4.0-21.5)	1/301	3.3 (0.6-18.6)	0.43
Recall[†]	10/541	18.5 (10.1-33.7)	6/301	19.9 (9.2-42.8)	1.00
False positives[‡]	5/536	9.3 (4.0-21.6)	5/300	16.7 (7.1-38.4)	0.34
PPV[§]	5/10	50.0 (23.7-76.3)	1/6	16.7 (3.0-56.4)	0.31
Both screening rounds – Mammography[‖]					
Cancer detection	5/3,697	1.4 (0.6-3.3)	5/1,563	3.2 (1.3-7.7)	0.17
Recall[†]	62/3,697	16.8 (13.1-21.5)	23/1,563	14.7 (9.8-22.1)	0.59
False positives[‡]	57/3,692	15.4 (11.9-20.0)	18/1,558	11.5 (7.3-18.3)	0.28
PPV[§]	5/62	8.1 (3.4-18.0)	5/23	21.7 (9.4-42.8)	0.09
Both screening rounds – MRI[‖]					
Cancer detection	20/2,932	6.8 (4.4-10.5)	6/1,222	4.9 (2.2-10.9)	0.48
Recall[†]	88/2,932	30.0 (24.4-36.8)	35/1,222	28.7 (20.7-39.7)	0.82
False positives[‡]	68/2,912	23.4 (18.5-29.5)	29/1,216	23.9 (16.6-34.1)	0.92
PPV[§]	20/88	22.7 (15.2-32.6)	6/35	17.1 (7.9-33.3)	0.50

PPV= positive predictive value

[†] The recall rates were calculated by the number of women who were recalled divided by the total number of screens.

[‡] False-positive rates were calculated by dividing the number of women who were recalled and did not have breast cancer, by the number of women who underwent the respective screening test and did not have a screen-detected cancer.

[§] The PPV was defined as the number of screen-detected cancers by the number of positive screens.

[‖] Data not validated were excluded.

Of the 8.2 (1.4+6.8) detected tumors per 1,000 women whose breast density remained extremely high, 83% (6.8/(1.4+6.8)) were detected with MRI, whereas in women whose density decreased, 60% (4.9/(3.2+4.9)) of the detected cancers were found with MRI.

DISCUSSION

After three consecutive screening rounds, 62% of the women with extremely dense breasts (VDG4) were still in this upper density category. The largest decrease in VBD was seen in the youngest women. Furthermore, perimenopausal status, a high BMI, having two or more children and an age below 30 years at first live birth were significantly associated with a decrease in breast density. In women whose breast density decreased to VDG3/2, mammography resulted in a slightly higher cancer detection rate compared to women whose density remained extremely high, but this difference was not statistically significant. Supplemental MRI resulted in a slightly lower cancer detection rate in women whose breast density decreased compared to those whose breasts remained extremely dense but also not statistically significant. Despite these difference in detection rates, the majority of cancers was detected with MRI and not with mammography in both women whose breast density remained extremely high and women whose breast density decreased.

Our results on declining breast density by age are in line with published literature[13,14] Previous research also suggested that women with a higher initial breast density had a faster rate of decreasing breast density,[14-16] and that older women had a lower rate of decreasing breast density.[17] We showed that older women and postmenopausal women who still have extremely dense breasts, have a low chance of going to a lower density category over time. This may be useful for participants in terms of expectation management, but also to estimate the screening capacity needed.

To the best of our knowledge, no previous longitudinal studies evaluated the performance of additional MRI screening in women who are on an MRI screening scheme and whose breast density may have decreased during the study. Previous longitudinal studies on MRI screening and breast density showed breast density at baseline only, and cross-sectional studies combined density categories (ACR breast density categories A+B; C+D).[18-21] A study by Kuhl et al. showed that supplemental MRI screening resulted in 26 detected cancer among 811 women with density category C (32/1,000), and 11 cancers in 471 women with density category D (23/1,000).[6] They only measured breast density at baseline and they did not show the number of screening examinations per density category, but their results showed that supplemental MRI yields substantial supplemental cancer detection in both density categories.[6]

An important strength of this study is the use of a large longitudinal dataset of women with extremely dense breast tissue. A limitation is the fact that the numbers of interval cancers were unknown. However, we expect the numbers of interval cancers to be low in both density groups, as only four (0.8/1,000) interval cancers occurred in the first round.[3] Furthermore, by applying density categories, large and small differences within categories are averaged out. Women with a VBD level close to the lower bound of VDG4 may not differ much from women with a VBD level close to the upper bound of VDG3, but they were stratified into different categories in our analyses.

In the DENSE trial, only women who had a negative mammography result were invited for supplemental MRI screening. This should be kept in mind when comparing our performance outcomes with trials in which women undergo both screening modalities. Due to our study design, we do not know which of the mammography-detected tumors would also have been detected with MRI.

Furthermore, comparing our mammography performance measures with that of mammography within the national screening program should be done with caution. All women who underwent mammography in the second (and third) round of the DENSE trial, had already undergone a previous screening MRI in the first round. Therefore, for example detection rates of mammography are likely to be lower in our study compared to women attending mammography screening who were not included in the DENSE trial. A previous study showed a cancer detection rate of mammography of 5.6 per 1000 screens in women with VDG4, which is higher compared to our results.[1]

In conclusion, younger women with extremely dense breasts showed the highest chance of a decrease in breast density over time. Furthermore, women with a high BMI, higher parity and women who were perimenopausal were more likely to have declining breast density as opposed to women with a low BMI, low parity and being postmenopausal. No large statistically significant differences in screening performance measures of both mammography and MRI were found between women whose density remained extremely dense and women whose density decreased to a lower category. The performance of mammography was slightly better in women whose breast density decreased to heterogeneously dense breasts, compared to women whose breasts remained extremely dense, but the differences were small. In women whose breast density decreased, MRI resulted in a slightly non-significantly lower cancer detection rate compared to women whose breast density remained extremely dense. However, in both density groups, the majority of the cancers was detected with MRI and not with mammography.

Acknowledgements

We thank the trial participants for their contributions; the regional screening organizations, Volpara Health Technologies, the Dutch Expert Center for Screening, and

the National Institute for Public Health and the Environment for their advice and in-kind contributions. We thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry.

REFERENCES

1. Wanders JO, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat* 2017; 162(1): 95-103.
2. Boyd NF, Rommens JM, Vogt K, et al. Mammographic breast density as an intermediate phenotype for breast cancer. *Lancet Oncol* 2005; 6(10): 798-808.
3. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med* 2019; 381(22): 2091-102.
4. Burton A, Maskarinec G, Perez-Gomez B, et al. Mammographic density and ageing: A collaborative pooled analysis of cross-sectional data from 22 countries worldwide. *PLoS Med* 2017; 14(6): e1002335.
5. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. *AJR Am J Roentgenol* 2012; 198(3): W292-5.
6. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women with Average Risk of Breast Cancer. *Radiology* 2017; 283(2): 361-70.
7. Kwon MR, Choi JS, Won H, et al. Breast Cancer Screening with Abbreviated Breast MRI: 3-year Outcome Analysis. *Radiology* 2021; 299(1): 73-83.
8. Veenhuizen SGA, de Lange SV, Bakker MF, et al. Supplemental Breast MRI for Women with Extremely Dense Breasts: Results of the Second Screening Round of the DENSE Trial. *Radiology* 2021: 203633.
9. Emaus MJ, Bakker MF, Peeters PH, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. *Radiology* 2015; 277(2): 527-37.
10. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. *IEEE Trans Med Imaging* 2006; 25(3): 273-82.
11. D'Orsi C, Sickles E, Mendelson E, Morris E. ACR BI-RADS atlas: breast imaging reporting and data system. 5th edition. Reston, VA: American College of Radiology, 2013.
12. Pepe MS, Anderson GL. A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in Statistics - Simulation and Computation* 1994; 23(4): 939-51.
13. McCormack VA, Perry NM, Vinnicombe SJ, Dos Santos Silva I. Changes and tracking of mammographic density in relation to Pike's model of breast tissue aging: a UK longitudinal study. *Int J Cancer* 2010; 127(2): 452-61.
14. Maskarinec G, Pagano I, Lurie G, Kolonel LN. A longitudinal investigation of mammographic density: the multiethnic cohort. *Cancer Epidemiol Biomarkers Prev* 2006; 15(4): 732-9.
15. Kelemen LE, Pankratz VS, Sellers TA, et al. Age-specific trends in mammographic density: the Minnesota Breast Cancer Family Study. *Am J Epidemiol* 2008; 167(9): 1027-36.
16. Liao YS, Zhang JY, Hsu YC, Hong MX, Lee LW. Age-Specific Breast Density Changes in Taiwanese Women: A Cross-Sectional Study. *Int J Environ Res Public Health* 2020; 17(9).
17. Lokate M, Stellato RK, Veldhuis WB, Peeters PH, van Gils CH. Age-related changes in mammographic density and breast cancer risk. *Am J Epidemiol* 2013; 178(1): 101-9.
18. Saadatmand S, Geuzinge HA, Rutgers EJT, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol* 2019; 20(8): 1136-47.

19. Comstock CE, Gatsonis C, Newstead GM, et al. Comparison of Abbreviated Breast MRI vs Digital Breast Tomosynthesis for Breast Cancer Detection Among Women With Dense Breasts Undergoing Screening. *JAMA* 2020; 323(8): 746-56.
20. Kriege M, Brekelmans CT, Obdeijn IM, et al. Factors affecting sensitivity and specificity of screening mammography and MRI in women with an inherited risk for breast cancer. *Breast Cancer Res Treat* 2006; 100(1): 109-19.
21. Berg WA, Zhang Z, Lehrer D, 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.

SUPPLEMENTARY APPENDIX

DENSE trial study group

University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands:
CH van Gils PhD, MF Bakker PhD, SV de Lange MD, SGA Veenhuizen MSc, WB Veldhuis MD PhD, RM Pijnappel MD PhD, MJ Emaus PhD, PHM Peeters MD PhD, EM Monninkhof PhD, MA Fernandez-Gallardo MD, WPThM Mali MD PhD, MAAJ van den Bosch MD PhD, PJ van Diest MD PhD

Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands:
RM Mann MD PhD, Roel Mus MD, N Karssemeijer PhD, M. Imhof-Tas MD

The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands:
CE Loo MD PhD, PK de Koekkoek-Doll MD, HAO Winter-Warnars MD PhD

Albert Schweitzer Hospital, Dordrecht, the Netherlands:
RHC Bisschops MD PhD, MCJM Kock MD PhD, RK Storm MD, PHM van der Valk MD

Maastricht University Medical Centre, Maastricht, the Netherlands:
MBI Lobbes MD PhD, S Gommers MD

Jeroen Bosch Hospital, 's-Hertogenbosch, the Netherlands:
MDF de Jong MD, MJCM Rutten MD PhD

Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands:
KM Duvivier MD, P de Graaf MD PhD

Hospital Group Twente (ZGT), Almelo, the Netherlands:
J Veltman MD PhD, RLJH Bourez MD

Erasmus Medical Center, Rotterdam, the Netherlands:
HJ de Koning MD PhD

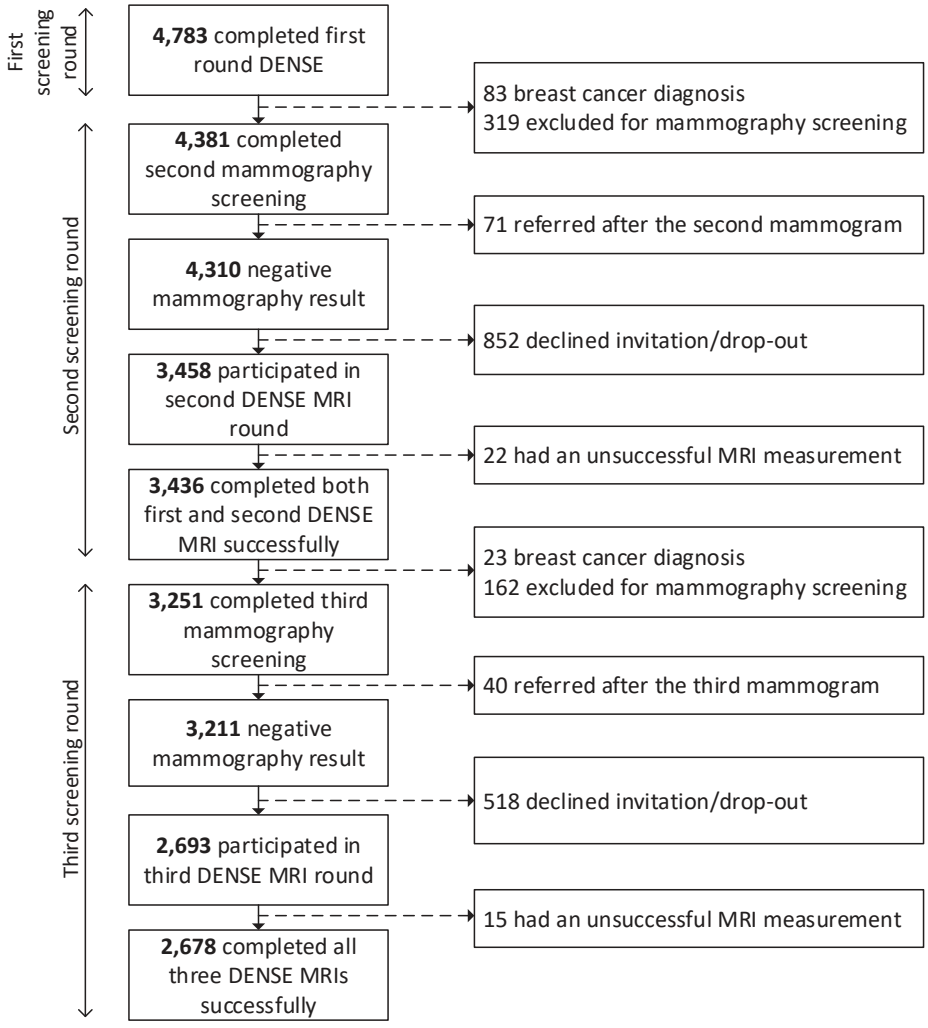


Figure S1. Flow chart of the DENSE trial

Table S1. Characteristics of MRI-participants at baseline (first screening round)

MRI-participants (n=4783)*	
Median age, yrs (IQR)	54.2 (51.3-59.7)
49-54	2,637 (47%)
55-59	991 (21%)
60-64	603 (13%)
65-69	407 (9%)
70-76	145 (3%)
Median BMI, kg/m2 (IQR)	22.0 (20.5-23.6)
<18.5	216 (5%)
18.5-24.9	3,694 (83%)
≥25	535 (12%)
Missing	338
Menopausal status†	
Premenopausal	473 (10%)
Perimenopausal	1,337 (29%)
Postmenopausal	2,768 (60%)
Missing	205
Hormone replacement therapy	
never	3,325 (74%)
in the past	550 (12%)
currently	623 (14%)
missing	285
First degree family member breast cancer history	
yes	746 (26%)
no	2,081 (74%)
missing	1956
Parity	
Nulliparous	1,051 (23%)
1 birth	540 (12%)
≥2 births	2,908 (65%)
Missing	284
Age at first live birth, yrs‡	
<25	804 (23%)
25-29	1,406 (41%)
≥30	1,221 (36%)
Missing	18
Alcohol consumption	
No	584 (13%)
<1 glass per week	1,155 (26%)
≥1 glass per week	2,754 (61%)
Missing	290

*274 women did not complete the questionnaire on baseline characteristics

† Women aged ≥60 years, or reporting having had a hysterectomy or bilateral oophorectomy or women reporting 0 periods within last 12 months without use of hormonal contraceptives were categorized as postmenopausal. Women reporting regular periods (12-18 times in last 12 months) without use of hormonal contraceptives were categorized as premenopausal. All other women were categorized as perimenopausal.

‡ if given live birth

Table S2. Median VBD by VDG-pattern over time

	VDG classification			Median VBD (IQR)		
No. women (n=27,616)	**Round 1**	**Round 2**	**Round 3**	**Round 1**	**Round 2**	**Round 3**
15,642 (58%)	VDG4	VDG4	VDG4	20.90 (18.5-24.0)	20.25 (18.1-23.3)	19.65 (17.5-22.7)
4,033 (15%)	VDG4	VDG4	VDG3/2	18.10 (16.8-20.0)	16.95 (16.1-18.3)	14.15 (13.0-14.9)
6,153 (23%)	VDG4	VDG3/2	VDG3/2	16.75 (16.0-18.0)	13.60 (12.4-14.6)	12.45 (10.9-13.8)
1,270 (5%)	VDG4	VDG3/2	VDG4	17.15 (16.2-18.6)	14.70 (14.0-15.2)	16.80 (16.1-18.1)
518*		Missing				

*Missings were caused by a missing VDG classification in the second round
VDG= Volpara density grade
VBD= volumetric breast density
VDG3= 7.5% ≤ VBD < 15.5%, VDG4: ≥ 15.5%

Table S3. Median VBD over time, by quartiles at baseline

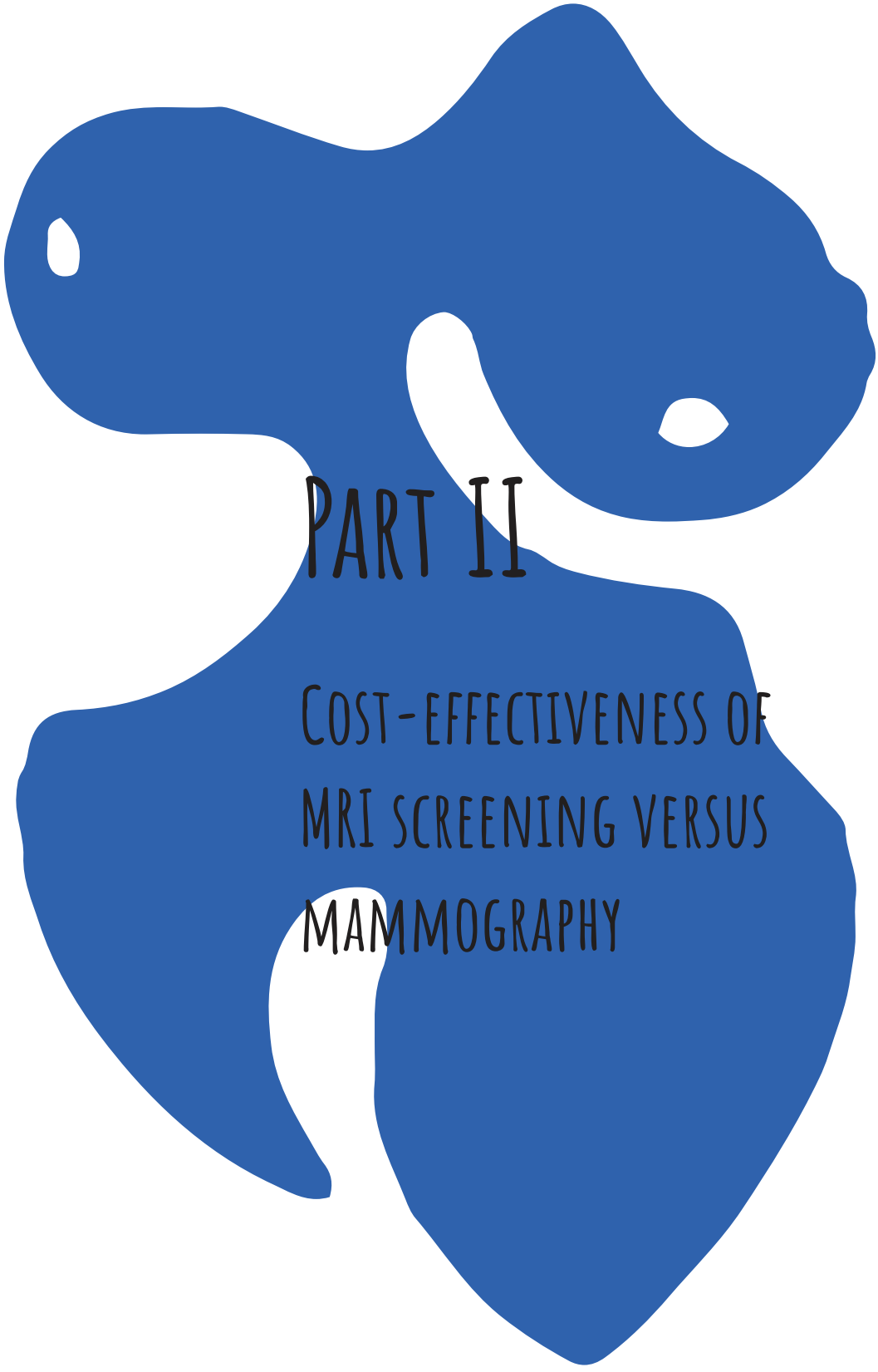
	Median VBD (IQR)			
Quartiles of VBD in the first round	**First round**	**Second round**	**Third round**	**Difference between the third and first round***
Q1	16.2 (15.8-16.5)	15.0 (13.3-16.6)	14.1 (12.0-16.2)	-2.0 (-4.0 - 0.1)
Q2	17.9 (17.4-17.9)	16.5 (14.7-18.3)	15.6 (13.4-17.9)	-2.3 (-4.5 - -0.1)
Q3	20.3 (19.6-21.1)	18.8 (16.8-20.8)	17.9 (15.3-20.3)	-2.5 (-5.0 - -0.2)
Q4	24.7 (23.2-27.0)	23.1 (20.5-25.9)	22.0 (18.8-25.2)	-3.3 (-6.2 - -0.6)

VBD= volumetric breast density
Q1=minimal value-25th percentile; Q2=25th percentile - median; Q3=median-75th percentile; Q4=75th percentile-maximum value
p-value trend analysis: <0.001

Table S4. Performance of mammography and MRI by VDG in both screening rounds, adjusted for age

	Rate (95%CI)/ 1000 screens		*P* value
Mammography[‖]	**VDG4**	**VDG3/2**	
Screen-detection	1.3 (0.5-3.1)	3.2 (1.3-7.6)	0.17
Recall[†]	16.2 (12.6-20.8)	13.6 (9.0-20.7)	0.59
False positives[‡]	14.4 (11.0-19.0)	10.3 (6.4-16.5)	0.28
PPV[§]	6.6 (2.6-15.6)	20.8 (8.6-42.4)	0.09
MRI[‖]			
Screen-detection	6.8 (4.4-10.5)	4.9 (2.2-11.0)	0.48
Recall[†]	28.6 (23.2-35.3)	26.1 (18.3-37.0)	0.82
False positives[‡]	20.9 (16.3-26.7)	20.1 (13.2-30.4)	0.92
PPV[§]	22.3 (14.5-32.1)	14.3 (5.5-32.4)	0.50

PPV= positive predictive value
[†] The recall rate was defined as the number of women with a positive screening result (BI-RADS ≥4 on mammography or BI-RADS ≥3 on MRI), divided by the total number of screens.
[‡] False-positive rates were calculated by dividing the number of women who were recalled and did not have breast cancer, by the number of women who underwent the respective screening test and did not have a screen-detected cancer.
[§] The PPV was defined as the number of screen-detected cancers by the number of positive screens.
[‖] Data not validated were excluded



Part II

Cost-effectiveness of MRI screening versus mammography

Chapter 5

Cost-effectiveness of breast cancer screening with magnetic resonance imaging for women at familial risk

Geuzinge HA, Obdeijn IM, Rutgers EJT, Saadatmand S, Mann RM, Oosterwijk JC, Tollenaar RAEM, de Roy van Zuidewijn DBW, Lobbes MBI, van 't Riet M, Hooning MJ, Ausems MGEM, Loo CE, Wesseling J, Luiten EJT, Zonderland HM, Verhoef C, Heijnsdijk EAM, Tilanus-Linthorst MMA, de Koning HJ, on behalf of the FaMRIsc study group

JAMA Oncology 2020;6(9):1381-1389

ABSTRACT

Importance: For women with a 20% or more familial risk of breast cancer without a known BRCA1/2 or TP53 variant, screening guidelines vary substantially, and cost-effectiveness analyses are scarce.

Objective: To assess the cost-effectiveness of magnetic resonance imaging (MRI) screening strategies for women with a 20% or more familial risk for breast cancer without a known BRCA1/2 or TP53 variant.

Design, setting and participants: In this economic evaluation, conducted from February 1, 2019, to May 25, 2020, microsimulation modeling was used to estimate costs and effectiveness on a lifetime horizon from age 25 years until death of MRI screening among a cohort of 10 million Dutch women with a 20% or more familial risk for breast cancer without a known BRCA1/2 or TP53 variant. A Dutch screening setting was modeled. Most data were obtained from the randomized Familial MRI Screening (FaMRIsc) trial, which included Dutch women aged 30 to 55 years. A health care payer perspective was applied.

Interventions: Several screening protocols with varying ages and intervals including those of the randomized FaMRIsc trial, consisting of the mammography (Mx) protocol (annual mammography and clinical breast examination) and the MRI protocol (annual MRI and clinical breast examination plus biennial mammography).

Main outcomes and measures: Costs, life-years, quality-adjusted life-years (QALYs), and incremental cost-effectiveness ratios (ICERs) were calculated and discounted by 3%. A threshold of €22,000 (US $24,796) per QALY was applied.

Results: This economic evaluation modeling study estimated that, on a lifetime horizon per 1000 women with the Mx protocol of the FaMRIsc trial, 346 breast cancers would be detected, and 49 women were estimated to die from breast cancer, resulting in 22,885 QALYs and total costs of €7,084,767 (US $7,985,135). The MRI protocol resulted in 79 additional QALYs and additional €2,657,266 (US $2 994 965). Magnetic resonance imaging performed only every 18 months between the ages of 35 and 60 years followed by the national screening program was considered optimal, with an ICER of €21,380 (US $24,097) compared with the previous nondominated strategy in the ranking, when applying the National Institute for Health and Care Excellence threshold. Annual screening alternating MRI and mammography between the ages of 35 and 60 years, followed by the national screening program, gave similar outcomes. Higher thresholds

would favor annual MRI screening. The ICER was most sensitive to the unit cost of MRI and the utility value for ductal carcinoma in situ and localized breast cancer.

Conclusions and relevance: This study suggests that MRI screening every 18 months between the ages of 35 and 60 years for women with a family history of breast cancer is cost-effective within the National Institute for Health and Care Excellence threshold for all densities. Higher thresholds would favor annual MRI screening. These outcomes support a change of current screening guidelines for this specific risk group and support MRI screening.

INTRODUCTION

Women with a family history of breast cancer have an increased risk of developing breast cancer, and an increased risk of developing it at a relatively young age.[1] In approximately 64% to 87% of these women, no causative hereditary gene variant has been found.[2] Because tumor stage at diagnosis is of importance for survival,[3] screening is advised, but guidelines differ substantially.

The American Cancer Society advises additional magnetic resonance imaging (MRI) for women with a lifetime risk of 20% or more of developing breast cancer,[4] whereas in the Netherlands and the UK, only mammography screening is advised for women at familial risk without a BRCA1/2 (BRCA1, OMIM 113705; and BRCA2, OMIM 114480) variant.[4,5] All guidelines recommend to start screening among women with a familial risk of breast cancer at a younger age than women at average risk.[4-6] However, younger women often have dense breast tissue,[7] which is associated with decreased mammographic sensitivity.[8] Magnetic resonance imaging screening has a high sensitivity, not affected by breast density.[9,10] However, MRI leads to more false-positive results and is associated with higher costs.[9-11] To our knowledge, little is known about the cost-effectiveness of MRI screening for women with a familial risk of breast cancer; one previous study evaluated the cost-effectiveness of MRI screening in women at familial risk of breast cancer without a known gene variant, showing by microsimulation modeling that MRI screening was very costly.[11] The model was based on data from a nonrandomized study.

Recently, the randomized Familial MRI Screening (FaMRIsc) trial showed higher breast cancer detection rates and detection of breast cancer at, on average, an earlier stage when screening with MRI in comparison with mammography in women at increased familial risk without a known BRCA or TP53 (OMIM 151623) gene variant.[9] In this study, we calculate real-life costs of MRI and mammography in the FaMRIsc trial. We estimate the cost-effectiveness by microsimulation modeling, and compare different screening scenarios by varying starting and stopping ages, screening intervals, and combinations of MRI and mammography.

METHODS

The FaMRIsc trial

In the multicenter randomized clinical FaMRIsc trial, Dutch women aged 30 to 55 years with a cumulative lifetime breast cancer risk of 20% or more due to a family history of breast cancer without a known BRCA1/2 or TP53 variant were randomly assigned into 2 screening groups after providing written informed consent.[12] The MRI group received annual MRI plus clinical breast examination (CBE), and mammography every 2 years. The

mammography (Mx) group received annual mammography and CBE, in accordance with the Dutch screening protocol.[5] Women refusing randomization could participate in a registration group (Reg-MRI group or Reg-Mx group) by providing consent for registration of their screening results. More details have been described elsewhere.[9,12] The FaMRIsc Study follows the Declaration of Helsinki[13] and was approved by the Erasmus University Medical Center Institutional Review Board (Rotterdam, the Netherlands; reference MEC-2010-292). The FaMRIsc trial is registered with the Netherlands Trial Register NL2661.

The microsimulation screening analysis model

In this economic evaluation, conducted from February 1, 2019, to May 25, 2020, we used the Microsimulation Screening Analysis (MISCAN) model, which simulates individual natural histories from birth to death and the natural history of breast cancer. We adjusted the version by Sankatsing et al[14] to extrapolate the findings of the FaMRIsc trial. To be able to model the difference in the numbers of detected ductal carcinoma in situ (DCIS) and T1a and T1b tumors between the 2 study groups,[9] 2 additional preclinical states were added to the original MISCAN model: DCIS_MRI and T1a/T1b_MRI (**Figure S1**). We assumed that DCIS and T1a and T1b tumors could for some time be detected only by MRI before they could also be detected by mammography or before they become clinically detectable. During all other preclinical states, the tumor could be detected with MRI as well as mammography or clinically diagnosed. Progression through the health states was modeled as a semi-Markov process. The model only takes into account first breast cancers and no contralateral breast cancers.

We assumed the mammographic sensitivity to be 15% lower than previously used in the model owing to the younger population we modeled.[15] We assumed that CBE would not lead to additional cancer detection.[16] Incidence, dwelling times, stage-specific sensitivities of MRI, and transition probabilities of the additional health state DCIS_MRI to DCIS and to T1a/T1b_MRI were estimated by calibration using the Nelder-Mead simplex optimization method.[17] We used data from all trial groups (Mx group + Reg-Mx group and MRI group + Reg-MRI group) to increase the amount of data for calibration. Model predictions were calibrated to the number of screening-detected breast cancers per T stage, the number of interval cancers, the number of detected cancers per 10-year age groups, and the number of screening-detected tumors during incident and prevalent rounds, all stratified by screening protocol as observed during the FaMRIsc trial. We aimed for all predicted numbers to fall within 95% Poisson CIs of the observed numbers of tumors.

Probabilities of (false) positive results and diagnostic procedures were obtained from the FaMRIsc trial, stratified by screening modality and by age (<50 and ≥50 years). Both true-positive and false-positive results were associated with diagnostic follow-up and

associated costs. For the screening period within the national breast cancer screening program, we applied the same probabilities as for the Mx protocol.

For the situation without screening, we assumed all women with a diagnosed breast cancer would undergo a diagnostic mammogram, CBE, biopsy, or fine needle aspiration, and all women with a diagnosed T2 or higher tumor would undergo an MRI. The percentage of ultrasonographic evaluations per formed in diagnosed cases in a situation without screening was assumed to be equal to the percentage of those performed among women with a diagnosed breast cancer within the Mx protocol.

Screening strategies

After calibration, we applied several screening strategies, varying in starting and stopping ages, intervals, and screening modalities. With stopping ages below the age of 75 years, we modeled the women to continue screening within the national screening program until the age of 75 years, consisting of biennial mammography at a local screening unit. Attendance rates were set at 100%.

Costs

We applied a health care perspective and considered only direct medical costs (converted to 2018 amounts; **Table S1**) and costs related to other causes of death. Costs of MRI, mammography in a hospital setting, and ultrasonography were derived from the tariff tool from an insurance company by calculating the mean of all published prices.[18] The price of mammography in a local screening unit was obtained from the Netherlands Comprehensive Cancer Organisation.[19] All other costs were obtained from a study by Saadatmand et al.[11] Costs of fine needle aspiration and biopsy were updated and adjusted by adding costs of pathologic examination of the specimen, obtained from the tariff tool.[18] Costs of breast-conserving surgery and mastectomy were adjusted assuming 1.5 consecutive hospital days with its price obtained from the Dutch costing manual.[20] Costs associated with breast cancer death were assumed to be €19,679 (US $22,180) and death due to other causes were assumed to be €15,044 (US $16,956).[21]

We multiplied costs with the resource use during the trial to calculate real-life costs. Mean treatment costs per TN stage were calculated by dividing total treatment costs per TN stage by the number of cancers. Model outcomes were multiplied with aforementioned prices to calculate costs per screening protocol.

Health state utilities

Utility values were obtained from the literature (**Table S2**). The utility value for the healthy state was based on a study by Versteegh et al.[22] Early-stage cancer was associated with disutility of 10%, regional cancer was associated with disutility of 25%, and metastasis was associated with disutility of 40%.[23] A disutility of 0.105 was applied for a positive

screening result with a duration of 5 weeks.[24] We did not apply a disutility for screening visits.[25]

Statistical analysis

We simulated the number of invitations, screening visits, screening-detected cancers, interval cancers, life-years, quality-adjusted life-years (QALYs), deaths from breast cancer, and deaths from other causes, all on a lifetime horizon from age 25 years for a cohort of 10 million Dutch women born in 1980. All results were scaled to 1000 women. Overdiagnosis was defined as detected cancers that would not have been diagnosed in a woman's lifetime in a situation without screening. Incremental cost-effectiveness ratios (ICERs) were calculated by dividing incremental costs by incremental QALYs. We plotted an efficiency frontier representing efficient strategies that are either less costly and more effective, or more costly but more cost-effective than those below the frontier. A cost- effectiveness threshold was based on the National Institute for Health and Care Excellence (NICE) threshold of £20,000 (€2,000 [US $24,796]). Average cost-effectiveness ratios were calculated by dividing additional costs by additional QALYs compared with a situation without screening. Costs and effects were discounted by 3%.[26]

One-way sensitivity analyses were performed for utility values, the price of MRI, and false-positive rates to analyze the association of these parameters with the ICER of the MRI protocol versus the Mx protocol. Utility values were varied ±10% of the base case values and the other parameters were varied ±20% of the base case values. Sensitivity analyses were discounted by 3%.

Five scenario analyses were performed to quantify methodological uncertainty. First, we applied discount rates of 4.0% for costs and 1.5% for effects, according to Dutch guidelines.[20] Second, we calculated the ICER without discounting. Third, we applied utility values based on a study by Lidgren et al[27] (**Table S3**). Fourth, we calculated the ICER without costs related to death from other causes. Fifth, we applied a disutility of 0.006 for 1 week for screening participation.[24] Scenario analyses were performed for the comparison of the 2 screening protocols of the FaMRIsc trial.

We calculated the risk of radiation-induced breast cancers for an optimal screening strategy with mammography compared with a strategy without mammography. We used the excess absolute risk model[28,29] with a glandular dose of a 2-view mammogram of 4.4 mGy.

RESULTS

Real-life results during the FaMRIsc trial

After a mean follow-up of 4.3 years, 41 tumors were detected in the MRI group, whereas 15 tumors were detected in the Mx group.[9] **Table 1** shows the number of detected tumors, woman-years at risk, and real-life screening costs according to group, age, and density during the FaMRIsc trial. The MRI protocol resulted in approximately 2 times higher costs of screening and additional investigation. Mean treatment costs are shown in **Table S4**.

Model calibration results

Figure S2 shows the number of observed breast cancers during the FaMRIsc trial according to T stage and the number of predicted cancers by our calibrated model. All predicted numbers were within the 95% CIs of the observed numbers.

Table 1. Real-life costs during the FaMRIsc trial by group, age and by density

			Costs, € (US $)	
	No. of breast cancers[a]	**Life-years at risk**	**Screening**	**Additional investigation**
MRI group by age				
<50 years	18	2106	740,188 (834,255)	171,054 (192,792)
≥50 years	23	1112	357,578 (403,021)	59,281 (66,815)
Total	41[b]	3218	1,097,766 (1,237,276)	230,335 (259,607)
Mx group by age				
<50 years	8	2099	341,568 (384,976)	87,576 (98,706)
≥50 years	7	1215	178,692 (201,401)	31,266 (35,239)
Total	15	3314	520,260 (586,377)	118,842 (133,945)
MRI group and registration MRI protocol by density				
BI-RADS A-C	38	2743	939,818 (1,059,255)	184,004 (207,388)
BI-RADS D	5	507	176,580 (199,021)	52,750 (59,454)
Total	43[b]	3249	1,116,397 (1,258,274)	236,754 (266 ,842)
Mx group and registration Mx protocol by density				
BI-RADS A-C	17	3648	567,145 (639,221)	120,408 (135,710)
BI-RADS D	6	659	105,386 (118,779)	42,227 (47,593)
Total	23	4308	672,531 (758,0007)	162,635 (183,304)

Abbreviations: BI-RADS, Breast Imaging Reporting and Data System Atlas; FaMRIsc, Familial MRI Screening; MRI, magnetic resonance imaging; Mx, mammography.

[a] Breast cancers include invasive breast cancers and ductal carcinoma in situ.

[b] One additional cancer was added in this article, which was excluded in the previous article.[9] This was an interval cancer between a mammogram and MRI in the first screening round in the MRI group.

Cost-effectiveness results

Table 2 and **Table S5** display the outcomes of all modeled strategies per 1000 women. With the Mx protocol of the FaMRIsc trial (strategy M), 346 breast cancers would be detected, and 49 women would die from breast cancer, resulting in 22 885 QALYs (discounted by 3%) and total costs of €7,084,767 (US $7,985,135) (discounted by 3%) and total costs of €23,497,356 (US $26,483,518) (undiscounted). With the MRI protocol of the FaMRIsc trial (strategy V), 377 breast cancers would be detected and 30 breast cancer deaths would occur, resulting in 79 additional QALYs (discounted by 3%) and additional costs of €2,657,266 (US $2,994,965) (discounted by 3%) and total costs of €28,024,674 (US $31,586,190) (undiscounted). Comparing these 2 protocols resulted in an ICER of €33,277 (US $37,506) per QALY gained (discounted).

Both screening protocols of the FaMRIsc trial were dominated by similar screening strategies without CBE (strategy B and U) (**Figure 1**). Strategies involving MRI resulted in fewer breast cancer deaths, lower numbers of interval cancers, and lower total treatment costs but more overdiagnosed cancers, compared with screening without MRI. The 2 strategies with intervals of 18 months were both on the efficiency frontier (**Figure 1**). Most strategies on the efficiency frontier consisted of screening from age 35 until 60 years, continued within the national screening program. Switching to screening within the national screening program before age 60 years resulted in higher numbers of clinically diagnosed cancers and breast cancer deaths, and were therefore dominated (**Table S5**).

Strategy D, consisting of MRI screening every 18 months between ages of 35 and 60 years followed by the national screening program had the highest acceptable ICER, €21,380 (US $24,002), when applying the NICE threshold of £20,000 (€22,000 [US $24,796]) and was considered optimal. Strategy E, consisting of alternating annual MRI or mammography between the ages of 35 and 60 years, was almost on the efficiency frontier. The effects of this strategy were similar to those of strategy D for somewhat higher cost. Strategies D and E, both followed by screening within the national breast cancer screening program, resulted in a reduction of 98 and 99 breast cancer deaths, respectively, and 65 or 66 overdiagnosed cases, respectively, when compared with a situation without screening.

Sensitivity and scenario analyses

Results of the deterministic sensitivity analyses are shown in **Figure 2**. The ICER was most sensitive to the price of MRI screening and the utility value for DCIS or localized breast cancer.

When applying Dutch discount rates, the ICER of the MRI protocol versus the Mx protocol became lower: €13,108 (US $14,774) per QALY gained. The difference in life-years and QALYs between the 2 protocols were 176 and 170, respectively, and the

Table 2. Modelled Effects and Costs per 1000 Women of Efficient Strategies With an ICER Below €100,000 (US $112,709), and the FaMRIsc Trial Strategies[a]

	No screening	Strategy A[b]	Strategy B[c]	Strategy C[d]	Strategy D[e]	Strategy M[f]	Strategy E[g]	Strategy F[h]	Strategy V[i]
Screening rounds	NA	22,296	27,136	18,895	18,706	27,136	26,196	25,939	25,924
Breast cancers	306	346	346	365	370	346	372	377	377
Screening detected	n/a	288	297	319	332	297	337	348	349
Clinically diagnosed	306	58	49	46	38	49	35	29	28
Breast cancer deaths	136	53	49	44	38	49	37	31	30
Reduction breast cancer deaths, compared to no screening, %	NA	-61%	-64%	-68%	-72%	-64%	-73%	-77%	-78%
False-positives	NA	1,331	1,578	1,514	1,885	2,436	2,086	2,629	3,825
Overdiagnosis, No. (% of screen-detected cancers)	NA	40 (14%)	40 (13%)	59 (18%)	65 (19%)	40 (13%)	66 (20%)	71 (20%)	71 (20%)
LYs	55,936	57,289	57,404	57,508	57,632	57,404	57,642	57,757	57,774
QALYs	47,450	48,774	48,885	48,992	49,113	48,876	49,122	49,236	49,242
Costs €, (US $)									
Screening tests	NA	1,937,798 (2,184,063)	2,382,952 (2,685,789)	2,980,816 (3,359,633)	4,274,868 (4,818,140)	3,994,553 (4,502,201)	4,251,282 (4,772,723)	6,141,510 (6,922,004)	8,677,921 (9,780,755)
Diagnosis	296,869 (334,597)	608,169 (685,458)	738,054 (831,850)	917,523 (1,034,126)	1,031,319 (1,162,384)	1,138,243 (1,282,897)	1,305,090 (1,470,947)	1,451,048 (1,635,454)	2,129,114 (2,399,693)
Treatment	4,329,646 (4,879,879)	3,367,954 (3,795 970)	3,295,594 (3,714,415)	2,798,040 (3,153,629)	2,512,550 (2,831,857)	3,295,594 (3,714,415)	2,514,195 (2,833,712)	2,268,554 (2,556,853)	2,242,559 (2,527,555)
Breast cancer death	2,684,157 (3,025,273)	1,046,478 (1,179,470)	972,526 (1,096,119)	866,085 (976,151)	744,323 (838,915)	972,526 (1,096,119)	723,739 (815,715)	615,290 (693,484)	598,822 (674,923)
Death other causes	12,816,630 (14,445,431)	14,040,806 (15,825,182)	14,096,439 (15,887,885)	14,176,051 (15,977,614)	14,267,489 (16,080,673)	14,096,439 (15,887,885)	14,282,608 (16,097,713)	14,364,011 (16,189,461)	14,376,257 (16,203,264)
Total	20,127,302 (22,685,180)	21,001,204 (23,670,142)	21,485,565 (24,216,058)	21,738,516 (24,501,155)	22,830,549 (25,731,969)	23,497,356 (26,483,518)	23,076,913 (26,009,643)	24,840,412 (27,997,256)	28,024,674 (31,586,190)
QALYs gained[j,k]	NA	283	315	338	365	311	366	393	390

	No screening	Strategy A[b]	Strategy B[c]	Strategy C[d]	Strategy D[e]	Strategy M[f]	Strategy E[g]	Strategy F[h]	Strategy V[i]
Total costs, € (US $)[k]	5,019,633 (5,657,553)	5,653,893 (6,372,418)	5,996,015 (6,758,019)	6,306,999 (7,108,524)	6,896,883 (7,773,373)	7,084,767 (7,985,135)	7,085,452 (7,985,907)	8,009,853 (9,027,785)	9,742,033 (10,980,099)
ACER, € (US $) [k]	NA	2,241 (2,526)	3,097 (3,491)	3,811 (4,295)	5,138 (5,791)	6,648 (7,493)	5,641 (6,358)	7,617 (8,585)	12,094 (13,631)
ICER, € (US $)[k]	**NA**	**2,241 (2,526)**	**10,588 (11,934)**	**13,812 (15,567)**	**21,380 (24,097)**	**Strongly dominated[l]**	**Weakly dominated[l]**	**40,919 (46,119)**	**Strongly dominated[l]**

Abbreviations: ACER, average cost-effectiveness ratio (comparison of a strategy with a situation without screening); FaMRIsc, Familial MRI Screening; ICER, incremental cost-effectiveness ratio (comparison of a strategy to the previous nondominated strategy in the ranking); LYs, life-years; MRI, magnetic resonance imaging; Mx, mammography; NA, not applicable; QALYs, quality-adjusted life-years.

[a] Breast cancers include invasive breast cancers and ductal carcinoma in situ. Results are without discounting. Outcomes contain the effects of both the described strategy and the subsequent national breast cancer screening program.

[b] Annual mammography between 40 and 60 years.

[c] Annual mammography between 35 and 60 years.

[d] Alternating MRI or mammography every 18 months between 35 and 60 years.

[e] Magnetic resonance imaging every 18 months between 35 and 60 years.

[f] Annual mammography and clinical breast examination between 35 and 60 years (Mx protocol in FaMRIsc trial).

[g] Alternating annual MRI or mammography between 35 and 60 years.

[h] Annual MRI between 35 and 60 years.

[i] Annual MRI plus clinical breast examination, and biennial mammography between 35 and 60 years (MRI protocol in FaMRIsc trial).

[j] Relative to a situation without screening.

[k] Discounted by 3%.

[l] Strategies which are more costly and less effective than another strategy are dominated

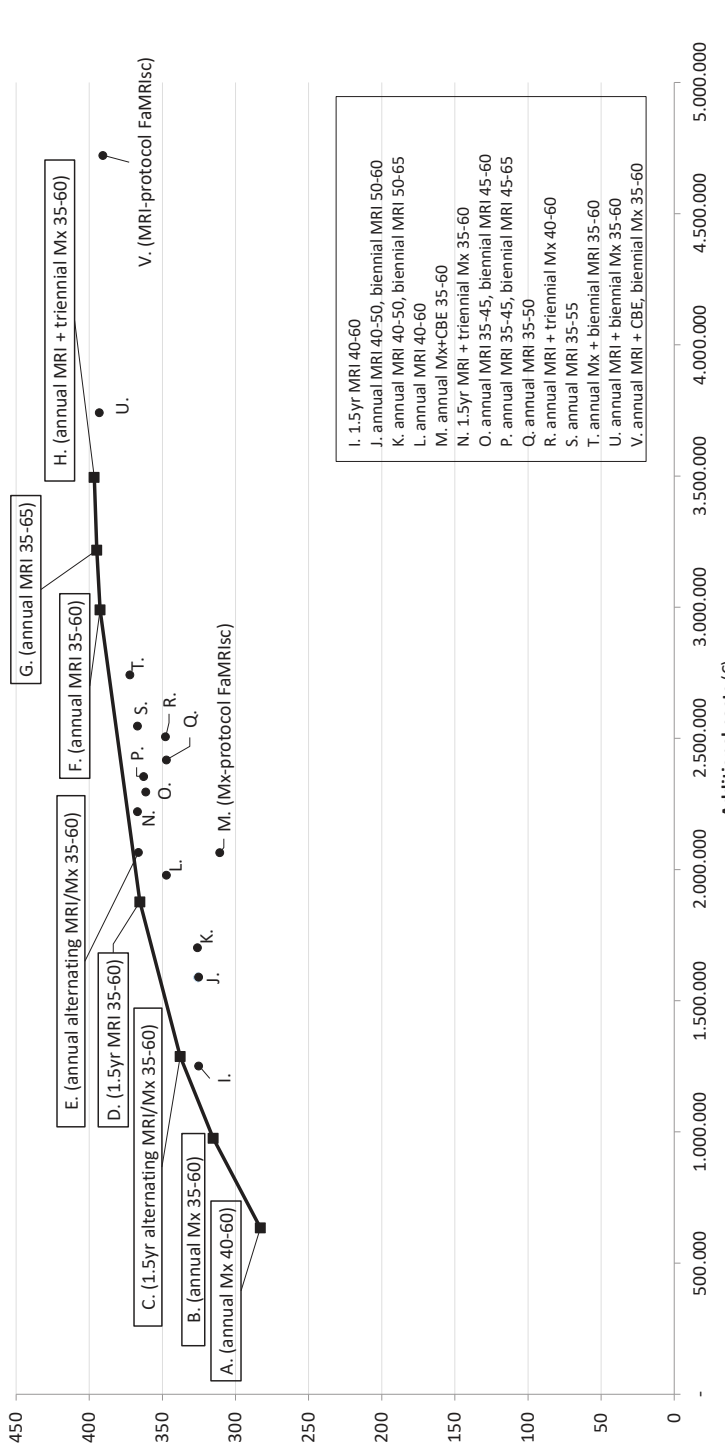


Figure 1. Efficiency Frontier of Screening Strategies
Number indicates interval, and ranges represent women's ages reported in years; all results are discounted by 3%. CBE: clinical breast examination, FaMRIsc: Familial MRI Screening Study, MRI: magnetic resonance imaging, Mx: mammography, QALY: quality-adjusted life-year. Additional costs €1 = $1.13 on July 1, 2020.

difference in costs was €2,234,665 (US $2,518,657). Without discounting, the ICER was €12,376 (US $13,949).

In the third scenario analysis, in which we applied a different set of utility values, the difference in QALYs between the MRI protocol and Mx protocol became 71, which was lower compared with the base case. Consequently, the ICER became larger: €37,489 (US $42,253) per QALY gained (discounted). When not applying costs related to death from other causes, the ICER became €32,712 (US $36,869) per QALY gained (discounted), which was similar to the ICER when including these costs. Applying a utility decrement for screening participation hardly affected the ICER, which became €33,534 (US $37,796) (discounted).

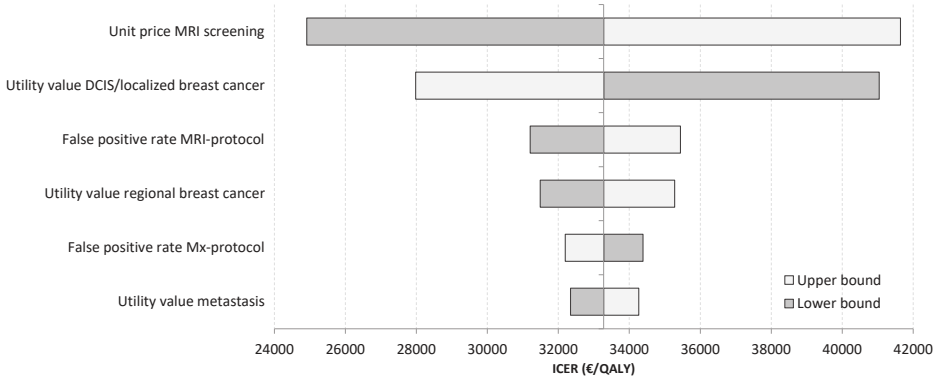


Figure 2. Tornado Diagram of the 1-Way Sensitivity Analyses
All results are discounted by 3%. DCIS: ductal carcinoma in situ, ICER: incremental cost-effectiveness ratio, MRI: magnetic resonance imaging, Mx: mammography, QALY: quality-adjusted life-year.

Radiation risk

In a situation with additional mammography to the optimal screening strategy (D) consisting of MRI every 18 months between the ages of 35 and 60 years, radiation would induce 0.94 breast cancers and 0.12 breast cancer deaths per 1,000 women. In this situation, 3 additional breast cancers would be detected by screening of which 1 would be overdiagnosed, and 2 breast cancer deaths would be prevented (undiscounted) compared with a strategy without additional mammography (strategy D).

DISCUSSION

This economic modeling study of data on Dutch women showed that the detection of more tumors at an early stage and fewer at a late stage by MRI9 could be a cost-effective method to reduce breast cancer mortality despite more overdiagnosis and higher costs

in comparison with mammography. Yearly MRI seems to bring the largest mortality reduction, but for an ICER higher than allowed by NICE guidelines.[30] Neither protocol of the FaMRIsc trial was on the efficiency frontier, mainly owing to the addition of CBE that proved to be inefficient.[9,16] Screening with MRI only every 18 months between the ages of 35 and 60 years and subsequent screening in the national screening program until age 75 years would be an efficient and cost-effective strategy, with an ICER just below the threshold of £20,000 (€22,000 [US $24,796]). We also found that the additional association of mammography with this strategy was limited. Screening consisting of alternating annual MRI and mammography between ages of 35 and 60 years, followed by screening within the national screening program until the age of 75 years was almost on the frontier, with similar effects and more costs as the previously mentioned strategy (MRI only every 18 months between the ages of 35 and 60 years). Most of the efficient strategies consisted of screening from 35 to 60 years of age, with continuation of screening within the national screening program. Furthermore, our results indicated that the switch to the national screening program should not take place before 60 years of age.

We modeled a Dutch health care setting but we expect the relative difference in health outcomes between our modeled strategies to be similar in other countries. In contrast, unit prices as well as cost-effectiveness thresholds vary substantially per country, which should be taken into account when generalizing our results to other countries.

We simulated one group of women with, on average, the same risk of breast cancer. However, starting screening at 35 years of age may not be beneficial for all women within this group, depending on the youngest age of breast cancer diagnosis of a family member and their individually calculated life-time risk.[31] Therefore, family history should be taken into account when choosing the starting age for screening.

To our knowledge, one previous study evaluated the cost-effectiveness of additional MRI screening for this group of women. Saadatmand et al[11] calibrated the MISCAN model on data from the 1999-2006 MRI Screening (MRISC) study. The breast cancer incidence in the FaMRIsc Study was higher than that in the MRISC study, and the sensitivity of both MRI and mammography were also higher.

Strengths and limitations

This study has some strengths, including the use of randomized clinical trial data for calibration, which has, to our knowledge, not been done before for this group of women. By using randomized clinical trial data, the model gets more information on the performance of MRI and mammography separately than when these screening modalities are performed simultaneously.

This study also has some limitations. First, the study sample of the FaMRIsc trial was still quite small for calibration. The numbers of observed cancers stratified by group and stage were small and therefore 95% CIs were large. Therefore, we added the data of the registration groups. However, there may have been a difference in population between women registered and those randomized. A second limitation is the assumption that there is no DCIS that is detectable only by mammography. Third, we were unable to model strategies by breast density categories as the numbers by breast density in the FaMRIsc trial were too small, albeit the associations of MRI with detection seem similar across density categories. A recent study showed the benefit of MRI screening in women with extremely dense breasts.[32] Fourth, we did not measure utility values within our study population. Utility values related to breast cancer vary significantly in the literature[33,34] and we are aware of the association of these values with the results, as shown in our analyses. Furthermore, we would like to point out the uncertainty of efficiency frontiers as such. Efficiency frontiers are sensitive to changes in underlying data and assumptions, and they do not display uncertainty.[35]

Downsides of MRI are its high costs, more false-positive results, and increased overdiagnosis. Overdiagnosis may be a result of excessive detection of low-grade tumors, but our model cannot distinguish between low-grade or high-grade tumors. Overdiagnosis is captured in our results and the same (dis)utility values were applied to all modeled breast cancer cases because one does not know whether a cancer is overdiagnosed or not.

Applying MRI screening may have some practical implications. Hospitals need to have enough capacity for the screening and for additional diagnostic testing due to more (false) positive results, to prevent waiting lists. In addition, radiologists may need additional training to guarantee good quality, as MRI screening requires expertise.

Currently, abbreviated MRI seems promising, which has shorter acquisition time and reading time while maintaining diagnostic accuracy.[36] A less time-consuming MRI will decrease the price of the test, which has a favorable association with the ICER, as shown in our sensitivity analyses.

Conclusions

Based on this cost-effectiveness analysis, MRI screening every 18 months or alternating annual MRI and mammography between the ages of 35 and 60 years may be recommended for women at increased familial risk of breast cancer, both followed by screening within the national screening program, when applying the NICE threshold. Annual MRI was associated with the largest mortality reduction, but for an ICER higher than allowed by NICE guidelines.

REFERENCES

1. Claus EB, Risch NJ, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol*. 1990;131(6):961-972.
2. Hartmann LC, Lindor NM. The role of risk-reducing surgery in hereditary breast and ovarian cancer. *N Engl J Med*. 2016;374(5):454-468.
3. Saadatmand S, Bretveld R, Siesling S, Tilanus-Linthorst MMA. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173,797 patients. *BMJ*. 2015;351:h4901.
4. Saslow D, Boetes C, Burke W, et al; American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin*. 2007;57(2):75-89.
5. Oncoline. Richtlijn mammacarcinoom (breast cancer national guideline). 2018; https://www.oncoline.nl/borstkanker. Accessed January 15, 2019.
6. American College of Radiology. Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas), 4th edition. Reston, VA; 2003.
7. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. AJR Am J Roentgenol. 2012;198(3):W292-W295.
8. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*. 2007;356(3):227-236.
9. Saadatmand S, Geuzinge HA, Rutgers EJT, et al; FaMRIsc study group. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol*. 2019;20(8):1136-1147.
10. Berg WA, Zhang Z, Lehrer D, et al; ACRIN 6666 Investigators. 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): 1394-1404.
11. Saadatmand S, Tilanus-Linthorst MMA, Rutgers EJT, et al. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst*. 2013;105(17):1314-1321.
12. Saadatmand S, Rutgers EJ, Tollenaar RA, et al. Breast density as indicator for the use of mammography or MRI to screen women with familial risk for breast cancer (FaMRIsc): a multicentre randomized controlled trial. *BMC Cancer*. 2012;12:440.
13. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191-2194.
14. Sankatsing VDV, Heijnsdijk EAM, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in the Netherlands. *Int J Cancer*. 2015;137(8):1990-1999.
15. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology*. 2002;225(1):165-175.
16. Roeke T, van Bommel AC, Gaillard-Hemmink MP, Hartgrink HH, Mesker WE, Tollenaar RA. The additional cancer yield of clinical breast examination in screening of women at hereditary increased risk of breast cancer: a systematic review. *Breast Cancer Res Treat*. 2014;147(1):15-23.
17. Barton RR, Ivey JS. Nelder-Mead simplex modifications for simulation optimization. *Manag Sci*. 1996;42(7):954-973.

18. CZ Healthcare Insurance. Tarieventool. 2018; https://www.cz.nl/ service-en-contact/zoek-tarieven. Accessed October 3, 2018.
19. Netherlands Comprehensive Cancer Organisation (IKNL). Monitor bevolkingsonderzoek borstkanker 2016. 2018; https://www.rivm.nl/sites/default/files/2018-11/Monitor%20 bevolkingsonderzoek%20borstkanker%202016%20webversie.pdf Accessed July 9, 2019.
20. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Tan S. Kostenhandleiding: methodologie van kostenonderzoek. In: Referentieprijzen Voor Economische Evaluaties in de Gezondheidszorg. Diemen; 2015.
21. Polder JJ, Barendregt JJ, van Oers H. Health care costs in the last year of life—the Dutch experience. *Soc Sci Med*. 2006;63(7):1720-1731.
22. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19(4):343-352.
23. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst*. 2006;98(11): 774-782.
24. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer*. 1991;49(4): 538-544.
25. Rijnsburger AJ, Essink-Bot ML, van Dooren S, et al. Impact of screening for breast cancer in high-risk women on health-related quality of life. *Br J Cancer*. 2004;91(1):69-76.
26. Gold M, Siegel J, Russel L, Weinstein M. Cost-Effectiveness in Health and Medicine. Oxford University Press; 1996.
27. Lidgren M, Wilking N, Jönsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Qual Life Res*. 2007;16(6):1073-1081.
28. Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice JD Jr. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. *Radiat Res*. 2002;158(2):220-235.
29. de Gelder R, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. *Br J Cancer*. 2011;104(7):1214-1220.
30. National Institute for Health and Care Excellence. Guide to the processes of technology appraisal. 2019; https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-apr-2018.pdf. Accessed June 25, 2020.
31. Tilanus-Linthorst MM, Lingsma HF, Evans DG, et al. Optimal age to start preventive measures in women with BRCA1/2 mutations or high familial breast cancer risk. *Int J Cancer*. 2013;133(1):156-163.
32. Bakker MF, de Lange SV, Pijnappel RM, et al; DENSE Trial Study Group. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med*. 2019;381(22):2091-2102.
33. Bromley HL, Petrie D, Mann GB, Nickson C, Rea D, Roberts TE. Valuing the health states associated with breast cancer screening programmes: a systematic review of economic measures. *Soc Sci Med*. 2019;228:142-154.
34. Peasgood T, Ward SE, Brazier J. *Health-state utility values in breast cancer*. Expert Rev Pharmacoecon Outcomes Res. 2010;10(5):553-566.
35. Stollenwerk B, Lhachimi SK, Briggs A, et al. Communicating the parameter uncertainty in the IQWiG efficiency frontier to decision-makers. *Health Econ*. 2015;24(4):481-490.
36. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers RD, Bieling HB. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection—a novel approach to breast cancer screening with MRI. *J Clin Oncol*. 2014;32(22):2304-2310.

SUPPLEMENTARY APPENDIX

Writing committee:

H Amarens Geuzinge (Erasmus University Medical Centre), Inge-Marie Obdeijn (Erasmus University Medical Centre), Eveline AM Heijnsdijk (Erasmus University Medical Centre), Emiel JT Rutgers (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital); Sepideh Saadatmand (Erasmus University Medical Centre), Ritse M Mann (Radboud University Hospital; The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), Jan C Oosterwijk (Medical Centre Leeuwarden; University Medical Centre Groningen), Rob AEM Tollenaar (Leiden University Medical Centre) Diderick BW de Roy van Zuidewijn (Medical Centre Leeuwarden), Marc BI Lobbes (Maastricht University Medical Center), Martijne van 't Riet (Reinier de Graaf Gasthuis), Maartje J Hooning (Erasmus University Medical Centre), Ingeborg Mares-Engelberts (Vlietland ziekenhuis), Margreet GEM Ausems (University Medical Centre Utrecht), Claudette E Loo (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), J Wesseling (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), Ernest JT Luiten(University Medical Centre Utrecht), Harmien M Zonderland (Amsterdam UMC, University of Amsterdam), Cees Verhoef (Erasmus University Medical Centre), Madeleine MA Tilanus-Linthorst (Erasmus University Medical Centre), Harry J de Koning (Erasmus University Medical Centre)

Other co-authors of the FaMRIsc study group:

CHM van Deurzen (Erasmus University Medical Centre), E Madsen (Erasmus University Medical Centre), J Rothbarth (Erasmus University Medical Centre), LB Koppert (Erasmus University Medical Centre Medical Centre), C de Monye (Erasmus University Medical Centre), MM van Rosmalen (Erasmus University Medical Centre), J Remmelzwaal (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital), M Schlooz-Vries (Radboud University Hospital), R Roi-Antonides (Medical Centre Leeuwarden), S van der Meij (Amsterdam University Medical Centre), WE Mesker (Leiden University Medical Centre), MNJM Wasser (Leiden University Medical Centre), K Keymeulen (Academic Hospital, Maastricht), WB Veldhuis (University Medical Centre Utrecht), AJ Witkamp (University Medical Centre Utrecht), E van Druten (Reinier de Graaf Gasthuis), E Tetteroo (Amphia Ziekenhuis), C Contant (Maasstad ziekenhuis), Nico Karssemeijer PhD[4]

List of participating hospitals

Erasmus University Medical Center, Antoni van Leeuwenhoek – the Netherlands Cancer Institute, Maastricht University Medical Center, Leiden University Medical Center, Radboud university medical center, Academic Medical Center and University Medical Center Utrecht, Medical Centre Leeuwarden, Reinier de Graaf Hospital Delft, Amphia Hospital Breda, and Vlietland Hospital Vlaardingen/Rotterdam, Maasstad Hospital Rotterdam

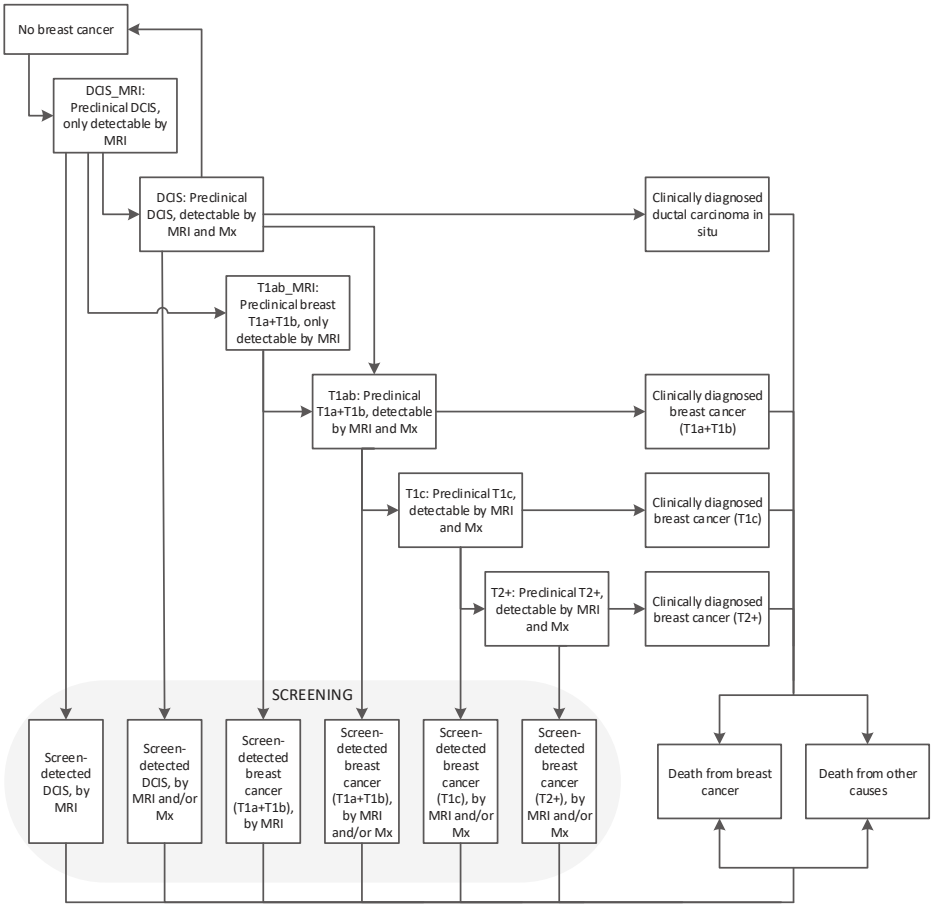


Figure S1. Model structure

Table S1. Unit prices per procedure associated with breast cancer screening, diagnosis and treatment

Procedure	Unit price (€)
Screening and diagnosis	
Mammography	91.97 [a]/69.10[b]
MRI	272.00
Consultation (with CBE)	72.57
Ultrasound	115.23
Fine needle aspiration	288.62
Biopsy	288.62
Surgery	
Breast conserving surgery	1452.49
Breast conserving surgery incl. sentinel	
node biopsy	1512.14
Mastectomy	1623.35
Mastectomy incl. sentinel node biopsy	1682.99
Lymph node dissection	884.98
Adjuvant therapy	
Radiotherapy	6885.05
Chemotherapy	3573.21
Chemotherapy followed by one year of Trastuzumab	25832.18
Hormonal therapy	2574.81

Prices derived from the study by Saadatmand et al [1] and the costing manual [2] were converted to Euro 2018 prices using consumer price indices [3].

[a] Mammography at a hospital (screening and diagnostic)

[b] Mammography within the national breast cancer screening programme

Table S2. Utility values and durations

Health state	Utility value	Duration
No breast cancer	0.858	n.a.
After a (false) positive screening result	0.105 (disutility)	5 weeks
DCIS/localized breast cancer	0.772	5 years
Regional breast cancer	0.644	5 years
Metastasis	0.515	Until death
Death	0	

Table S3. Alternative utility values and durations used in a scenario analysis

Health state	Utility value	Duration
No breast cancer	0.858	n.a.
After a (false) positive screening result	0.105 (disutility)	5 weeks
DCIS/localized breast cancer, first year	0.696	1 year
DCIS/localized breast cancer, after the first year	0.779	10 years
Regional breast cancer	0.685	11 years
Metastasis	0.515	Until death
Death	0	

Table S4. Mean costs per tumor stage

	No. of tumors	Mean additional investigation costs	Mean surgery costs	Mean radiotherapy costs	Mean systemic therapy costs	Mean total costs
DCIS	25	433	1,554	3,305	0	5,292
T1a, N-	7	624	1,601	984	0	3,209
T1b, N-	8	627	1,604	2,582	0	4,813
T1c, N-	12	560	1,540	4,590	7,511	14,200
T2+, N-	1	669	1,512	6,885	6,148	15,214
T1a, N+	0	n.a.	n.a.	n.a.	n.a.	n.a.
T1b, N+	2	433	1,512	6,885	3,074	11,904
T1c, N+	5	571	1,854	5,508	4,403	12,336
T2+, N+	6	890	2,372	6,885	6,148	16,296

In this paper, one additional cancer was added, which was excluded in the previous paper [4]. This was an interval cancer between a mammogram and MRI in the first screening round in the MRI-arm.

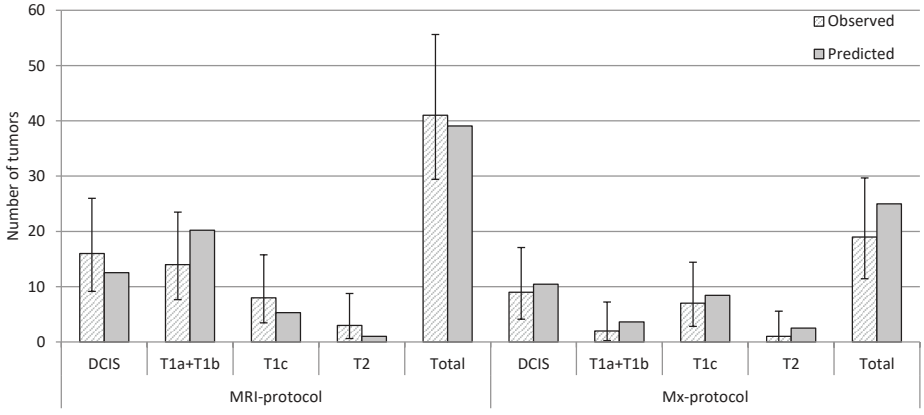


Figure S2. Observed and predicted screen-detected cancers according to T-stage

Table S5. Modelled effects and costs per 1000 women of dominated screening protocols and of protocols with an ICER >100,000 euro per QALY gained

	I	J	K	L	N	O	P	Q	R
Screening rounds	16,256	17,595	17,520	21,152	18,706	20,011	19,901	22,784	21,137
Breast cancers	372	371	380	376	370	368	381	349	377
Screen detected	325	323	334	337	332	327	343	301	338
Clinically diagnosed	47	48	46	40	38	41	38	48	38
Breast cancer deaths	42	43	42	37	38	40	38	46	36
Reduction breast cancer deaths, compared to no screening	-69%	-68%	-69%	-73%	-72%	-71%	-72%	-66%	-74%
False-positives	1623	1,762	1,781	2,131	1759	2,020	2,010	2,240	2,715
Overdiagnosis (% of screen-detected cancers)	66 (20%)	65 (20%)	74 (22%)	70 (21%)	64 (19%)	62 (19%)	75 (22%)	43 (14%)	71 (21%)
LYs	57,492	57,490	57,500	57,594	57,635	57,609	57,627	57,527	57,607
QALYs	48,977	48,972	48,979	49,079	49,117	49,086	49,099	49,002	49,088
Costs	3,498,785	3,855,876	4,225,106	4,839,223	4,725,348	4,625,117	4,866,252	4,388,062	5,600,707
Screening tests (€)									
Diagnosis (€)	866,028	966,838	979,856	1,148,518	1,111,498	1,132,147	1,138,247	1,251,846	1,472,032
Treatment (€)	2,604,847	2,639,056	2,574,523	2,408,956	2,509,938	2,589,918	2,489,337	2,980,128	2,381,939
Breast cancer death (€)	826,524	843,959	823,942	719,972	742,591	785,312	752,129	904,354	705,282
Death other causes (€)	14,205,372	14,192,464	14,207,268	14,285,271	14,268,542	14,236,904	14,261,547	14,147,949	14,296,118
Total costs (€)	22,001,557	22,498,193	22,810,695	23,401,940	23,357,918	23,369,399	23,507,512	23,672,340	24,456,077
QALYs gained[#]*	325	325	326	347	367	361	363	347	348
Total costs (€)*	6,271,826	6,610,610	6,722,959	6,998,746	7,241,465	7,316,046	7,375,201	7,438,477	7,526,288
ACER (€/QALY)*	3,853	4,894	5,228	5,701	6,057	6,359	6,496	6,969	7,209
ICER (€/QALY)*	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated

	S	T	G	H	U
Screening rounds	22,690	26,202	27,061	25,928	25,924
Breast cancers	358	370	385	377	377
Screen detected	315	335	360	349	349
Clinically diagnosed	42	34	25	28	28
Breast cancer deaths	40	36	29	31	30
Reduction breast cancer deaths, compared to no screening	-71%	-74%	-79%	-77%	-78%
False-positives	2,266	2,466	2,771	2,444	3,380
Overdiagnosis (% of screen-detected cancers)	52 (17%)	64 (19%)	79 (22%)	71 (20%)	71 (20%)
LYs	57,622	57,668	57,777	57,769	57,774
QALYs	49,100	49,145	49,252	49,250	49,246
Costs (€)					
Screening tests (€)	4,884,424	5,243,477	6,819,899	6,796,109	7,123,245
Diagnosis (€)	1,269,612	1,574,109	1,524,262	1,593,060	1,876,975
Treatment (€)	2,678,728	2,493,810	2,176,410	2,249,955	2,242,559
Breast cancer death (€)	792,564	707,330	580,499	604,132	598,822
Death other causes (€)	14,231,504	14,294,899	14,389,857	14,372,210	14,376,257
Total costs (€)	23,856,832	24,313,625	25,490,928	25,615,466	26,217,858
QALYs gained#*	367	372	395	397	393
Total costs (€)*	7,567,680	7,762,521	8,237,080	8,514,625	8,761,895
ACER (€/QALY)*	6,945	7,372	8,149	8,813	9,525
ICER (€/QALY)*	Dominated	Dominated	101,489	161,008	Dominated

Breast cancers include invasive breast cancers and DCIS

Outcomes contain the effects of both the described strategy and the subsequent National breast cancer screening programme

LYs: life-years; QALYs: quality-adjusted life-years; ACER: average cost-effectiveness ratio (comparison of a strategy to a situation without screening); ICER: incremental cost-effectiveness ratio (comparison of a strategy to the previous non-dominated strategy in the ranking)

Relative to a situation without screening

*Discounted by 3.0%

I. MRI every 18 months between age 40 and 60

J. annual MRI between age 40 and 50, biennial MRI between age 50 and 60

K. annual MRI between age 40 and 50, biennial MRI between age 50 and 65

L. annual MRI between age 40 and 60

N. MRI every 18 months and triennial mammography between age 35 and 60

O. annual MRI between age 35 and 45, biennial MRI between age 45-60

P. annual MRI between age 35 and 45, biennial MRI between age 45-65

Q. annual MRI between age 35 and 50

R. annual MRI and triennial mammography between age 40 and 60

S. annual MRI between age 35 and 55

T. annual mammography and biennial MRI between age 35 and 60

G. annual MRI between age 35 and 60

H. annual MRI, and triennial mammography between age 35 and 60

U. annual MRI and biennial mammography between age 35 and 60

References

1. Saadatmand S, Tilanus-Linthorst MMA, Rutgers EJT, et al. Cost-Effectiveness of Screening Women With Familial Risk for Breast Cancer With Magnetic Resonance Imaging. *J Natl Cancer Inst.* 2013;105(17):1314-1321.
2. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Tan S. Kostenhandleiding: Methodologie van Kostenonderzoek En Referentieprijzen Voor Economische Evaluaties in de Gezondheidszorg. Diemen; 2015.
3. Statistics Netherlands. Annual rate of change CPI; since 1963. 2018; http://statline.cbs.nl/Statweb/publication/?DM=SLEN&PA=70936ENG&D1=0&D2=623,636,649,662,675,688,701,714&LA=EN&VW=T.
4. Saadatmand S, Geuzinge HA, Rutgers EJT, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol.* 2019;20(8):1136-1147.

Chapter 6

Cost-effectiveness of magnetic resonance imaging screening for women with extremely dense breast tissue

Geuzinge HA, Bakker MF, Heijnsdijk EAM, van Ravesteyn NT, Veldhuis WB, Pijnappel RM, de Lange SV, Emaus MJ, Mann RM, Monninkhof EM, de Koekkoek-Doll PK, van Gils CH, de Koning HJ, on behalf of the DENSE trial study group

Journal of the National Cancer Institute. 2021;113(11):1476-1483

ABSTRACT

Background: Extremely dense breast tissue is associated with increased breast cancer risk and limited sensitivity of mammography. The DENSE trial showed that additional MRI screening in women with extremely dense breasts resulted in significantly fewer interval cancers. The cost-effectiveness of MRI screening for these women is unknown.

Methods: We used the MISCAN-breast microsimulation model to simulate several screening protocols containing mammography and/or MRI to estimate long-term effects and costs. The model was calibrated using results of the DENSE trial, and adjusted to incorporate decreases in breast density with increasing age. Screening strategies varied in the number of MRIs and mammograms offered to ages 50-75. Outcomes were numbers of breast cancers, life-years, quality-adjusted life-years (QALYs), breast cancer deaths and overdiagnosis. Incremental cost-effectiveness ratios (ICERs) were calculated (3% discounting), with a willingness-to-pay threshold of €22,000.

Results: Calibration resulted in a conservative fit of the model regarding MRI detection. Both strategies of the DENSE trial were dominated (biennial mammography; biennial mammography plus MRI). MRI alone every four years was cost-effective with €15,620/QALY. Screening every three years with MRI alone resulted in an ICER of €37,181/QALY. All strategies with mammography and/or a two-year interval were dominated because other strategies resulted in more additional QALYs per additional euro. Alternating mammography and MRI every two years was close to the efficiency frontier.

Conclusions: MRI screening is cost-effective for women with extremely dense breasts, when applied at a four-year interval. For a higher willingness to pay than €22,000/QALY gained, MRI at a three-year interval is cost-effective too.

INTRODUCTION

Approximately 8% of Dutch women aged 50-74 years have extremely dense breast tissue.[1] Women with extremely dense breast tissue have approximately a twofold higher risk of developing breast cancer than the average screening population.[2] At the same time, dense breast tissue limits the sensitivity of mammography, resulting in high numbers of interval cancers.[3,4] In contrast to mammography, the effect of breast density on the sensitivity of MRI is limited.[5] However, in most Western countries, women with average breast cancer risk, including women with dense breast tissue, are currently screened with mammography only.

Recently, the multicenter randomized Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial showed that additional MRI screening for women with extremely dense breast tissue resulted in significantly fewer interval cancers.[6] Furthermore, screen-detected tumors were on average smaller among MRI participants than those among women receiving mammography alone. However, MRI screening also resulted in more false positive results,[6] which will lead to additional costs.

Several modelling studies have shown that MRI screening can be cost-effective among high-risk women, especially women with a BRCA1 mutation.[7-10] It is unknown whether MRI can be cost-effective for women with extremely dense breasts who are currently screened within the Dutch national mammography screening program. As MRI screening is more expensive than mammography, which can lead to an increase in health care spending, a cost-effectiveness analysis is needed to evaluate whether the additional effects are worth the money.

In this study, we estimate the cost-effectiveness of MRI screening compared to mammography in women with extremely dense breast tissue by using the results of the DENSE trial and microsimulation modelling. We quantify the effects and costs of several different screening scenarios by varying the screening interval between MRIs and mammograms offered aged 50-75.

METHODS

DENSE trial

The DENSE trial is embedded within the Dutch biennial mammography screening program, for women aged 50-75 years. Women with extremely dense breasts (Volpara density grade 4) and a negative ('normal') mammography result (Breast Imaging-Reporting and Data System [BI-RADS] category 1 or 2) were randomly assigned to two groups: the MRI invitation group (n=8,061) and the control group (n=32,312).[6] Women assigned to the MRI invitation group were offered additional MRI screening (women

who accepted this offer are referred to as the MRI participants, n=4,783).[6] Women in the control group did not receive additional screening. Breast density was measured using Volpara imaging software. Volpara density grades (VDG 1 to 4) correspond to the categories of the fourth BI-RADS edition.[11] All MRI examinations were performed on 3.0 Tesla MRI systems and the macrocyclic gadolinium based contrast agent gadobutrol (Gadovist; Bayer AG, Leverkusen, Germany) was used in all examinations. More details of the DENSE trial have been described previously.[6,12] The study has been approved by the Dutch Minister of Health.

MISCAN model

To extrapolate the findings of the DENSE trial, we used an updated version of the Microsimulation Screening Analysis (MISCAN) model by Sankatsing et al.[13] The MISCAN model simulates individual life histories from birth to death and the natural history of breast cancer. A subset of women have an onset of breast cancer; the probability of onset increases with age. At each breast cancer stage (ductal carcinoma in situ (DCIS), T1A, T1B, T1C, T2+)[14] the tumor can be preclinical in which it may grow to the subsequent stage, or become clinically diagnosed, or screen-detected. The model structure has been published previously.[13]

The MISCAN model was used to model women with extremely dense breast tissue. Incidence, dwelling times (the time between transitions from one stage to the next), and stage-specific sensitivities of MRI and mammography were estimated by calibration. Model predictions were calibrated to the numbers of screen-detected cancers during the first (prevalent) and second (incident) round, and interval cancers during the first round as observed during the DENSE trial among the MRI participants and the control group (further specified in the Appendix).[6] We aimed to model the predictions with 95% Poisson confidence intervals of the observed numbers.

After calibration, the model was adjusted to incorporate decreases in breast density over time. Based on Dutch data[15], 21.9% of the women with VDG4 at the age of 50 was modelled to remain at that level. For 78.1% of the women, a decline in breast density was modelled, at either the age of 55 and/or 65 (Table 1). Decreasing breast density over time was assumed to be associated with increasing sensitivity of mammography, decreasing breast cancer incidence and decreasing numbers of false positive mammography results (Table 1).[1,16] All other parameters were assumed to be equal across density categories.

Probabilities of additional investigations and false positive results in the MRI participants were obtained from the DENSE trial (Table S1 and S2). These probabilities were not measured in the control arm. The probability of a false positive mammogram in extremely dense breasts was based on Wanders et al.[1] Based on published data,[17-19] expert opinion and according to Dutch practice, we assumed that all women with a positive mammogram would be referred to a hospital to all undergo tomosynthesis, and

88% undergo an ultrasound, 38% an ultrasound-guided biopsy and 8% a stereotactic biopsy (for estimations see Table S3).

In the DENSE trial, 22% of the MRI participants with a positive MRI underwent an ultrasound of the axilla. Imaging the axilla is not part of the screening work-up, but is performed for efficiency reasons in women receiving a biopsy: if the biopsy is positive the woman does not have to be recalled for a staging ultrasound of the axilla. We applied an equal probability after a positive mammogram. Based on Dutch guidelines, we assumed all women to undergo a PET-CT after a T2+N+ diagnosis.[20] Furthermore, we modelled 26%, 27% and 38% of women with a DCIS, T1A-T1C and T2+ diagnosis respectively to undergo a pre-operative MRI.[21,22]

Table 1. Adjustments to the MISCAN model to simulate decreases in breast density

% of women with VDG4 at age 50	VDG density according to age			Factor difference of the sensitivity of mammography[1]			Factor difference of the onset of breast cancer[16]			Factor difference of the probability of a false positive mammogram[1]		
	50-54	55-64	65+	50-54	55-64	65+	50-54	55-64	65+	50-54	55-64	65+
21.9%	4	4	4	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
18.8%	4	4	3	Ref.	Ref.	1.14	Ref.	Ref.	0.94	Ref.	Ref.	0.76
40.8%	4	3	3	Ref.	1.14	1.14	Ref.	0.94	0.94	Ref.	0.76	0.76
18.4%	4	3	2	Ref.	1.14	1.27	Ref.	0.94	0.60	Ref.	0.76	0.63

Reference values were calibrated

Costs

We applied a health care perspective and only considered direct medical costs and costs related to other causes of death.

Most unit prices were derived from a previous cost-effectiveness study.[10] The price of tomosynthesis was assumed to be equal to mammography within a hospital setting. Unit prices are shown in Table S4.

A telephone consult with the general practitioner was modelled after a positive screening result, which reflects current practice in the Netherlands.

Mean treatment costs were calculated using previously published prices[10] and the quantity of each treatment type per T-stage in 2011 (Table S5). These data were obtained from the Netherlands Comprehensive Cancer Registration (IKNL). Subsequently, mean treatment costs were multiplied by the modelled numbers of tumors by T-stage (Table S6).

Costs in the last year of life were derived from Polder et al[23] and converted to the price level of 2018.

Utilities

Utility values (quality of life) were obtained from the literature (Table S7). A disutility of 10% was applied for DCIS, and a disutility of and 25% for regional breast cancer[24], with durations of 5 years. A disutility of 50% was applied for metastasized breast cancer until death.[24] We applied a disutility associated with screening participation of 0.006 for 1 week, and a disutility associated with a positive screen of 0.105 for 5 weeks.[25]

Screening strategies

Several screening strategies were simulated with varying intervals (see supplementary figures). The MRI participants and control group of the DENSE trial correspond with '2Mx_2MRI' (mammography plus MRI at a two-year interval) and '2Mx' (mammography at a two-year interval), respectively.

Each modelled strategy started with mammography at age 50, because women always undergo mammography as their first screening (as density is unknown). In case of a decrease in breast density from category 4 to 3, women switched to mammography at a two-year interval. This switch was assumed to take place after the first screening following the modelled breast density drop. We assumed that breast density can be measured with mammography and MRI. In the screening strategies containing mammography and MRI together in one screening round, we assumed women undergo the mammogram first and the MRI one month later, which allows for a cancellation of the MRI when a drop in breast density was shown on the mammogram. A screening attendance rate of 100% was modelled.

Analyses

A cohort of 10 million Dutch women, born in 1965, was simulated from age 25 until death. Outcomes were the number of screening mammograms, screening MRIs, screen-detected cancers, interval cancers, life-years (LYs), quality-adjusted life-years (QALYs), deaths from breast cancer and deaths from other causes. Overdiagnosis was defined as detected cancers that would not have been diagnosed in a woman's lifetime in a situation without screening. Strategies were ranked by total costs. A strategy with fewer QALYs than the previous strategy in the ranking was considered 'strongly dominated'. The incremental cost-effectiveness ratio (ICER) of a strategy in comparison with the previous strategy was calculated by dividing incremental costs by incremental LYs and incremental QALYs. A strategy with a higher ICER than the next strategy was considered 'weakly dominated'. All results were scaled to 1000 women. Both costs and effects were discounted at 3%. A willingness-to-pay threshold of €22,000 (£20,000) was applied, based on the lower bound of the National Institute for Health and Care Excellence (NICE) threshold range.[26]

One-way sensitivity analyses were performed by varying utility values ±10%, probabilities of diagnostic procedures after a positive mammogram ±25%, and the price and false positive rate of MRI ±25%. Furthermore, we adjusted the price of tomosynthesis +25% because tomosynthesis may be more expensive. Since axillary ultrasound is relatively often performed in the Netherlands, we performed a sensitivity analysis in which this was only performed after a proven malignancy. We applied these adjustments separately to all strategies to analyze the effect on the ICERs.

Three scenario analyses were performed to quantify methodological uncertainty. First, we applied a discount rate of 4.0% for costs and 1.5% for effects, based on Dutch guidelines.[27] Second, we assumed breast cancer incidence would not decrease with decreasing breast density.[28] Third, we applied different utility values (see Table S8).

RESULTS

Calibration results

Figures S1-S6 show the numbers observed and simulated screen-detected and interval cancers by T-stage. Most of the simulated numbers fell within the 95% confidence intervals of the observed numbers. Overall, the simulated tumor size in the model was slightly larger than observed. The number of screen-detected T2+ tumors in de MRI-group was overestimated, as well as the number of interval T2+ tumors in the control group. The number of interval T1C tumors in the control group was underestimated.

Outcomes of different screening strategies

Discounted and undiscounted results are shown in Table 2, in order of lowest to highest total costs. Biennial mammography alone resulted in 69 screen-detected breast cancers and 43 breast cancer deaths per 1000 women. Adding MRI every other screening round (2Mx_4MRI) resulted in 24 additional screen-detected cancers and 7 fewer breast cancer deaths. The addition of MRI every screening round (2Mx_2MRI) resulted in another 4 additional screen-detected cancers and 1 fewer breast cancer death. Leaving out mammography, MRI alone every two years, yielded 100 screen-detected cancers, and 97 screen-detected cancers when offered every three years. Numbers of overdiagnosis were similar across all strategies containing MRI: equaling 20-21 cases, compared to 17 with biennial mammography. When moving from the strategy consisting of alternating mammography and MRI at a two-year interval (2Mx/MRI) to a more expensive strategy, no additional breast cancer deaths were averted.

Screening every two years with mammography alone (2Mx) resulted in the lowest total costs and the lowest number of QALYs compared to all other screening strategies (Figure 1). Additional MRI every two years (2Mx_2MRI) resulted in the highest costs but

not the highest number of QALYs and was therefore strongly dominated. Most strategies containing mammography were dominated, because of the limited sensitivity of mammography compared to MRI. However, alternating mammography and MRI every 2 years was close to the efficiency frontier. Screening with MRI alone was efficient with various intervals. Lengthening the intervals resulted in lower total costs, and only a few cancers not being screen-detected. When applying the NICE threshold, quadrennial MRI (4MRI) had the highest acceptable ICER. Screening every three years with MRI alone resulted in an ICER of €37,181 per QALY gained.

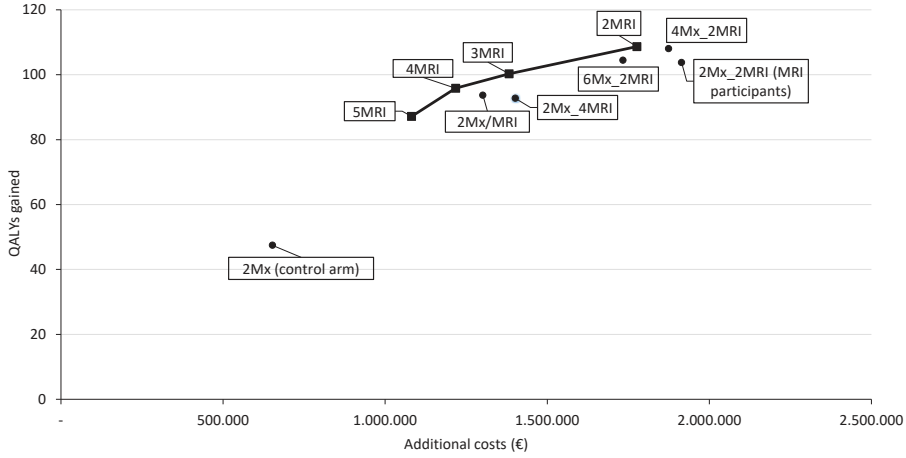


Figure 1. Efficiency frontier (3% discounting)
2Mx: mammography every two years; 5MRI: MRI every five years; 4MRI: MRI every four years; 2Mx/MRI: screening every two years with alternating mammography and MRI; 3MRI: MRI every three years; 2Mx_4MRI: mammography every two years and MRI every four years; 6Mx_2MRI: mammography every six years and MRI every two years; 2MRI: MRI every two years; 4Mx_2MRI: mammography every four years and MRI every two years; 2Mx_2MRI: MRI and mammography every two years

Sensitivity and scenario analyses

In all one-way sensitivity analyses, MRI screening every four years remained cost-effective with the highest acceptable ICER (Table 3). The ICERs were most sensitive to the unit price of MRI.

When applying discount rates of 1.5% for effects and 4.0% for costs, the ICERs became lower. The strategy consisting of quadrennial MRI screening remained the highest acceptable ICER of €9,836 per QALY gained.

Table S9 shows the results when the breast cancer incidence was assumed not to decrease with decreasing breast density. Overall, more cancers were detected among all strategies but the ICERs were fairly similar as those presented in Table 2.

When applying a different set of utility values, the ICERs remained similar as well.

Table 2. Results of all screening strategies, per 1000 women from age 25 until death

	No Screening	2Mx	5MRI	4MRI	2Mx/MRI	3MRI	2Mx_4MRI	6Mx_2MRI	2MRI	4Mx_2MRI	2Mx_2MRI
Undiscounted results											
Tumors	152	169	173	173	173	173	172	173	173	173	173
Screen-detected	-	69	94	96	94	97	93	98	100	99	97
Overdiagnosis [a]	-	17	21	21	21	21	20	21	21	21	21
Breast cancer deaths	54	43	37	35	36	35	36	34	34	34	35
False positives	-	141	203	217	245	239	270	288	283	308	330
Number of mammograms	-	11,114	6,483	5,328	8,735	6,006	11,024	7,446	5,865	8,137	11,006
Number of MRIs	-	-	2,891	3,779	3,242	4,324	3,079	5,542	6,088	5,926	5,383
Screening (€)	-	768,004	1,234,301	1,395,979	1,485,393	1,591,067	1,599,214	2,022,061	2,061,187	2,174,171	2,224,690
Diagnostics (€)	207,544	222,857	238,966	244,556	260,337	256,639	273,806	284,184	280,881	294,094	298,100
Treatment (€)	1,230,641	1,338,814	1,374,366	1,366,305	1,362,604	1,362,531	1,360,721	1,355,903	1,355,403	1,354,624	1,354,264
Breast cancer death (€)	1,512,425	1,209,474	1,030,535	994,921	1,009,505	982,002	1,014,825	966,768	949,821	952,233	969,311
Death other causes (€)	20,281,602	20,512,479	20,648,984	20,676,144	20,665,053	20,685,991	20,660,998	20,697,618	20,710,555	20,708,688	20,695,666
Discounted results [b]											
Total costs (€)	10,029,132	10,681,842	11,110,699	11,246,367	11,330,895	11,411,722	11,431,163	11,763,234	11,805,633	11,903,811	11,943,649
Life-years	57,816	57,870	57,913	57,921	57,919	57,926	57,918	57,929	57,933	57,933	57,929
QALYs	49,473	49,520	49,560	49,569	49,566	49,573	49,565	49,577	49,581	49,581	49,576
ICER (€/QALY)	n/a	Weakly dominated	12,410	15,620	Strongly dominated	37,181	Strongly dominated	Weakly dominated	46,971	Strongly dominated	Strongly dominated

[a]Due to rounding, the numbers of overdiagnosis may not be precisely the difference in the numbers of tumors with and without screening as shown in this table.

[b] Discounted by 3%

2Mx: mammography every two years; 5MRI: MRI every five years; 4MRI: MRI every four years; 2Mx/MRI: screening every two years with alternating mammography and MRI; 3MRI: MRI every three years; 2Mx_4MRI: mammography every two years and MRI every four years; 6Mx_2MRI: mammography every six years and MRI every two years; 2MRI:MRI every two years; 4Mx_2MRI: mammography every four years and MRI every two years; 2Mx_2MRI: MRI and mammography every two years

Note that some differences in outcomes between strategies do not seem logical, but these are caused by mammography showing a possible drop in breast density: in strategies containing mammography and MRI in one screening round, a decrease in density was detected by mammography, resulting in dropping the subsequent MRI in that same screening round. In strategies not containing mammography, a drop in density was observed on the MRI. Therefore, in strategy 2MRI, more MRIs were performed compared to 2Mx_2MRI.

Table 3. Results of the one-way sensitivity analyses and scenario analyses

	Value in sensitivity analysis	Strategy with the highest acceptable ICER	ICER (€/QALY)
Unit cost MRI +25%	€340	4MRI	21,267
Unit cost MRI -25%	€204	4MRI	10,074
Utility value DCIS/localized breast cancer +10%	0.849	4MRI	14,722
Utility value DCIS/localized breast cancer -10%	0.695	4MRI	16,749
Utility value regional breast cancer +10%	0.708	4MRI	16,243
Utility value regional breast cancer -10%	0.579	4MRI	15,137
Utility value metastasis +10%	0.566	4MRI	15,983
Utility value metastasis -10%	0.463	4MRI	15,369
Probabilities of a false-positive MRI result +10%	First MRI: 9.8% Subsequent MRIs: 3.3%	4MRI	16,013
Probabilities of a false-positive MRI result -10%	First MRI: 5.9% Subsequent MRIs: 1.9%	4MRI	15,331
Unit cost tomosynthesis +25%	€115	4MRI	15,590
Probability diagnostic ultrasound after a positive mammogram +25%	100%	4MRI	15,602
Probability diagnostic ultrasound after a positive mammogram -25%	66%	4MRI	15,653
Probability stereotactic biopsy after a positive mammogram +25%	11%	4MRI	15,608
Probability stereotactic biopsy after a positive mammogram -25%	6%	4MRI	15,632
Probability ultrasound-guided biopsy after a positive mammogram +25%	48%	4MRI	15,590
Probability ultrasound-guided biopsy after a positive mammogram -25%	29%	4MRI	15,650
Ultrasound axilla only after a proven malignancy	17%	4MRI	15,627
Scenario analyses			
No decrease in breast cancer incidence	n/a	4MRI	15,467
Different discount rates	Costs: 4% Effects: 1.5%	4MRI	9,836 [a]
Different set of utility values	See Table S5.	4MRI	15,955

[a] The ICER of the next strategy (3MRI) on the frontier was just above the threshold with a value of 24,835/QALY
4MRI: MRI every four years

DISCUSSION

The aim of this study was to evaluate the cost-effectiveness of several screening strategies containing (additional) MRI screening for women with extremely dense breast tissue.

We found that using screening with MRI alone every four years resulted in the highest acceptable ICER when applying the NICE threshold. When applying a higher threshold, MRI at an interval of two or three years can be considered cost-effective too. Strategies containing mammography were dominated due to more clinically diagnosed cancers, resulting in more breast cancer deaths and less QALYs, compared to strategies with MRI.

To our knowledge, this is the first study evaluating the cost-effectiveness of MRI screening for women with extremely dense breasts. One previous study evaluated costs and QALYs associated with MRI but there was no comparison strategy and they used a relatively short time horizon.[29] In prior cost-effectiveness studies, either the target groups were high risk women,[7-10] or the cost-effectiveness of screening modalities other than MRI were evaluated.[30,31] Shortening the screening interval of mammography from two years to one year was shown not to be cost-effective for women with dense breasts.[30] Additional ultrasonography after a negative mammogram was also not cost-effective due to relatively small benefits and high costs.[31] A study by Lee et al. concluded that a combination of tomosynthesis and mammography is likely to be cost-effective for this group of women, with an ICER of $54,000.[32]

An important strength of this study is the use of data of incident and prevalent screening rounds of a large randomized controlled trial. In addition, we used a well-established microsimulation model to extrapolate the findings of this trial. By calibration, dwell times and sensitivities of mammography and MRI were estimated, which allowed us to model several screening strategies, expanding the DENSE trial. An important limitation is that most of our MRI detection estimates were higher than the observed numbers. However, most numbers were within the confidence limits of the observed data. This was not the case for T2+ and T1C tumors. We overestimated the number of screen-detected T1C tumors and underestimated the number of screen-detected T2+ tumors by mammography in the control arm. Also, we overestimated the number of MRI-detected T2+ tumors, although the number of estimated T2+ interval tumors was within the confidence limits. Overall, we expect this to result in conservative model predictions for the effects of MRI screening, mainly due to the overestimated number of screen-detected T2+ tumors by MRI, since T2+ tumors are associated with a relatively poor survival. The fact that numbers of interval cancers during the second round were unknown is also a limitation. By varying dwelling times and sensitivities of mammography and MRI, we performed several calibrations of which the fit closest to the target outcomes was used in our analyses. Another limitation may be that we assumed that pre-operative MRIs were performed in 26%-38% of the detected tumors,[21,22] independent of the detection mode. However, in reality, when a tumor is detected by MRI, a pre-operative MRI may not be necessary anymore. We also had to make several assumptions on diagnostic procedures after a positive mammogram, but these hardly affected the ICERs.

Based on our results, screening with MRI alone every four years would be recommended from a cost-effectiveness perspective. However, when women know they have extremely dense breasts and thereby an elevated breast cancer risk, they may want to be screened more often than once every four years. This may result in opportunistic screening. Opportunistic screening, however, was not incorporated in our model. In case a two-year interval is preferred by policy makers, alternating mammography and MRI can be an alternative.

Approximately 8% of women aged 50-75 years have extremely dense breasts.[1] Even though we showed MRI screening is cost-effective for these women, it would create a burden on health care budgets. Furthermore, screening these women within a hospital setting may lead to capacity problems. Implementation of MRI screening would lead to a need of more MRI machines and more (trained) personnel.

We modelled only women with extremely dense breasts getting MRI screening, but the sensitivity of mammography is also low among women with heterogeneously dense breasts (VDG3).[1] However, also expanding MRI screening to these women will create a larger burden on health care budgets and screening capacities, as 29% of the screening population falls in this category.[1] Also, the benefit of MRI may be lower for this group because the sensitivity of mammography is higher among women with heterogeneously dense breasts compared to women with extremely dense breasts.[1]

We modelled Dutch women with extremely dense breasts within the Dutch health care setting but we assume that relative differences in health outcomes between the modelled screening strategies will be approximately similar in other countries. Since health care prices and cost-effectiveness thresholds vary per country, this should be kept in mind when translating our ICERs to other countries.

In our analyses, the unit cost of MRI was €272, and the ICER was highly sensitive to this. In the near future, we expect several technological developments, such as artificial intelligence and abbreviated MRI, which could reduce false positive diagnoses, and both acquisition and reading time.[33,34] Therefore, we expect MRI screening to become less expensive in the future.

In conclusion, this study showed that MRI screening every four years for women with extremely dense breast tissue was cost-effective and had the highest acceptable ICER. If decision-makers are willing to pay more than €22,000 per QALY gained, MRI every three or two years can also become cost-effective.

Acknowledgements

We thank the trial participants for their contributions; the regional screening organizations, Volpara Health Technologies, the Dutch Expert Center for Screening, and the National Institute for Public Health and the Environment for their advice and in-kind contributions. We thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry.

REFERENCES

1. Wanders JO, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat.* 2017;162(1):95-103.
2. Price ER, Hargreaves J, Lipson JA, et al. The California breast density information group: a collaborative response to the issues of breast density, breast cancer risk, and breast density notification legislation. *Radiology.* 2013;269(3):887-892.
3. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med.* 2007;356(3):227-236.
4. Kerlikowske K. The mammogram that cried Wolfe. *N Engl J Med.* 2007;356(3):297-300.
5. Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. *Ann Intern Med.* 2008;148(9):671-679.
6. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.
7. Pataky R, Armstrong L, Chia S, et al. Cost-effectiveness of MRI for breast cancer screening in BRCA1/2 mutation carriers. *BMC Cancer.* 2013;13:339.
8. Plevritis SK, Kurian AW, Sigal BM, et al. Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. *JAMA.* 2006;295(20):2374-2384.
9. Lee JM, McMahon PM, Kong CY, et al. Cost-effectiveness of breast MR imaging and screen-film mammography for screening BRCA1 gene mutation carriers. *Radiology.* 2010;254(3):793-800.
10. Geuzinge HA, Obdeijn IM, Rutgers EJ, et al. Cost-effectiveness of breast cancer screening with MRI in women at familial risk: the randomized FaMRIsc trial. *JAMA Oncol.* 2020;6(9):1381-1389.
11. American College of Radiology. Breast Imaging Reporting and Data System Atlas (BIRADS atlas*)*. Reston, VA; 2003.
12. Emaus MJ, Bakker MF, Peeters PH, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. *Radiology.* 2015;277(2):527-537.
13. Sankatsing VDV, Heijnsdijk EAM, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. *International Journal of Cancer.* 2015;137(8):1990-1999.
14. American Joint Committee on Cancer (AJCC). *AJCC Cancer Staging Manual.* 8th ed. New York, NY: Springer; 2017.
15. Wanders JOP, Holland K, Karssemeijer N, et al. Changes in volumetric breast density and the association with breast cancer risk. *Submitted.*
16. Roman M, Sala M, Bare M, et al. Changes in mammographic density over time and the risk of breast cancer: An observational cohort study. *Breast.* 2019;46:108-115.
17. Netherlands Comprehensive Cancer Organisation (IKNL). Monitor bevolkingsonderzoek borstkanker 2017/2018. 2018.
18. Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM, et al. The Breast Imaging Reporting and Data System (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. *Eur Radiol.* 2012;22(8):1717-1723.
19. Weigel S, Decker T, Korsching E, Hungermann D, Bocker W, Heindel W. Calcifications in digital mammographic screening: improvement of early detection of invasive breast cancers? *Radiology.* 2010;255(3):738-745.

20. Nationaal Borstkanker Overleg Nederland (NABON). Borstkanker, landelijke richtlijn, versie 2.0. 2020; https://www.lrcb.nl/resources/uploads/2017/02/Richtlijn-Mammacarcinoom.pdf.
21. Keymeulen KBIM, Geurts SME, Lobbes MBI, et al. Population-based study of the effect of preoperative breast MRI on the surgical management of ductal carcinoma in situ. *Br J Surg.* 2019;106(11):1488-1494.
22. Lobbes MB, Vriens IJ, van Bommel AC, et al. Breast MRI increases the number of mastectomies for ductal cancers, but decreases them for lobular cancers. *Breast Cancer Res Treat.* 2017;162(2):353-364.
23. Polder JJ, Barendregt JJ, van Oers H. Health care costs in the last year of life--the Dutch experience. *Soc Sci Med.* 2006;63(7):1720-1731.
24. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst.* 2006;98(11):774-782.
25. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer.* 1991;49(4):538-544.
26. Appleby J, Devlin N, Parkin D. NICE's cost effectiveness threshold. *BMJ.* 2007;335(7616):358-359.
27. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Tan S. Kostenhandleiding: Methodologie van Kostenonderzoek En Referentieprijzen Voor Economische Evaluaties in de Gezondheidszorg. Diemen 2015.
28. Kerlikowske K, Ichikawa L, Miglioretti DL, et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. *J Natl Cancer Inst.* 2007;99(5):386-395.
29. Froelich MF, Kaiser CG. Cost-effectiveness of MR-mammography as a solitary imaging technique in women with dense breasts: an economic evaluation of the prospective TK-Study. *Eur Radiol.* 2020.
30. 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. *Ann Intern Med.* 2011;155(1):10-20.
31. Sprague BL, Stout NK, Schechter C, et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. *Ann Intern Med.* 2015;162(3):157-166.
32. Lee CI, Cevik M, Alagoz O, et al. Comparative effectiveness of combined digital mammography and tomosynthesis screening for women with dense breasts. *Radiology.* 2015;274(3):772-780.
33. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers RD, Bieling HB. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection-a novel approach to breast cancer screening with MRI. *J Clin Oncol.* 2014;32(22):2304-2310.
34. Verburg E, van Gils CH, Bakker MF, et al. Computer-Aided Diagnosis in Multiparametric Magnetic Resonance Imaging Screening of Women With Extremely Dense Breasts to Reduce False-Positive Diagnoses. *Invest Radiol.* 2020;55(7):438-444.

SUPPLEMENTARY APPENDIX

DENSE trial study group

University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands:
CH van Gils PhD, MF Bakker PhD, SV de Lange MD, SGA Veenhuizen MSc, WB Veldhuis MD PhD, RM Pijnappel MD PhD, MJ Emaus PhD, PHM Peeters MD PhD, EM Monninkhof PhD, MA Fernandez-Gallardo MD, WPThM Mali MD PhD, MAAJ van den Bosch MD PhD, PJ van Diest MD PhD

Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands:
RM Mann MD PhD, Roel Mus MD, N Karssemeijer PhD, M. Imhof-Tas MD

The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands:
CE Loo MD PhD, PK de Koekkoek-Doll MD, HAO Winter-Warnars MD PhD

Albert Schweitzer Hospital, Dordrecht, the Netherlands:
RHC Bisschops MD PhD, MCJM Kock MD PhD, RK Storm MD, PHM van der Valk MD

Maastricht University Medical Centre, Maastricht, the Netherlands:
MBI Lobbes MD PhD, S Gommers MD

Jeroen Bosch Hospital, 's-Hertogenbosch, the Netherlands:
MDF de Jong MD, MJCM Rutten MD PhD

Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands:
KM Duvivier MD, P de Graaf MD PhD

Hospital Group Twente (ZGT), Almelo, the Netherlands:
J Veltman MD PhD, RLJH Bourez MD
Erasmus Medical Center, Rotterdam, the Netherlands:
HJ de Koning MD PhD

Data used for calibration

Control arm:

- Interval cancers during the first round, according to age (5-year age groups)
- Interval cancers according to T-stage during the first round
- Screen-detected cancers by mammography according to T-stage during the second round

MRI participants:

- Screen-detected cancers by MRI according to T-stage during the first round
- Interval cancers according to T-stage during the first round
- Screen-detected cancers by MRI according to age during the first round
- Screen-detected cancers by mammography according to T-stage during the second round
- Screen-detected cancers by MRI according to T-stage during the second round

Results of the first round of the DENSE trial have been published[1], and results of the second round are currently submitted for publication.

Table S1. Probabilities of a false-positive test result per screen, by screening modality and screening round

	First round		Second round	
	Mammography	**MRI**	**Mammography**	**MRI**
MRI participants	4.0%*	7.8%	1.3%	2.6%
Control arm	4.0%*	n/a	1.3%	n/a

*Obtained from the literature in same source population as the DENSE trial[2]

Table S2. Probabilities of diagnostic procedures by screening modality and screening round

	First round, after a positive MRI		Second round, after a positive MRI		
	BI-RADS 3*	**BI-RADS 4/5***	**BI-RADS 3***	**BI-RADS 4/5***	**Positive mammogram**
Diagnostic mammogram/ tomosynthesis	3.3%	19.5%	0%	17.5%	100%
Ultrasound	8.6%	38.1%	14.3%	30.2%	88%
Ultrasound-guided biopsy	11.2%	58.9%	0%	63.5%	38%
Diagnostic MRI	76.3%	11.3%	71.4%	23.8%	0%
MRI-guided biopsy	9.9%	44%	14.3%	39.7%	0%
Stereotactic biopsy	0%	0%	0%	0%	8%

* Observed in the DENSE trial[1]

Percentages equal more than 100% because multiple tests can be performed after a positive MRI.

Tabel S3. Calculations diagnostic procedures

Diagnostic procedure	Calculation/explanation
Ultrasound	Based on expert opinion, we assumed all women with a BI-RADS 4/5 mammography result to get an ultrasound, and 80% of the women with a BI/RADS 0 result. Based on data of the whole Dutch breast cancer screening program, the distribution of BI-RADS 4/5 versus BI-RADS 0 results is 50%:50%.[3] Based on expert opinion, we assumed this distribution to be slightly different among women with extremely dense breasts: 40%:60% - BI-RADS 0: 0.6*0.8=48% - BI-RADS 4/5: 0.4*1.0= 40% - In total 88%
Biopsies	A study by Timmers et al.[4] showed the percentage of women getting a biopsy according to BI-RADS category (mammography result). In our opinion, the percentage of women with a BI-RADS 0 mammography results undergoing a biopsy was quite high (i.e. 47.3%) so we assumed this to be 30% in the current setting.[3] Using this, in combination with the previously mentioned distribution of BI-RADS 4/5 vs. 0 results, we calculated the weighted average of women getting a biopsy: - BI-RADS 0: 0.3*0.6=18% - BI-RADS 4: 0.633 * (578/(578+156))*0.4=19.9% - BI-RADS 5: 0.955*(156/(578+156))*0.4 = 8.1% - In total 46% The type of biopsy radiologists conduct depends on what is seen on the mammogram: calcification or a lesion. Based on a study by Wiegel et al.[5] we assumed this distribution to be 23% calcifications vs. 77% lesions. Furthermore, we assumed 80% of the calcifications to get a stereotactic biopsy, and 20% of the calcifications and 100% of the lesions to get a ultrasound-guided biopsy, based on expert opinion. - Stereotactic biopsy: 0.46*0.227*0.80=8.4% - Ultrasound-guided biopsy: 0.46*((0.227*0.2)+0.773)=38%

Table S4. Unit prices

Procedure	Price (€)	Source
Screening, diagnosis and staging		
Mammography	91.97[a]	CZ tariff tool[6]
Mammography	69.10[b]	IKNL[7]
Tomosynthesis	91.97	Assumption
MRI	272.00	CZ tariff tool[6]
Telephone consultation with general practitioner	17.69	Dutch costing manual[8]
Standard consultation with general practitioner	34.34	Dutch costing manual[8]
Ultrasound	115.23	CZ tariff tool[6]
Biopsy, ultrasound-guided	246.44	Erasmus MC; CZ tariff tool[6]
Biopsy, MRI-guided	599.41	Erasmus MC; CZ tariff tool[6]
Biopsy, stereotactic	452.31	Erasmus MC; CZ tariff tool[6]
PET-CT	1069.76	Erasmus MC
Surgery		
Breast conserving surgery	1452.49	Saadatmand et al[9]; Dutch costing manual[8]
Mastectomy	1623.35	Saadatmand et al[9]; Dutch costing manual[8]
Adjuvant therapy		
Radiotherapy	6885.05	Saadatmand et al[9]
Chemotherapy	3573.21	Saadatmand et al[9]
Hormonal therapy	2574.81	Saadatmand et al[9]

a. Mammography at a hospital (screening and diagnostic)

b. Mammography within the national breast cancer screening program

All prices are converted to Euro 2018 prices using consumer indices[10]

Table S5. Treatment according to T-stage (Dutch data from 2011, IKNL)

	DCIS	T1AN-	T1AN+*	T1BN-	T1BN+	T1CN-	T1CN+	T2+N-	T2+N+
Hormonal therapy	0%	2%	27%	3%	16%	15%	23%	28%	27%
Chemotherapy	0%	2%	54%	6%	79%	56%	81%	61%	59%
Radiotherapy	58%	69%	44%	82%	79%	75%	72%	59%	62%
Mastectomy	35%	30%	38%	17%	23%	23%	32%	46%	59%
Breast conserving surgery	64%	69%	25%	82%	73%	76%	64%	51%	29%
Other/no treatment	1%	0%	25%	0%	1%	0%	1%	1%	3%

*Also includes T0N+ and T1 micro invasive

Women can get more than one type of treatment so the total of all percentages by T stage is greater than 100%.

Table S6. Treatment costs according to T-stage

	Costs per stage (€)
DCIS	5,520
T1AN-	6,376
T1AN+	6,617
T1BN-	7,441
T1BN+	10,110
T1CN-	9,073
T1CN+	9,901
T2+N-	8,480
T2+N+	8,448

Table S7. Utility values and durations

Health state	Utility value	Duration	Source
No breast cancer	0.858	n.a.	Versteegh et al[11]
After a (false) positive screening result	0.105 (disutility)	5 weeks	de Haes et al[12]
After undergoing screening	0.006 (disutility)	1 week	de Haes et al[12]
DCIS/localized breast cancer	0.772	5 years	Stout et al[13]
Regional breast cancer	0.644	5 years	Stout et al[13]
Metastasis	0.515	Until death	Stout et al[13]
Death	0		

Table S8. Alternative utility values and durations used in scenario analysis

Health state	Utility value	Duration	Source
No breast cancer	0.858	n.a.	Versteegh et al[11]
After a (false) positive screening result	0.105 (disutility)	5 weeks	de Haes et al[12]
After undergoing screening	0.006 (disutility)	1 week	de Haes et al[12]
DCIS/localized breast cancer, first year	0.696	1 year	Lidgren et al[14]
DCIS/localized breast cancer, after the first year	0.779	10 years	Lidgren et al[14]
Regional breast cancer	0.685	11 years	Lidgren et al[14]
Metastasis	0.515	Until death	Lidgren et al[14]
Death	0		

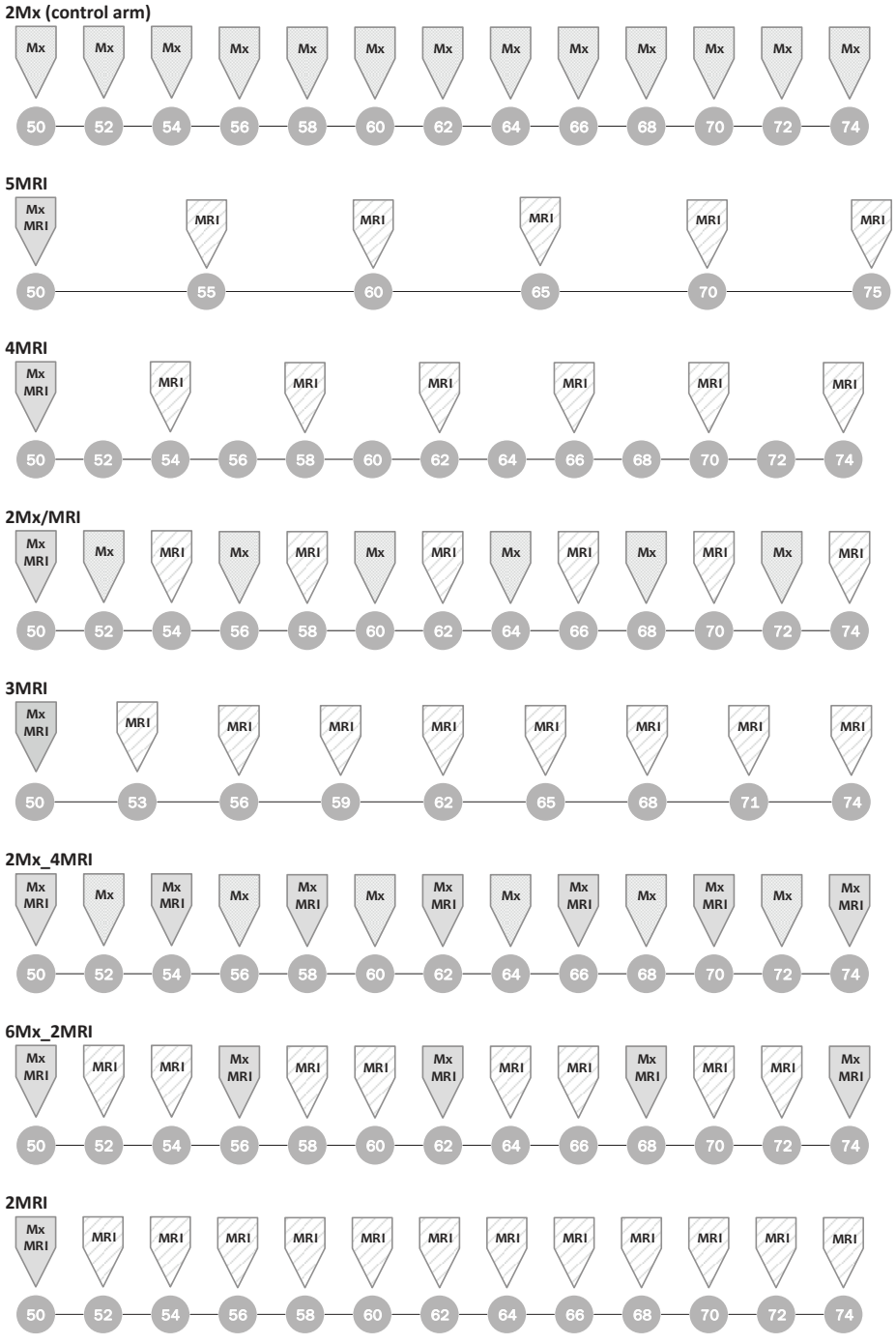
Table S9. Results of all screening strategies, modelled without a decrease in breast cancer risk, per 1000 women

	No Screening	2Mx	5MRI	4MRI	2Mx/MRI	3MRI	2Mx_4MRI	2MRI	6Mx_2MRI	4Mx_2MRI	2Mx_2MRI
Undiscounted results											
Tumors	163	181	186	185	185	185	185	185	185	185	185
Screen-detected	n/a	74	99	101	99	102	98	105	103	104	102
Overdiagnosis	n/a	18	23	22	22	22	22	22	22	22	22
Breast cancer deaths	57	45	39	38	38	37	38	36	37	36	37
Discounted results*											
Total costs (€)	10,093,306	10,747,990	11,176,466	11,311,654	11,396,124	11,476,749	11,496,511	11,807,894	11,828,006	11,968,418	12,024,719
Life-years	57,805	57,861	57,905	57,914	57,911	57,918	57,910	57,923	57,921	57,925	57,922
QALYs	49,459	49,508	49,548	49,557	49,555	49,562	49,554	49,568	49,566	49,570	49,567
ICER (€/QALY)	n/a	Weakly dominated	12,042	15,467	Strongly dominated	35,166	Strongly dominated	53,766	Strongly dominated	96,830	Strongly dominated

*Discounted by 3%

2Mx:mammography every two years; 5MRI:MRI every five years; 4MRI:MRI every four years; 2Mx/MRI: screening every two years with alternating mammography and MRI; 3MRI: MRI every three years; 2Mx_4MRI: mammography every two years and MRI every four years; 2MRI:MRI every two years; 6Mx_2MRI: mammography every 6 years and MRI every 2 years; 4Mx_2MRI:MRI every two years and mammography every four years; 2Mx_2MRI: MRI and mammography every two years

All screening strategies visualized



4Mx_2MRI

2MRI_2Mx (MRI participants)

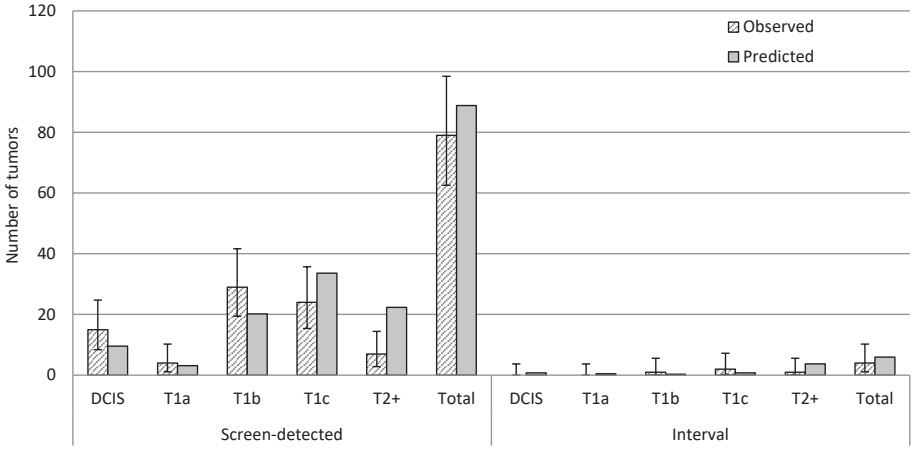


Figure S1. Numbers of screen-detected tumors and interval tumors in the MRI participants during the first round. 95% confidence intervals are presented around the observed data.

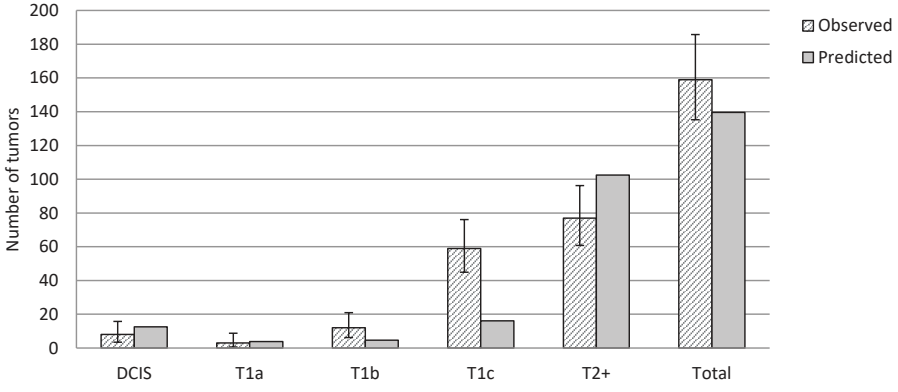


Figure S2. Numbers of interval tumors in the control arm during the first round. 95% confidence intervals are presented around the observed data.

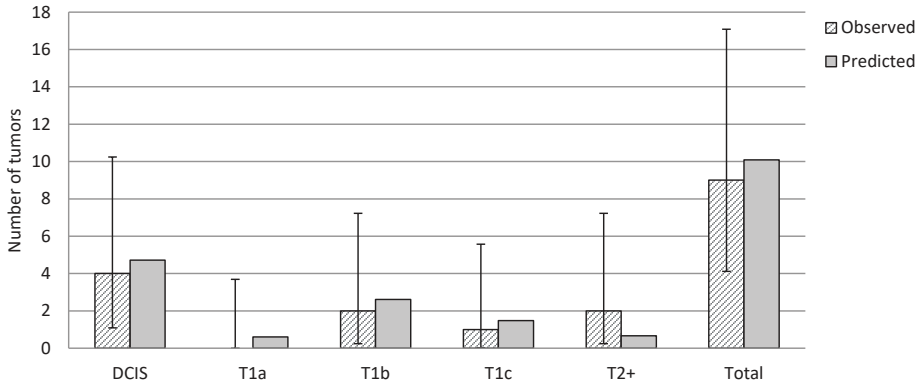


Figure S3. Numbers of screen-detected tumors by mammography in the MRI participants during the second round. 95% confidence intervals are presented around the observed data.

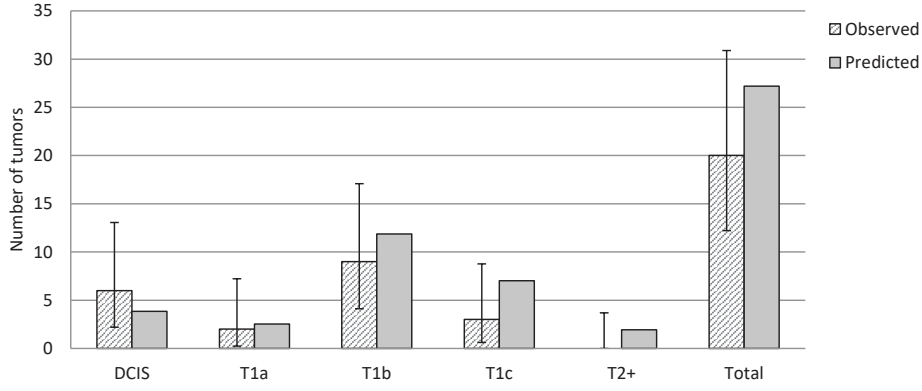


Figure S4. Numbers of screen-detected tumors by MRI in the MRI participants during the second round. 95% confidence intervals are presented around the observed data.

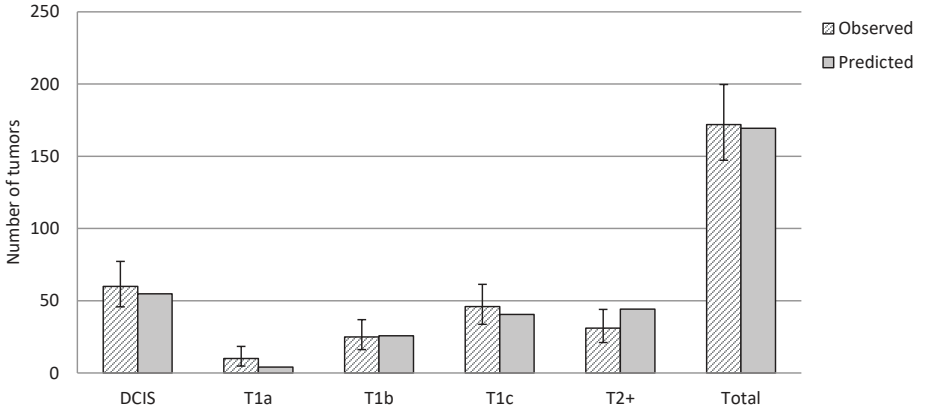


Figure S5. Numbers of screen-detected tumors by mammography in the control arm during the second round. 95% confidence intervals are presented around the observed data.

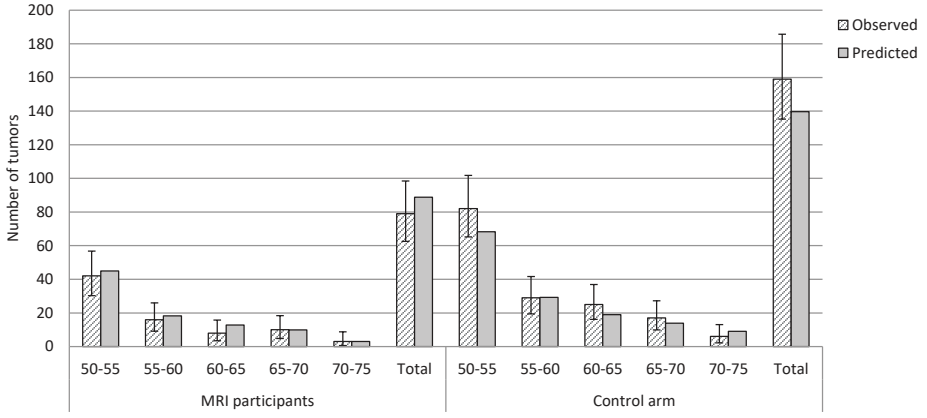


Figure S6. Numbers of tumors according to age during the first round. 95% confidence intervals are presented around the observed data.

References

1. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.
2. Wanders JO, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat.* 2017;162(1):95-103.
3. Netherlands Comprehensive Cancer Organisation (IKNL). Monitor bevolkingsonderzoek borstkanker 2017/2018. 2018.
4. Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM, et al. The Breast Imaging Reporting and Data System (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. *Eur Radiol.* 2012;22(8):1717-1723.
5. Weigel S, Decker T, Korsching E, Hungermann D, Bocker W, Heindel W. Calcifications in digital mammographic screening: improvement of early detection of invasive breast cancers? *Radiology.* 2010;255(3):738-745.
6. CZ. Tarieventool. https://www.cz.nl/service-en-contact/zoek-tarieven. Accessed October 3, 2018.
7. Netherlands Comprehensive Cancer Organisation (IKNL). Monitor bevolkingsonderzoek borstkanker 2016. 2018; https://www.rivm.nl/sites/default/files/2018-11/Monitor%20bevolkingsonderzoek%20borstkanker%202016%20webversie.pdf. Accessed 9 July 2019.
8. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Tan S. Kostenhandleiding: Methodologie van Kostenonderzoek En Referentieprijzen Voor Economische Evaluaties in de Gezondheidszorg. Diemen 2015.
9. Saadatmand S, Tilanus-Linthorst MMA, Rutgers EJT, et al. Cost-Effectiveness of Screening Women With Familial Risk for Breast Cancer With Magnetic Resonance Imaging. *J Natl Cancer Inst.* 2013;105(17):1314-1321.
10. Statistics Netherlands. Annual rate of change CPI; since 1963. 2018; http://statline.cbs.nl/Statweb/publication/?DM=SLEN&PA=70936ENG&D1=0&D2=623,636,649,662,675,688,701,714&LA=EN&VW=T. Accessed 9 July 2019.
11. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. *Value Health.* 2016;19(4):343-352.
12. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer.* 1991;49(4):538-544.
13. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst.* 2006;98(11):774-782.
14. Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Qual Life Res.* 2007;16(6):1073-1081.

Part III

What do women want?

Chapter 7

Experiences, expectations and preferences regarding MRI and mammography as breast cancer screening tools in women at familial risk

HA Geuzinge, EAM Heijnsdijk, AIM Obdeijn, HJ de Koning, MMA Tilanus-Linthorst, on behalf of the FaMRIsc-study group

The Breast. 2021;56:1-6

ABSTRACT

Background: Several studies have investigated MRI breast cancer screening in women at increased risk, but little is known about their preferences. In this study, experiences, expectations and preferences for MRI and mammography were evaluated among women undergoing screening with MRI and/or mammography in the randomized FaMRIsc trial.

Methods: A 17-item questionnaire was sent to 412 women in the FaMRIsc trial. Participants were aged 30-55 years, had a ≥20% cumulative lifetime risk, but no *BRCA1/2* or *TP53* gene variant, and were screened outside the population-based screening program. Women received annual mammography (mammography-group), or annual MRI and biennial mammography (MRI-group). We asked whether women trust the screening outcome, what they consider as (dis)advantages, which screening they prefer and what they expect of the early detection by the screening tools.

Results: 255 (62%) women completed our questionnaire. The high chance of early cancer detection was the most important advantage of MRI screening (MRI-group: 95%; mammography-group: 74%), while this was also the main advantage of mammography (MRI-group: 57%; mammography-group: 72%). Most important disadvantages of MRI were the small tunnel and the contrast fluid (for 23-36%), and of mammography were its painfulness and X-radiation (for 48-60%). Almost the whole MRI-group and half the mammography-group preferred screening with MRI (either alone or with mammography).

Discussion: Most women would prefer screening with MRI. The way women think of MRI and mammography is influenced by the screening strategy they are undergoing. Our outcomes can be used for creating information brochures when MRI will be implemented for more women.

INTRODUCTION

Breast cancer is the most common cancer among women. Many countries offer breast cancer screening in order to detect breast cancer at an early stage. High risk women are often offered breast cancer screening with mammography and/or MRI outside population-based screening programs.[1-3] Recently, research on MRI screening efficacy is extended to other subgroups of women with increased risk. A large randomized controlled trial in women with extremely dense breast tissue in the population-based screening program showed that additional MRI led to less interval cancers.[4] Another randomized controlled trial, investigating MRI screening in women with a family history of breast cancer but without a pathogenic gene variant (the FaMRIsc trial), also showed a higher sensitivity of MRI, and cancers being detected at an earlier stage, compared to mammography.[5] Unfortunately, MRI screening also leads to more false positive screening results, which was also seen in these trials,[4,5] and it is more expensive than mammography.[6,7]

Whether women should be screened with MRI is mostly based on cohort studies, randomized controlled trials and cost-effectiveness analyses. Little is known about preferences of women themselves. However, participant acceptability is crucial for a possible implementation of MRI screening. Due to the above mentioned randomized trials, and increasing MRI expertise and technologic advances over the years, it is possible that this modality will be implemented for a greater amount of women in the future.[8]

Several studies on population-based mammography screening showed that women regard the possibility of an earlier diagnosis as more important than the risk of false-positive screening results or overdiagnosis.[9] A study by Phillips et al. investigated patient preferences and attitudes towards contrast-enhanced spectral mammography (CESM) and MRI, and found that most high-risk women in their study preferred CESM over MRI if the exams had equal sensitivities.[10] Another study also showed a preference towards CESM over MRI.[11] In contrast, a study by Essink-Bot et al. showed that women with an increased risk for breast cancer undergoing MRI screening mainly preferred MRI as a screening test over mammography when assuming equal performance.[12] They also showed that these women experienced 'lying in the tunnel', 'noise of the machine' and the fact that they were not allowed to move during the procedure as important burdens of MRI.[12] To our knowledge, no previous studies investigated what women who were randomized to either MRI or mammography screening expect and think of both tools.

In our study, we compare experiences, expectations and preferences for MRI and mammography among women with a family history of breast cancer who were either screened with mammography or with a combination of MRI and mammography during the FaMRIsc study.[5,13]

METHODS

Study population

The Familial MRI Screening (FaMRIsc) study was a multicenter randomized controlled trial assessing the efficacy of MRI screening in comparison to mammography in women with a family history of breast cancer, and assessing the influence of breast density.[13] Women aged 30-55 years with a cumulative lifetime risk for breast cancer of ≥20% due to a family history without a known *BRCA1/2* or *TP53* variant were randomly assigned to two groups: 1) the mammography-group: screening consisting of annual mammography, and 2) the MRI-group: screening consisting of annual MRI and biennial mammography. Both groups also received annual clinical breast examination (CBE). Women who did not want to be randomized but who provided consent for registration of their screening results were grouped as the registration group (231/1586=15%) and could either be screened according to the mammography-protocol (218 of 231=94%) or MRI-protocol (13 of 231=6%).[13] In this paper, all women who were screened according to the MRI-protocol are referred to as the MRI-group, and women screened according to the mammography-protocol are referred to as the mammography-group.

During the final months of the FaMRIsc trial (end of 2017), 412 of 1586 (26%) participants were sent a letter in which they were asked to complete a questionnaire.[13] The letter contained a code to log in to a website to complete the questionnaire. Participants could also request a printed version of the questionnaire. The invitation letter for the questionnaire was sent randomly to an equal number of women per screening protocol who filled in a previous questionnaire. This was done to increase the likelihood of reaching a high response rate. We aimed to also include all women with a breast cancer diagnosis, so we also invited participants who were diagnosed with an invasive cancer during the trial. In the letter we highlighted that we were not testing their knowledge but that we were interested in their opinion.

Questionnaire

The questionnaire was developed by three researchers of the FaMRIsc study (of which one was also a clinician) and was discussed in a group of five other breast cancer screening researchers. The questionnaire was sent to women in November 2017. In January 2018, a reminder was sent to all women who did not respond to the first invitation.

The questionnaire contained 17 questions and an open space to fill in the year of birth. The questions included in this paper encompassed four categories: 1) breast cancer (screening) history; 2) advantages and disadvantages; 3) expectations; 4) preference. The questionnaire can be found in the Supplementary Appendix. Questions 1-4 were related to the history of breast cancer screening and a possible breast cancer diagnosis. The category 'advantages and disadvantages' consisted of four questions (questions 11, 12,

13 and 14), all containing a list of advantages or disadvantages of MRI or mammography screening. In these questions, participants were asked to assign all options from the list that they consider important advantages or disadvantages of MRI and mammography. In case multiple answers were chosen, participants were asked to indicate which answer was most important to them. The category 'expectations' contained a question about screening in general (question 5), a question about early detection of MRI in comparison with mammography (question 10), and two questions about trust in the findings of mammography and MRI (questions 8 and 9), both on a Likert scale with a range of 0-4. The category 'preference' contained a question to obtain participants' preferences for a screening modality and questions evaluating preference with regard to the ability of early detection, the chance of false-positive results and costs (question 6, 15, 16 and 17). One question was neither about mammography or MRI (question 7), and was therefore not included in this manuscript (but the outcomes can be found in the Supplementary appendix, table S8).

Statistical analyses

Outcomes were stratified according to the screening protocols of the study: the MRI-group and the mammography-group. Furthermore, outcomes of the preference for a screening modality (question 6) were stratified by women experiencing a false alarm yes/no (question 3), and by women who had ever had undergone MRI yes/no (question 4) when participating in the mammography-group. The latter stratification was also performed for the question about trust in the findings of MRI (question 9). Fisher's exact tests were performed to examine differences in the answers between the groups. A *p*-value of 0.05 or less was considered statistically significant. Missing data were taken into account when analyzing the data, and were included in the tables, however these were not included in the Fisher's exact tests.

RESULTS

A total of 412 participants of the FaMRIsc study were sent a letter in which we asked them to fill in our questionnaire. After receiving the first letter, 178 women filled in the questionnaire, and another 77 women filled it in after the reminder was sent. This resulted in a response rate of 62% (255/412). Two women did not answer the question regarding the screening protocol they were assigned to. Therefore, we excluded the outcomes of these two women from the analyses. Of the respondents, 43% (108/253) was screened according to the MRI-protocol, and 57% (145/253) according to the mammography-protocol. Most of these women (241/253: 95% underwent randomization to these protocols, and the other 12 women participated in the registration group. Of the women

who were screened according to the mammography-protocol, 36% (49/145) previously had a breast MRI, either for diagnostics or for screening. Women in the MRI-group were on average 50 years old (SD:6.3), and women in the mammography-group 51 years old (SD:6.4). Seven respondents were diagnosed with breast cancer (MRI-group: 5; mammography-group: 2), of which six were screen-detected within the FaMRIsc study, and one was an interval cancer. Four women were diagnosed with a precursor of breast cancer (ductal carcinoma in situ) before the FaMRIsc study, and two women were undergoing additional diagnostic testing due to a positive screening result (both in the MRI-group). All outcomes of the questions regarding breast cancer (screening history) can be found in the supplementary appendix (tables S1-S4).

Advantages and disadvantages

Table 1 shows how often specific advantages and disadvantages of mammography and MRI were chosen per group. In the MRI-group more women called 'the high chance of early detection' an advantage or MRI than of mammography (95% vs. 57% respectively), while in the mammography-group the advantage of 'high chance of early detection' was given as frequently for MRI as for mammography (74% vs 72% respectively). In both groups, the high chance of early detection was the most frequently mentioned advantage for both mammography and MRI. The two groups also agreed on the most important disadvantages of mammography screening: 1) 'it is painful', 2) 'radiation risk', and 3) 'it does not detect all breast cancers'. Women who chose the option 'Other, …' for the questions about advantages and disadvantages of MRI, mostly wrote that they never had a breast MRI and therefore did not know what to answer. The disadvantage 'it causes a false alarm sometimes' was not frequently chosen, neither for MRI nor for mammography (percentages ranging from 5% to 20%).

When asking the participants to indicate which advantage and disadvantage of mammography screening was most important for them, most women ranked the chance of early detection by mammography as the most important advantage (mammography-group: 65%(67/103); MRI-group: 53% (39/74)), and the fact that mammography can be painful as the most important disadvantage (mammography-group: 41% (46/111); MRI-group: 29% (23/80)). Women of the MRI-group ranked the disadvantages 'radiation risk' and 'it does not detect all breast cancers' also as important disadvantages of mammography, with 26% (21/80) and 28% (22/80) respectively. When it comes to advantages of MRI, both groups ranked the early detection of breast cancers as most important (mammography-group: 60% (61/101); MRI-group: 81% (69/85)). The groups also agreed on the most important disadvantage of MRI: 'you have to lie in a small tunnel' (mammography-group: 26% (24/91); MRI-group: 24% (20/85)).

Table 1. Advantages and disadvantages of mammography and MRI

Mammography screening					
Advantages of mammography	**MRI-group (N=108)**	**Mammography-group (N=145)**	**Disadvantages of mammography**	**MRI-group (N=108)**	**Mammography-group (N=145)**
High chance of early detection of breast cancer	61 (57%)	105 (72%)	It is painful	65 (60%)	83 (57%)
You can get the screening result quickly	58 (54%)	98 (68%)	You get X-radiation	52 (48%)	71 (49%)
It does not take much time	43 (40%)	67 (46%)	It does not detect all breast cancers	52 (48%)	46 (32%)
I was already familiar with mammography	30 (28%)	46 (32%)	I do not see disadvantages of mammography	16 (15%)	29 (20%)
It has a small chance of a false alarm	14 (13%)	20 (14%)	It causes a false alarm sometimes	22 (20%)	26 (18%)
I can get it close to where I live	9 (8%)	17 (12%)	Other, …	4 (4%)	6 (4%)
I do not see advantages of mammography	10 (9%)	10 (7%)	It takes long before I get the result	1 (1%)	5 (3%)
It is not expensive	8 (7%)	8 (6%)	It takes (too) much time	0	1 (1%)
Other, …	9 (8%)	7 (5%)	I have to take off my clothes	0	0
MRI screening					
Advantages of MRI	**MRI-group (N=108)**	**Mammography-group (N=145)**	**Disadvantages of MRI**	**MRI-group (N=108)**	**Mammography-group (N=145)**
High chance of early detection of breast cancer	103 (95%)	107 (74%)	You have to lie in a small tunnel	39 (36%)	47 (32%)
You don't get X-radiation	43 (40%)	60 (41%)	The infusion of contrast fluid is unpleasant	35 (32%)	34 (23%)
It does not cause pain	45 (42%)	55 (38%)	The noise is unpleasant	35 (32%)	31 (21%)
You can get the screening result quickly	19 (18%)	33 (23%)	It takes a lot of time	27 (25%)	22 (15%)
It has a small chance of a false alarm	27 (25%)	21 (15%)	I do not see disadvantages of MRI	26 (24%)	19 (13%)
Other, …	1 (1%)	20 (14%)	Other, …	10 (9%)	41 (28%)
I can keep some clothes on	7 (7%)	6 (4%)	I have to wait more than one day for the result	20 (19%)	14 (10%)
			Some contrast fluid may remain in my body, even though no side effects of this are known	16 (15%)	12 (8%)
			It is expensive	10 (9%)	18 (12%)
			It does not detect all breast cancers	12 (11%)	12 (8%)
			It causes a false alarm sometimes	11 (10%)	7 (5%)
			It is far from home which causes travel time	3 (3%)	2 (1%)

Percentages are calculated with the number of women as denominator. Since women were allowed to choose more than one option, the sum of all percentages is higher than 100%.

Expectations

Less than 2% of the women did not expect the chance to detect breast cancer at an early stage to be higher with screening than without screening (see appendix). Similar proportions of both groups thought that MRI has a higher chance of detecting breast cancer in an early stage than mammography (MRI-group: 84%; mammography-group: 81%). However, more women of the MRI-group thought that MRI has a much higher chance of detecting breast cancer in an early stage than women of the mammography-group (**Table 2**). The difference in expectation was statistically non-significant (p=0.098).

Table 2. The chance of detecting breast cancer early by MRI is […] than with mammography

	MRI-group (N=108)	Mammography-group (N=145)
Much smaller	0	1 (1%)
Smaller	2 (2%)	2 (1%)
Similar	15 (14%)	22 (15%)
Slightly higher	39 (36%)	69 (48%)
Much higher	52 (48%)	48 (33%)
Missing	0	3 (2%)

In total, 85% of the MRI-group and 92% of the mammography-group had quite some trust or a lot of trust in mammography screening (**Table 3**). The proportion of women with a lot of trust in mammography was relatively large in the mammography-group, compared to the MRI-group (57% versus 37%). The difference in trust in mammography was statistically significant (p=0.014) between the groups. Higher proportions of women had a lot of trust in MRI, compared to mammography. However, a similar proportion of the mammography-group had a lot of trust in MRI (61%) as they had in mammography (57%). A relatively high proportion of the MRI-group had a lot of trust in MRI compared to the mammography-group (82% versus 61%). The difference in trust in MRI was significantly different (p<0.001) between the groups. Subgroup analyses of trust in the findings of MRI, stratified by prior experience with MRI are shown in **Table S5**. A higher proportion of women who had prior experience with breast MRI had a lot of trust in MRI (38/49: 78%) than women who never had a breast MRI (45/89: 51%).

Preference

Preference for a screening strategy varied significantly per screening group (p<0.001), as shown in **Table 4**. Relatively few women of the MRI-group (6%) preferred screening with only mammography, and 31% of the mammography-group preferred this strategy. Half of the MRI-group (54 of 108) and approximately a third of the mammography-group (50 of 145) preferred screening with both MRI and mammography.

Table 3. Do you trust the finding that you do/do not have breast cancer after only mammography/MRI?

	Do you trust the finding that you do/do not have breast cancer after only mammography?		Do you trust the findings that you do/do not have breast cancer after only MRI?	
	MRI-group (N=108)	Mammography-group (N=145)	MRI-group (N=108)	Mammography-group (N=145)
No trust	4 (4%)	2 (1%)	0	0
A little trust	8 (7%)	4 (3%)	2 (2%)	2 (1%)
Neutral	4 (4%)	6 (4%)	1 (1%)	31 (21%)
Quite some trust	52 (48%)	50 (35%)	17 (16%)	13 (9%)
A lot of trust	40 (37%)	82 (57%)	88 (82%)	88 (61%)
Missing	0	1 (1%)	0	11 (8%)

Table 4. Preference of screening modality

	MRI-group (N=108)	Mammography-group (N=145)
Mammography	6 (6%)	45 (31%)
MRI	41 (38%)	26 (18%)
Mammography and MRI	54 (50%)	50 (35%)
In the national breast cancer screening program	0	5 (3%)
No preference	6 (6%)	8 (6%)
No screening at all	0	0
Other, …	1 (1%)	9 (6%)
Missing	0	2 (1%)

Subgroup analyses of preference by women who ever had a false alarm and those who had not, are shown in **Table S6**. Women within the mammography-group who ever had a false alarm, had slightly less often a preference for mammography only (27%), compared to women who never had a false alarm within the mammography-group (33%). Women within the MRI-group who ever had a false alarm had more often a preference for screening with MRI only (45%) compared to women who never had a false alarm (34%) in the MRI-group. It is important to mention that in these analyses we do not know by which screening tool (mammography or MRI) false alarms were caused in the MRI-group because the questionnaire was anonymous.

Table S7 shows the preference outcomes stratified by prior experience with MRI of women in the mammography-group. Of the women having prior experience with MRI, 18% had a preference for MRI screening and 39% preferred a combination of mammography and MRI. Of the women not having prior experience with MRI, also 18% had a preference for MRI, and 30% preferred a combination of mammography and MRI.

Answers to questions 15 and 16 showed that most women (i.e. 36-39%) in the MRI-group preferred screening with MRI regardless of how much better the early detection or the chance of getting a false-positive result are (see **Tables S10, S11**). 41% of the Mx-group preferred MRI screening in case MRI would case a false alarm as often as mammography. Only a few women (11-15%) preferred MRI screening if it causes a false alarm two or three times as often as mammography. Question 17 showed that for most women (approximately 75%), their preference for MRI or mammography seemed not influenced by its costs.

DISCUSSION

Women who were screened with both MRI and mammography had a different view on these screening tools than women who were screened with mammography only. A higher proportion of women in the MRI-group valued the advantage of the high chance of early detection of MRI important compared to women in the mammography-group. Also, more women of the MRI-group thought that MRI has a much higher chance of detecting breast cancer in an early stage than women of the mammography-group. Furthermore, women screened with MRI plus mammography were having less trust in the results of a mammogram and more trust in the results of MRI than women screened with mammography only. The preference for screening strategy differed also between the two groups: almost all women of the MRI-group preferred screening with either MRI only or a combination of MRI and mammography, whereas half of the mammography-group preferred a screening strategy with MRI. Most participants in our study understood the aim of screening very well and indicated the early diagnosis of breast cancer by mammography and MRI as the most important advantage. This is in line with previous studies in women at average breast cancer risk when choosing between mammography and no screening, showing that early diagnosis is of most importance to them.[9] A previous study, during the early days of MRI screening, with a sample of 178 high risk women all undergoing mammography and MRI, showed that 44% preferred MRI as a screening test and 14% preferred mammography when equal performance of these tests was assumed. Furthermore, they showed that 64% of the participants was completely reassured by a negative MRI test result, but only 40% for mammography.[12] In our study, trust in MRI was higher: 82% and 61% of the MRI-group and mammography-group respectively had a lot of trust in the findings of MRI. Fewer women of the MRI-group (37%) had a lot of trust in mammography compared to the mammography-group (57%).

This study has some strengths and limitations. A strength of this study is the fact that two groups of women filled in our questionnaire, so answers by women who were

screened with MRI and mammography could be compared with answers by women who were not screened with MRI. To our knowledge, previous studies focused on women who were all screened with both screening modalities. A limitation of this study is the fact that the questionnaire was not pilot tested. Especially in the Mammography-group, women indicated that they were not able to say what they thought were advantages and disadvantages of MRI, since they never had an MRI. By pilot testing the questionnaire we may have been able to prevent this by stating these questions differently. Another limitation of our study is the risk of response bias. Women with a strong opinion on MRI and mammography may have filled in the questionnaire more often than women who did not have a strong opinion. As the response rates were different between the two groups, a response bias may have been the case in our study. A third limitation is the fact that overdiagnosis was not listed in our questionnaire. However, in the literature we found that women have limited awareness of overdiagnosis.[9]

When interpreting the outcomes on respondents perception on the price of MRI, it is important to mention that costs of the MRI were provided by the FaMRIsc study. In case women were referred for further assessment after a positive MRI, insurers were billed for these costs. In case MRI becomes part of screening guidelines in the Netherlands, the costs of MRI will also be billed to insurers. Currently, in the Netherlands, people pay a deductible of at least 385 euros per year, so therefore sometimes women will have to pay for the MRI themselves when MRI is implemented.

Almost all respondents (95%) were randomly assigned to one of the two screening protocols. In case women did not want to be screened according to the MRI protocol, we assume they would have refused randomization. However, we do not know whether respondents in the Mammography-group would have accepted MRI screening eventually, and we did not ask this in our questionnaire. Therefore, we cannot tell which of the disadvantages of MRI would be reasons for non-participation. A study on reasons for declining or not completing MRI screening among women with an elevated breast cancer risk showed that the most important reason for not undergoing MRI was claustrophobia (11%).[14] In our study, we did not use the word 'claustrophobia' but the description that 'you have to lie in a small tunnel', which was the most important disadvantage of MRI in both groups.

Outcomes of our study can be used in creating information brochures for women undergoing MRI screening, or even tailoring brochures and screening invitations to prior breast MRI screening experience of the women. The outcomes of our study and the outcomes of previous studies evaluating reasons for not participating can be used to inform women, especially those who have never had a breast MRI. Our findings suggest that women's thoughts on MRI screening change after getting MRI screening. Future research is needed to evaluate the influence of preferences and perceptions on actual screening attendance outside a clinical study.

Conclusion

Our outcomes show that women in the FaMRIsc trial who received MRI screening, have a preference for this screening tool and that they have a lot of trust in the screening results of MRI. Women not undergoing MRI screening seem to be positive towards MRI screening as well but they also had considerable confidence in mammography screening. Overall, most women would accept a screening strategy with MRI as this was preferred the most. The way women think of MRI and mammography as screening tools depends on the screening strategy they are undergoing.

LITERATURE

1. NICE. Clinical Guideline 164. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. 2017; https://www.nice.org.uk/guidance/cg164/chapter/Recommendations#surveillance-and-strategies-for-early-detection-of-breast-cancer. Accessed 2018-04-15.
2. NABON. Richtlijn mammacarcinoom (Breast Cancer National Guideline). 2018; https://www.oncoline.nl/borstkanker. Accessed January 15, 2019.
3. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer Journal for Clinicians.* 2007;57(3):75-89.
4. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.
5. Saadatmand S, Geuzinge HA, Rutgers EJT, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol.* 2019;20(8):1136-1147.
6. Lee JM, McMahon PM, Kong CY, et al. Cost-effectiveness of breast MR imaging and screen-film mammography for screening BRCA1 gene mutation carriers. *Radiology.* 2010;254(3):793-800.
7. Geuzinge HA, Obdeijn IM, Rutgers EJ, et al. Cost-effectiveness of breast cancer screening with MRI in women at familial risk: the randomized FaMRIsc trial. *JAMA Oncol.* 2020;6(9):1381-1389.
8. Mann RM, Kuhl CK, Moy L. Contrast-enhanced MRI for breast cancer screening. *J Magn Reson Imaging.* 2019;50(2):377-390.
9. Mathioudakis AG, Salakari M, Pylkkanen L, et al. Systematic review on women's values and preferences concerning breast cancer screening and diagnostic services. *Psychooncology.* 2019;28(5):939-947.
10. Phillips J, Miller MM, Mehta TS, et al. Contrast-enhanced spectral mammography (CESM) versus MRI in the high-risk screening setting: patient preferences and attitudes. *Clin Imaging.* 2017;42:193-197.
11. Hobbs MM, Taylor DB, Buzynski S, Peake RE. Contrast-enhanced spectral mammography (CESM) and contrast enhanced MRI (CEMRI): Patient preferences and tolerance. *J Med Imaging Radiat Oncol.* 2015;59(3):300-305.
12. Essink-Bot ML, Rijnsburger AJ, van Dooren S, de Koning HJ, Seynaeve C. Women's acceptance of MRI in breast cancer surveillance because of a familial or genetic predisposition. *Breast.* 2006;15(5):673-676.
13. Saadatmand S, Rutgers EJ, Tollenaar RA, et al. Breast density as indicator for the use of mammography or MRI to screen women with familial risk for breast cancer (FaMRIsc): a multicentre randomized controlled trial. *BMC Cancer.* 2012;12:440.
14. Berg WA, Blume JD, Adams AM, et al. Reasons women at elevated risk of breast cancer refuse breast MR imaging screening: ACRIN 6666. *Radiology.* 2010;254(1):79-87.

SUPPLEMENTARY APPENDIX

Writing committee:

H Amarens Geuzinge (Erasmus University Medical Centre), Eveline AM Heijnsdijk (Erasmus University Medical Centre), A Inge-Marie Obdeijn (Erasmus University Medical Centre), Harry J de Koning (Erasmus University Medical Centre), Madeleine MA Tilanus-Linthorst (Erasmus University Medical Centre).

Other members:

Jan C Oosterwijk (Medical Centre Leeuwarden; University Medical Centre Groningen, the Netherlands), Ingeborg Mares-Engelberts, (Vlietland Ziekenhuis, Rotterdam, The Netherlands), Emiel JT Rutgers (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands); Sepideh Saadatmand (Erasmus University Medical Centre, Rotterdam, the Netherlands), Ritse M Mann (Radboud University Hospital, Nijmegen; The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands), Rob AEM Tollenaar (Leiden University Medical Centre, Leiden, the Netherlands) Diderick BW de Roy van Zuidewijn (Medical Centre Leeuwarden, Leeuwarden, the Netherlands), Marc BI Lobbes (Maastricht University Medical Center, Maastricht, the Netherlands), Martijne van 't Riet (Reinier de Graaf Gasthuis, Delft, the Netherlands), Maartje J Hooning (Erasmus University Medical Centre, Rotterdam, the Netherlands), Margreet GEM Ausems (University Medical Centre Utrecht, Utrecht, the Netherlands), Claudette E Loo (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands), J Wesseling (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands), Ernest JT Luiten (University Medical Centre Utrecht, Utrecht, the Netherlands), Harmien M Zonderland (Amsterdam UMC, University of Amsterdam, the Netherlands), Cees Verhoef (Erasmus University Medical Centre, Rotterdam, the Netherlands), Carolien HM van Deurzen (Erasmus University Medical Centre, Rotterdam, The Netherlands), Eva Madsen (Erasmus University Medical Centre, Rotterdam The Netherlands), J Rothbarth (Erasmus University Medical Centre, Rotterdam, The Netherlands), Linetta B Koppert (Erasmus University Medical Centre Rotterdam, The Netherlands), Cecile de Monye (Erasmus University Medical Centre, Rotterdam, The Netherlands), Mandy M van Rosmalen (Erasmus University Medical Centre, Rotterdam, The Netherlands), Jolanda Remmelzwaal (The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands), Margrethe Schlooz-Vries (Radboud University Hospital, Nijmegen, The Netherlands), Nico Karssemeijer (Radboud University Hospital, Nijmegen, The Netherlands), Roelie la Roi-Antonides (Medical Centre Leeuwarden, Leeuwarden, The Netherlands), Suzan van der Meij (Amsterdam University Medical Centre, Amsterdam, The Netherlands), Titia Lans (Amsterdam University Medical

Centre, Amsterdam, The Netherlands), Wilma E Mesker (Leiden University Medical Centre, Leiden, The Netherlands), Kristien Keymeulen (Academic Hospital, Maastricht, The Netherlands), Wouter B Veldhuis (University Medical Centre Utrecht, The Netherlands), Arjen J Witkamp (University Medical Centre Utrecht, the Netherlands), Edith van Druten (Reinier de Graaf Gasthuis, Delft, The Netherlands), Eric Tetteroo (Amphia Ziekenhuis, Breda, The Netherlands), Carolien Contant (Maasstad ziekenhuis, Rotterdam, The Netherlands).

Author contributions:

HAG and MMAT-L developed the questionnaire and were responsible for the study design, HAG and EAMH did the data analyses, and made the first draft of the manuscript. HAG, EAMH, I-MO, HJdK, and MMAT-L were responsible for data interpretation and critical reading of the manuscript. All authors read and gave final approval of the submitted manuscript.

Questionnaire

In which year were you born?

Q1. In which screening group of the FaMRIsc-study did you participate?

- o The mammography group (annual mammography and clinical breast examination)
- o The MRI group (annual MRI and clinical breast examination, biannual mammography)
- o I did not participate in one of the above mentioned groups but I gave permission for my mammography screening results to be registered
- o I did not participate in one of the above mentioned groups but I gave permission for my MRI screening results to be registered
- o Other, …

Q2. Have you been diagnosed with breast cancer?

- o No
- o Yes, it was detected with screening as part of the FaMRIsc-study
- o Yes, this was detected during the FaMRIsc-study but outside the study setting (for example is a different hospital, or at the national breast cancer screening programme)
- o Yes, before the start of the FaMRIsc trial a pre-cancerous lesion has been detected
- o I am currently receiving additional diagnostic testing due to a suspicious screening result of the FaMRIsc-study

- o I am currently receiving additional diagnostic testing due to a suspicious finding which was detected outside the FaMRIsc-study
- o Other, …

Q3. Have you ever had a false alarm (you were referred for further diagnostic testing but it turned out not to be a cancer)?

- o No
- o Yes

Q4. If you were not participating in the MRI-group of the FaMRIsc-study, have you have had a breast MRI in your life?

- o No
- o Yes

Q5. I expect that by screening the chance that breast cancer will be detected early and therefore be curable:

- o Is much higher than without screening
- o Is a little higher than without screening
- o Is not higher than without screening

Q6. I prefer screening:

- o With mammography
- o With MRI
- o With both mammography and MRI
- o In the national breast cancer screening program
- o No preference
- o No screening at all
- o Other, …

Q7. After only clinical breast examination, do you trust that the findings (that you do or do not have breast cancer) are correct? Please circle the number that best represents your opinion:

0----------1----------2----------3----------4

No trust A lot of trust

Q8. After only mammography, do you trust that the findings (that you do or do not have breast cancer) are correct? Please circle the number that best represents your opinion:

0----------1----------2----------3----------4

No trust A lot of trust

Q9. After only MRI, do you trust that the findings (that you do or do not have breast cancer) are correct? Please circle the number that best represents your opinion:

0----------1----------2----------3----------4

No trust A lot of trust

Q10. I expect that the chance of detecting breast cancer early by MRI […] is than with mammography

- o Much smaller
- o Smaller
- o Similar
- o Slightly higher
- o Much higher

Q11. Advantages of mammography are for me […]

(Please provide at 'Ranking' with number 1 which advantage is of most importance for you).

	Ranking
o High chance of early detection of breast cancer	…
o It does not take much time	…
o You can get the screening result quickly	…
o It has a small chance of a false alarm	…
o I can get it close to where I live	…
o It is not expensive	…
o I was already familiar with mammography	…
o I do not see advantages of mammography	…
o Other, …	…

Q12. Disadvantages of mammography are for me […]
(Please provide at 'Ranking' with number 1 which advantage is of most importance for you).

		Ranking
o	It is painful	…
o	You get X-radiation	…
o	It does not detect all breast cancers	…
o	It sometimes causes a false alarm	…
o	It takes (too) much time	…
o	It takes long before I get the result	…
o	I have to take off my clothes	…
o	I do not see disadvantages of mammography	…
o	Other, …	…

Q13. Advantages of MRI are for me […]
(Please provide at 'Ranking' with number 1 which advantage is of most importance for you).

		Ranking
o	High chance of early detection of breast cancer	…
o	You don't get X-radiation	…
o	You can get the screening result quickly	…
o	It has a small chance of a false alarm	…
o	It does not cause pain	…
o	I can keep some clothes on	…
o	Other, …	…

Q14. Disadvantages of MRI are for me […]
(Please provide at 'Ranking' with number 1 which advantage is of most importance for you).

		Ranking
o	The infusion of contrast fluid is unpleasant	…
o	You have to lie in a small tunnel	…
o	The noise is unpleasant	…
o	It takes a lot of time	…
o	It does not detect all breast cancers	…
o	It sometimes causes a false alarm	…
o	Some contrast fluid may remain in my body, even though no side effects of this are known	…
o	It is far from home which causes travel time	…
o	I have to wait more than one day for the result	…
o	It is expensive	…
o	I do not see disadvantages of MRI	…

- o Other, … …

Q15. I would only have a preference for MRI if it, in comparison with mammography, at least … (please choose one answer):

- o Detects breast cancer at an early stage as often as mammography
- o Detects breast cancer twice as often as mammography at an early stage
- o Detects breast cancer three times as often as mammography at an early stage
- o I never prefer MRI
- o I always prefer MRI

Q16. I prefer MRI screening, even if, in comparison with mammography… (please choose one answer):

- o It causes as often a false alarm as mammography
- o It causes a false alarm twice as often as mammography
- o It causes a false alarm three times as often as mammography
- o I never prefer MRI
- o I always prefer MRI

Q17. I prefer MRI screening, even if, in comparison with mammography… (please choose one answer):

- o It has the same price
- o Also if MRI is twice as expensive
- o Also if MRI is five times as expensive
- o The price is not important to me
- o I never prefer MRI
- o I always prefer MRI

Here is room for comments about breast cancer screening that were not captures in above questions, but that you find important:

Outcomes of questions regarding breast cancer (screening) history of participants

Table S1. Answers to question 1: In which screening group of the FaMRIsc-study did you participate?

	Total group (N=255)
The mammography group (annual mammography and clinical breast examination)	136 (53.3%)
The MRI group (annual MRI and clinical breast examination, biannual mammography)	105 (41.2%)
I did not participate in one of the above mentioned groups but I gave permission for my mammography screening results to be registered	9 (3.5%)
I did not participate in one of the above mentioned groups but I gave permission for my MRI screening results to be registered	3 (1.2%)
Other, …	2 (0.8%)

Table S2. Answers to question 2: Have you been diagnosed with breast cancer?

	MRI-group (N=108)	Mammography-group (N=145)
No	97 (89.8%)	137 (94.5%)
Yes, it was detected with screening as part of the FaMRIsc-study	4 (3.7)	2 (1.4%)
Yes, this was detected during the FaMRIsc-study but outside the study setting (for example is a different hospital, or at the national breast cancer screening programme)	1 (0.9%)	0
Yes, before the start of the FaMRIsc trial a pre-cancerous lesion has been detected	1 (0.95)	3 (2.1%)
I am currently undergoing additional diagnostic testing due to a suspicious screening result of the FaMRIsc-study	0	0
I am currently undergoing additional diagnostic testing due to a suspicious finding which was detected outside the FaMRIsc-study	2 (1.9%)	0
Other, …	3 (2.8%)	3 (2.1%)

Table S3. Answers to question 3: Have you ever had a false alarm?

	MRI-group (N=108)	Mammography-group (N=145)
No	68 (63.0%)	104 (71.7%)
Yes	40 (37.0%)	41 (28.3%)

Table S4. Answers to question 4: If you were not participating in the MRI-group of the FaMRIsc-study, have you have had a breast MRI in your life?

	Mammography-group (N=145)
No	89 (61.4%)
Yes	49 (33.8%)
Missing	7 (4.8%)

Outcomes of stratifications based on screening history

Table S5. Answers to the question 'Do you trust in the findings that you do/do not have breast cancer after only MRI', by prior experience with MRI

	Mammography-group (N=145)*	
	Prior experience MRI (n=49)	**No prior experience MRI (n=89)**
No trust	0	0
A little trust	2 (4%)	0
Neutral	4 (8%)	26 (29%)
Quite some trust	4 (8%)	8 (9%)
A lot of trust	38 (78%)	45 (51%)
Missing	1 (2%)	10 (11%)

*7 women did not answer the question about prior experience with MRI

Table S6. Preference for screening modality, by screening group and by experience of a false alarm

	MRI-group (N=108)		Mammography-group (N=145)	
Preference for screening modality	**False alarm ever (n=40)**	**False alarm never (n=68)**	**False alarm ever (n=41)**	**False alarm never (n=104)**
Mammography	0	6 (9%)	11 (27%)	34 (33%)
MRI	18 (45%)	23 (34%)	9 (22%)	17 (16%)
Mammography and MRI	19 (48%)	35 (52%)	15 (37%)	35 (34%)
In the national breast cancer screening program	0	0	1 (2%)	4 (4%)
No preference	3 (8%)	3 (4%)	1 (2%)	7 (7%)
No screening at all	0	0	0	0
Other, …	0	1 (2%)	3 (7%)	6 (6%)
Missing	0	0	1(2%)	1 (1%)

p-value MRI-group: 0.29; *p*-value Mammography-group: 0.72

Table S7. Preference of screening modality, by prior experience with MRI

	Mammography-group (N=145)*	
Preference for screening modality	**Prior experience MRI (n=49)**	**No prior experience MRI (n=89)**
Mammography	14 (29%)	29 (33%)
MRI	9 (18%)	16 (18%)
Mammography and MRI	19 (39%)	27 (30%)
In the national breast cancer screening program	3 (6%)	2 (2%)
No preference	1 (2%)	7 (8%)
No screening at all	0	0
Other, …	2 (4%)	7 (8%)
Missing	1 (1%)	1 (1%)

*7 women did not answer the question about prior experience with MRI
p-value: 0.53

Outcomes of questions not shown in the manuscript

Table S8. Answers to question 5: I expect that the chance of detecting breast cancer in an early stage and therefore better curable, is […] with screening than without screening.

	MRI-group (N=108)	Mammography-group (N=145)
Much higher	88 (81.5%)	116 (80.0%)
A little higher	18 (16.7%)	27 (18.6%)
Not higher	2 (1.9%)	2 (1.4%)

Table S9. Answers to question 7: After only clinical breast examination, do you trust that the findings (that you do or do not have breast cancer) are correct?

	MRI-group (N=108)	Mammography-group (N=145)
No trust	20 (18.5%)	16 (11.0%)
A little trust	41 (38.0%)	40 (27.6%)
Neutral	14 (13.0%)	14 (9.7%)
Quite some trust	26 (24.1%)	53 (36.6%)
A lot of trust	7 (6.5%)	21 (14.5%)
Missing	0	1 (0.7%)

Table S10. Answers to question 15: I would only have a preference for MRI if it, in comparison with mammography at least […]

	MRI-group (N=108)	Mammography-group (N=145)
Detects breast cancer as often in an early stage as mammography	35 (32.4%)	38 (26.2%)
Detects breast cancer twice as often in an early stage as mammography	26 (24.1%)	54 (37.2%)
Detects breast cancer three times as often in an early stage as mammography	7 (6.5%)	17 (11.7%)
I never prefer MRI	0	9 (6.2%)
I always prefer MRI	39 (36.1%)	23 (15.9%)
Missing	1 (0.9%)	4 (2.8%)

Table S11. Answers to question 16: I prefer MRI screening, even if, in comparison with mammography…

	MRI-group (N=108)	Mammography-group (N=145)
It causes as often a false alarm as mammography	39 (36.1%)	60 (41.4%)
It causes twice as often a false alarm as mammography	8 (7.4%)	12 (8.3%)
It causes three times as often a false alarm as mammography	8 (7.4%)	4 (2.8%)
I never prefer MRI	8 (7.4%)	28 (19.3%)
I always prefer MRI	42 (38.9%)	29 (20.0%)
Missing	3 (2.8%)	12 (8.3%)

Table S12. Answers to question 17: I prefer MRI screening, even if, in comparison with mammography…

	MRI-group (N=108)	Mammography-group (N=145)
It has the same price	16 (14.8%)	23 (15.9%)
Also if MRI is two times as expensive	7 (6.5%)	5 (3.4%)
Also if MRI is five times as expensive	2 (1.9%)	1 (0.7%)
The price is not important to me	43 (39.8%)	60 (41.4%)
I never prefer MRI	5 (4.6%)	27 (18.6%)
I always prefer MRI	32 (29.6%)	20 (13.8%)
Missing	3 (2.8%)	9 (6.2%)

Chapter 8

General discussion

DISCUSSION

This thesis aimed to evaluate the effectiveness, cost-effectiveness and women's opinions regarding MRI relative to mammography for breast cancer screening in women at increased risk. Our analyses attempted to optimize screening guidelines for women with a family history of breast cancer and women with extremely dense breasts in the Netherlands, and for women carrying a pathogenic variant in ATM, CHEK2 and PALB2 genes in the US.

Part one: effectiveness of MRI screening versus mammography

Screening performance in women at familial risk

In **chapter 2** we evaluated the performance of annual MRI plus clinical breast examination (CBE) and biennial mammography (referred to as the MRI protocol) versus annual mammography plus CBE (the mammography protocol) in a large prospective randomized controlled trial. We showed that the MRI protocol resulted in a higher screen-detection rate and a higher program sensitivity (97.5%) compared with the mammography protocol (87%). The MRI protocol resulted also in more false positives and thereby a lower specificity (84%) compared with the mammography protocol (91%). Despite the higher false-positive rate, the positive predictive value (PPV) of the MRI protocol was higher (8.0%), compared with the mammography protocol (4.5%). In a previous meta-analysis on the accuracy of screening in women at familial risk without a known gene mutation, similar specificities were shown and slightly higher PPVs.[1] Instead of program sensitivities, comparative sensitivities were published and could therefore not be compared with our results.

Diagnostic performance – a role for breast density

When stratifying screening performance outcomes by breast density category (**chapter 2**), higher incidence rates were found in BI-RADS density categories C and D (respectively 11.2 and 10.8 per 1,000 screening rounds) compared with categories A and B (respectively 5.0 and 9.1 per 1,000 screening rounds). In the MRI group, breast cancer detection was higher in each breast density category compared with the mammography group. However, the numbers were small, and we did not find a significant trend in the sensitivity of the screening protocols by density categories as was previously shown in the literature.[2] However, we did show a significant trend in specificity of both screening strategies: both the MRI protocol and the mammography protocol had a higher specificity in women with lower breast density, compared to women with higher breast density.

In women with extremely dense breast tissue who participated in the DENSE trial, supplemental MRI screening resulted in an interval cancer rate of 2.5 per 1,000 screens,

and 5.0 per 1,000 screens among women who underwent mammography only. When correcting for the participation rate of 59%, the interval cancer rate of supplemental MRI screening decreased to 0.8 per 1,000 screens. The program sensitivity of MRI screening was 95.2%.[3], whereas the program sensitivity of mammography in women with extremely dense breasts is 61%.[2] Not surprisingly, also in the DENSE trial supplemental MRI screening led to an increase in false positive results.

In **chapter 4** we evaluated the performance of mammography and supplemental MRI (after a negative mammogram) in women whose breast density decreased (when ageing) from extremely dense (VDG4) to heterogeneously dense (<VDG4) and in women whose breast density remained extremely high (VDG4) from the first to the second and third screening round of the DENSE trial. We found that differences in performance of both mammography and MRI between the two groups were small and not statistically significant. We found that mammography resulted in a slightly higher screen-detection rate (3.2 versus 1.4; p=0.17), and a higher PPV (21.7 versus 8.1; p=0.09) in women whose breast density decreased to VDG3/2 compared to women whose breasts remained extremely dense. Supplemental MRI resulted in a slightly lower detection rate (4.9 versus 6.8; p=0.48) and a slightly lower PPV (17.1 versus 22.7; p=0.50) in women whose breast density decreased compared to those who remained extremely dense. In both density groups, the majority of cancers was found with MRI and not with mammography. The fact that differences in performance across the two groups were minor, could be explained by the level of volumetric breast density in women who decreased to VDG3, which was still quite high and close to the upper bound of VDG3 (7.5% ≤ volumetric breast density < 15.5%).

Prevalent versus incident screening rounds

Most of the MRI-detected tumors in the FaMRIsc trial occurred in the first round: the prevalent round. In the subsequent screening rounds – the incident screening rounds – the difference in detection between the MRI protocol and the mammography protocol became smaller. Also in the DENSE trial, a prevalence peak was shown. The screen-detection rate of MRI was 16.5 per 1,000 screens during the prevalent round, but this decreased to 5.8 per 1,000 screens during the second round.[4]

Not only detection rates differed per screening round, but also the false positive rates. In the FaMRIsc trial the false positive rate was 159 per 1,000 screening rounds (both prevalent and incident), while this rate was 124 per 1,000 screening rounds when only considering the incident rounds. The same pattern was observed in the DENSE trial: 79.8 false positives per 1,000 prevalent screens,[3] compared to 26.3 per 1,000 incident screens.[4] This decrease was probably caused by gained experience of reading MRI scans, and by the availability of prior MRI examinations, which can be used for comparison.[4]

Stage shift

The goal of MRI screening is to detect breast cancer in an earlier stage than mammography. In the FaMRIsc trial, more small tumors and more node negative tumors were detected with MRI than with mammography. When looking at the distributions of the detected tumors, we could argue that the MRI protocol resulted in a stage shift: 12% of the detected tumors in the MRI group were ≥T2, whereas 25% of the detected tumors in the mammography group were ≥T2. However, when looking at the raw numbers, it could also be concluded that more small size tumors were found in the MRI group but not less ≥T2 tumors. From that perspective, we could argue that no stage shift was found. With regards to nodal status: 4 (4/24: 17%) node positive tumors were detected in the MRI group, whereas 5 (5/8: 63%) in the mammography group. Again, this is a large difference when looking at the distributions but when looking at the raw numbers, this difference is much smaller. When considering results of the incident rounds only, the differences in raw numbers were slightly larger, with 1 versus 2 ≥T2 tumors and 2 versus 5 with a positive nodal status in the MRI group versus the mammography group. However, it should be noted that the numbers were small, and are therefore uncertain. In the MRISC study, published in 2004, a similar pattern was found: of the cancers found with MRI screening, 25% was >T2, whereas 47.5% of the tumors in the control group was >T2. Unfortunately, the authors neither stated the number of women in the control group nor detection rates within that group, so we could not argue whether MRI screening resulted in less large tumors in their study.[5] A previous study published by Warner et al. convincingly showed that MRI screening resulted in a reduction of late stage tumors in comparison with mammography in women with a BRCA1/2 pathogenic variant.[6] It could have been the case that the follow-up time of the FaMRIsc trial was too short to prove a stage shift based on the raw numbers and not only by the distributions. The median follow-up in the FaMRIsc trial was 5.2 years,[7] so cancers not detected in the mammography group during the trial, may have appeared after the trial, in a larger stage compared to those detected in the MRI group. Unfortunately we cannot prove whether this suggestion holds true. In summary, when looking at distributions only, we conclude that a stage shift was shown, but when looking at the raw data of the FaMRIsc study (and MRISC study) there is not enough evidence to say that this conclusion holds true.

One of the study end-points of the DENSE trial was also stage distribution. Unfortunately, results on this have not yet been published because the numbers of interval cancers in the second and third round of the study are still unknown.[4] However, in a recently published paper on the second screening round of the DENSE trial, it was shown that none of the MRI screen-detected cancers were late stage and none were node positive.[4] In our opinion this points towards a stage shift but more data are needed to conclude this.

Overdiagnosis

In **chapter 2**, we showed that the MRI protocol in the FaMRIsc trial resulted in higher cancer detection compared with the mammography protocol in each of the incident screening rounds. This also means that MRI screening may result in overdiagnosis. Based on trial results only, it is difficult to estimate which cancers are overdiagnosed and which cancers are not.

The proportion of overdiagnosis can be estimated by modelling the natural history of breast cancer over time, with screening, and using clinical data. This can be done by assessing the number of detected tumors in a situation with screening, which would never have been detected in a situation without screening. In both **chapter 5 and 6**, in which microsimulation modeling was used, we estimated that the absolute numbers of overdiagnosed breast cancers were higher for MRI screening strategies compared with mammography only. Of the screen-detected cancers with the mammography protocol of the FaMRIsc trial, 14% were overdiagnosed, compared with 20% of the screen-detected cancers with the MRI protocol (**chapter 5**). In **chapter 6** and based on unpublished results according to **chapter 3**, similar proportions were shown: 21% of the MRI screen-detected cancers were overdiagnosed in women with extremely dense breasts, and 21-23% of the detected cancers by MRI plus mammography in ATM, CHEK2 and PALB2 carriers (results not published in the manuscript).

Effect on breast cancer mortality

Similar to what we have stated about overdiagnosis, we cannot conclude from our short-term trial results whether MRI screening results in a decrease in breast cancer mortality. We concluded that MRI contributes to early detection of breast cancer, and previous work has shown that tumor stage at detection influences survival and cure.[8] Moreover, a previous study showed that MRI screening improves metastasis-free survival in women at familial risk at a median follow-up of nine years (hazard-ratio: 0.21, 95% CI: 0.04-0.95).[9]

In our modelling study in **chapter 3**, we showed that screening carriers of a pathogenic variant in ATM, CHEK2 and PALB2 resulted in a breast cancer mortality reduction: mammography starting at age 40 reduced breast cancer mortality by 36-39% compared to no screening. Screening with MRI at age 30 followed by mammography and MRI at age 40 reduced mortality by 55-60%. In **chapter 5** we estimated this breast cancer mortality reduction to be even larger, i.e. up to 78%, mainly due to different assumptions, which will be discussed in the paragraph 'Limitations and methodological considerations' of this chapter.

Limited benefit of mammography and CBE

In **chapter 2** (the FaMRIsc trial), we showed that three invasive cancers (13%: n=3/23) in the MRI group were detected by mammography only, and one (4%: n=1/23) by CBE only, which raises the question of the additional benefit of mammography and CBE to MRI.

Guidelines on MRI screening for increased risk groups, often recommend to perform mammography as well.[10-12] Mammography can detect calcified breast lesions, which cannot be seen with MRI. A previous Dutch study evaluating screening outcomes of mammography and MRI in women at increased risk found that mammography mostly resulted in additional detection of DCIS.[13] The authors suggested to increase the starting age of mammography to age 40 in women at increased risk, due to a limited benefit of mammography below this age. Other studies proposed the same.[14,15] This would also result in a slight decrease in the numbers of false positive and biopsies.[13] Other studies also concluded that the additional benefit of mammography to MRI is limited.[16,17] In **chapter 3** we showed that mammography before age 40 had a limited effect on breast cancer mortality, while increasing the number of false positive screens.

Guidelines are inconclusive about the use of CBE for women with a family history of breast cancer. A previous study already concluded that CBE resulted in a negligible benefit in BRCA1/2 mutation carriers who are screened with mammography and MRI.[18] Based on our results of the FaMRIsc trial, omitting CBE would have resulted in missing one cancer, and preventing false positive findings by 13% when screening with both mammography and MRI. This is in line with previous findings.[19]

Currently, no evidence is available on the additional benefit of mammography in women with extremely dense breast tissue who are screened with MRI. In the DENSE trial, women with a positive mammogram were not offered supplemental MRI so it is unknown which of the mammography-detected tumors would have been detected by MRI as well.

Balancing harms and benefits

Choosing the most intensive screening strategy most often leads to the highest numbers of breast cancer deaths averted and life years gained. Unfortunately, those strategies are most often also associated with high numbers of false positives and benign biopsies, and more overdiagnosis compared to less intensive strategies. This has already partly been discussed with regards to the additional benefit of mammography and CBE when screening with MRI, but also the starting ages of screening play a role. In **chapter 3**, incremental harm-benefit ratios were calculated to evaluate which screening strategies were efficient with regards to incremental false positives and benign biopsies per life year gained in ATM, CHEK2 and PALB2 carriers. We showed that starting with annual MRI at the early age of 25 and annual mammography plus MRI at age 40 resulted the highest ratios: additional 47.0-57.9 false positive screens and 18.0-22.2 benign biopsies

per life year gained, compared to postponing MRI to age 30. These relatively high ratios are partly caused by the lower breast cancer risk at age 25 compared to age 30, and demonstrates that starting screening at younger ages is not always beneficial (even when not considering radiation risks). Starting with MRI at age 30 and both mammography plus MRI at age 40 would be an efficient alternative with seemingly reasonable ratios (12.8-15.2 false positives and 4.9-5.9 benign biopsies per life year gained, in comparison with mammography only at age 40). Postponing the starting age of MRI to a later age (i.e. age 40), would not be efficient because this resulted in relatively more false-positive screens and benign biopsies per life year gained.

Part two: cost-effectiveness of MRI screening versus mammography

Since health care expenditure is rising in the Netherlands and in many other countries, cost-effectiveness analyses are needed to inform priority setting. In **chapter 5 and 6** we evaluated whether MRI screening is cost-effective for women at familial risk without a known BRCA1/2 or TP53 pathogenic variant, and for women with extremely dense breast tissue.

Women at familial risk

In **chapter 5** we provided evidence that MRI is cost-effective for women at familial risk with a cumulative lifetime risk of ≥ 20% without a known BRCA1/2 or TP53 pathogenic variant. We evaluated which starting ages, stopping ages and screening intervals of mammography and MRI would result in the most cost-effective screening strategy when applying a threshold of €22.000 per QALY gained. We showed that screening these women yearly with MRI only, would result in the highest breast cancer mortality reduction (-78%) of the non-dominated strategies, but would not be cost-effective due to its high costs. The most cost-effective strategy was MRI only at an interval of 18 months between age 35-60, followed by biennial mammography within the national screening program. In both the Netherlands and the UK, switching from intensified screening to the national screening program at a certain age is recommended for several high risk groups.[11,12] Due to a generally decreasing tumor growth rate and also decreasing breast density with increasing age, this is considered safe. Our results indicated that the switch to the national screening program should not take place before age 60, because switching before age 60 would lead to higher numbers of clinically diagnosed cancers and breast cancer deaths.

Screening strategies containing CBE were all dominated due to its limited benefit in the FaMRIsc trial (**chapter 2**), and the additional costs. Furthermore, we also showed that the additional effect of adding mammography to the most cost-effective strategy was limited. Screening consisting of alternating annual MRI and mammography between age 35-60, followed by screening within the national screening program until age 75,

resulted in approximately similar effects and costs as the most cost-effective strategy (MRI only every 18 months between age 35-60).

To our knowledge, only one previous study evaluated the cost-effectiveness of MRI screening for this group of women, which was based on data of the MRISC study from 1999-2006.[20] The breast cancer incidence and the sensitivity of both mammography and MRI were higher in the FaMRIsc study than in the MRISC study. Furthermore, in the previous cost-effectiveness analysis only five strategies were simulated all starting at age 35 and stopping at age 60 after which screening within the national screening program was modelled.[20]

Women with extremely dense breast tissue

In **chapter 6** we showed that MRI screening is also cost-effective in women with extremely dense breasts at average risk. We evaluated several screening strategies containing mammography and MRI, all starting at age 50 and stopping at age 74. Data from the DENSE trial were used to calibrate and update the MISCAN-breast model. Our results showed that MRI only at a four-year interval was cost-effective with the highest acceptable incremental cost-effectiveness ratio (ICER) (€15,620) when applying a willingness-to-pay threshold of €22,000. Other strategies consisting of MRI only with shorter intervals were all efficient and thereby considered good alternatives but were not cost-effective. Our threshold was based on the lower bound of the NICE threshold. When applying the upper bound of the threshold of approximately €33,000, the ICER of MRI at a three-year interval would be just above the threshold.

We modelled that women whose breast density dropped to heterogeneously dense (BI-RADS category C; VDG 3), were no longer eligible for MRI screening and would therefore continue with mammography only to age 74. However, in **chapter 4** we concluded that these women still benefit from MRI screening with regards to screen-detection. This was concluded after performing our cost-effectiveness analysis, and was therefore not included in our analysis. Since we showed that levels of volumetric breast density in women who decreased to a lower density category in a period of four years were still relatively high, and that screening performance measures were slightly but non-significantly different between these two groups, we assume that the ICERs would not be much affected by offering supplemental MRI for one or two additional screening rounds after a decrease in breast density.

To our knowledge, only one previous cost-effectiveness analysis has been performed, also based on published data of the DENSE trial, and by modelling women in the US. They showed an ICER of $8,797 per QALY gained for biennial MRI in comparison with biennial mammography.[21] They only used data from the first round of the DENSE trial and they did not model other optional screening strategies. We showed that the ICER of biennial MRI was €46,971 per QALY gained. The most important reason for this large

difference is that we applied an efficiency frontier, evaluating all relevant screening strategies. Therefore, our ICER was calculated in comparison with triennial MRI instead of biennial mammography. If we would have calculated the ICER on the comparison of biennial MRI versus biennial mammography, our ICER would have been much lower (€18,422 per QALY gained) compared to what we have published now (€46,971 per QALY gained). In our opinion, the correct approach to calculate the cost-effectiveness of screening strategies, is to evaluate many possible strategies and applying an efficiency frontier.[22]

Combining family history and breast density

In our cost-effectiveness analysis of MRI screening in women with a family history of breast cancer, we did not adjust screening strategies to breast density. The data from the FaMRIsc trial stratified by breast density were unfortunately not sufficiently large enough to use for our microsimulation model. We believe that tailoring intensified high risk screening strategies to breast density, could make MRI screening even more cost-effective.

Currently in the UK, breast density plays a role in the recommendation of MRI screening for women who have a 30% chance or higher of having a BRCA1/2 or TP53 pathogenic variant.[11] Previous research has already shown that tailoring mammography screening intervals to breast cancer risk and breast density results in a similar or better balance of harms and benefits, and is likely to be cost-effective.[23-25] In the paragraph 'Future developments' of this chapter, tailored screening will be discussed.

Part three: what do women want?

Apart from the opinions of radiologists, other clinicians, health economists and policy makers, it is also important to evaluate the opinions of women themselves regarding the screening they are recommended to undergo. In **chapter 7** we showed that women with a family history of breast cancer who were participating in the FaMRIsc trial, had more trust in the findings by MRI than the findings by mammography. We also showed that the most important disadvantages of MRI were the small tunnel (32-36%) and the contrast fluid (23-32%). Most important disadvantages of mammography were its painfulness (57-60%) and radiation (48-49%). Since half of the women in the FaMRIsc trial underwent mammography plus MRI screening and the other half only mammography (both with CBE), we stratified the outcomes by screening scheme to evaluate whether opinions were different. More women of the MRI-group (48%) thought that MRI has a much higher chance of detecting breast cancer in an early stage than women of the mammography-group (33%). In both screening groups, most women preferred screening with mammography and MRI. This preference was most apparent in the MRI group, of which half of the women preferred screening with both MRI and

mammography, whereas 35% of the mammography-group. A previous study, on women's acceptance of MRI screening in a cohort of women at increased risk showed that 44% of the women preferred MRI, 41% preferred clinical breast examination and 14% mammography in case performance of the screening tests was equal.[26]

It is of importance to mention that women with different breast cancer risk may have different views on breast cancer screening. Women with a breast cancer family history, of whom for example the mother or sister has previously been diagnosed with breast cancer, may have a different view on screening than women who do not have family members with a breast cancer history.[27,28] Therefore, results of our questionnaires in **chapter 7** cannot just be generalized to women without a breast cancer family history.

Of the women invited to be randomly assigned to one of the two screening groups in the FaMRIsc trial, 77% accepted the invitation. In the DENSE trial, this percentage was 59%, which was similar to the participation rate of an MRI screening study in the United States (the ACRIN 6666 study) inviting women with elevated breast cancer risk.[29,30] A study evaluating the willingness to undergo MRI screening in the DENSE trial, showed that most important reasons for declining the invitation were 1) MRI inconveniences and/or self-reported (27%) contra-indications such as claustrophobia and refusing contrast agent, 2) anxiety regarding the result of supplemental screening (21%), and 3) personal reasons (21%) such as other health concerns and/or low estimate of own risk.[29] Most of the reasons for not undergoing MRI screening within the ACRIN 6666 study were claustrophobia, time constraints and/or other priorities, and financial concerns.[30]

It is difficult to generalize participation rates of a clinical study to general participation rates once MRI screening would be implemented. Reasons why women may behave differently could be that it is not yet proven that MRI screening is actually beneficial for them when participating within a clinical study, and some women would like to contribute to science by participating in a clinical trial.[29]

Besides women's opinions on the screening tools, the screening interval should also be acceptable for them. Short intervals might result in low participations rates at follow up screens, whereas a long screening interval may result in opportunistic screening: screening outside the organized screening program. In **chapter 6** we concluded from an economic perspective that MRI screening at a four-year interval would be optimal for women with extremely dense breast tissue. However, women may not be willing to be screened only once every four years after hearing about their increased breast cancer risk (due their extremely dense breast tissue). To the best of our knowledge, no studies examined whether women at increased breast cancer risk would accept a longer screening interval while being screened with MRI. A study on preferences towards personalized breast cancer screening within the general breast cancer screening population showed that of women at a hypothetical low risk only 19% would accept a screening interval of 4 or 5 years (mammography screening).[31] It is difficult to translate

this to women at increased risk being screened with MRI which is more sensitive than mammography, but it could give some sense of their thoughts of a four-year interval. The same study also showed that most women with a hypothetical high risk preferred a screening interval of six months (31%) or one year (51%), and a starting age of 40 years (59%) instead of 50 years (with mammography).[31]

Apart from whether women participate in screening, it is also important that their decision to (not) participate is an informed one. Currently, women being invited for the national mammography screening program receive an information leaflet explaining the harms and benefits of mammography screening. A study evaluating whether women in the Netherlands actually make an informed choice when intending to (not) participate showed that 88% made an informed choice.[32] To also reach this high proportion for MRI screening, it is important to inform women about all benefits and harms of MRI screening and also of mammography screening, in order to enable them to make an informed decision on whether they want to be screened with MRI and/or mammography. A decision to (not) participate in screening is the right decision when a woman has all decision-relevant knowledge.

Limitations and methodological considerations

There are several limitations and methodological considerations to keep in mind when interpreting the results from this thesis.

The FaMRIsc trial

The FaMRIsc trial was set up in a period when pathogenic variants in several genes, such as ATM, PALB2 and CHEK2, were not yet associated with an increased breast cancer risk and genetic testing for these genes was not yet offered in the Netherlands. Nowadays the population participating in the FaMRIsc trial could be further subdivided into multiple risk groups. In **chapter 3** we already showed that breast cancer risk estimates were different across ATM, PALB2 and CHEK2 carriers. In our cost-effectiveness analysis in **chapter 5**, we modelled women with a breast cancer family as a whole, while it may be more cost-effective to tailor screening strategies according to pathogenic variants and associated breast cancer risk.

Even though the FaMRIsc trial was a large trial, the number of women attending was too low and/or the follow-up period was too short to detect a difference in interval tumors between the two groups. Furthermore, it was not sufficiently powered to evaluate differences in sensitivity by breast density.

Microsimulation modelling

In this thesis, three studies were performed using microsimulation modelling. Microsimulation models are a useful tool to overcome the limitations of randomized

controlled trials, but it also has some limitations. Not all parameters in the model can directly be measured and obtained from data. Therefore, assumptions need to be made. This is especially the case for the natural history parameters: once a woman is diagnosed with breast cancer she gets treatment, which makes it impossible to observe tumor progression rates in the absence of treatment. To overcome this issue, calibration is used. In each chapter using microsimulations modelling, different data was used for calibration. The quality of input data in the model directly determines the quality of the model output.

In **chapter 3**, breast cancer risk estimations from the CARRIERS consortium were obtained, using rather large amounts of data.[33] Unfortunately they did not collect screening data. Therefore, we used published estimates of MRI and mammography performance within the High Risk Ontario Breast Screening Program (OBSP).[34] Thereby, we assumed that screening performance in women carrying several types of pathogenic variants in Ontario (Canada) were generalizable to the population we were modelling: ATM, CHEK2 and PALB2 carriers in the US. In **chapter 5** we used data from the FaMRIsc trial for calibration. Even though the FaMRIsc trial is the largest trial available, the amount of data was rather small for calibration of subgroups. Since low numbers of interval cancers and screen-detected cancers by T-stage were obtained, confidence intervals around these numbers were large, implying uncertainty. Our goal was to obtain model estimates within the confidence limits of the observed data. To reach this goal we had to add additional tumor stages to the model: DCIS and T1a/T1B being only detectable with MRI. Hereby we assumed that DCIS and T1a/T1b could for some time only be detected by MRI before it could also be detected by mammography or before it could become clinically detectable. In **chapter 6** data from the DENSE trial was used for calibration. Since the DENSE trial was larger than the FaMRIsc trial, confidence intervals were smaller. Unfortunately in this calibration, not all estimated outcomes were within the confidence limits of the observed data. This was considered a weakness of **chapter 6**. The tumor stages we added to the model in **chapter 5** for the FaMRIsc trial, were not needed to improve calibration outcomes of the DENSE trial.

To evaluate the accuracy of a microsimulation model, validation is needed. Comparing calibrated outcomes with the input is defined as internal validation. Ideally, a model should be externally validated as well: i.e. comparing model output with other sources of data not used as model input. Due to the fact that we modelled quite specific risk groups, there was hardly any data available for external validation. We were only able to evaluate the external validity of breast cancer risk estimates in **chapter 3**, showing good agreement with published data[35] (results not published).

To evaluate the impact of certain parameters on model outcomes, deterministic sensitivity analyses were performed. We have done this mainly for parameters which we were most uncertain about, due to necessary assumptions. Unfortunately, such

deterministic sensitivity analyses do not take into account any correlations and non-linearities in the model, and do not incorporate how likely it is that a parameter will have a specific value.[36] Probabilistic sensitivity analysis can overcome these shortcomings. However, performing probabilistic sensitivity analysis is computationally intensive, and was therefore not performed in our studies.

Comparing model outcomes

Not only the input data differed between our modelling studies, but also some underlying assumptions in the models. Due to these differences, outputs are also different.

Comparing outcomes of **chapter 5** (MISCAN-breast, Dutch women at familial risk) with **chapter 3** (MISCAN-Fadia, model W-H, US women carrying a pathogenic variant in ATM, CHEK2 and PALB2 genes), large differences between breast cancer deaths averted are seen (and thereby also in life years gained), as shown in **Table 1**. There are several reasons for this. First, breast cancer mortality in the absence of screening is considerably lower in **chapter 3** (2.4-9.1%) than in **chapter 5** (13.6%), which is mainly caused by underlying assumptions of treatment: optimal treatment effects in **chapter 3** and actual treatment effects in **chapter 5**. With a lower breast cancer mortality in the absence of screening, screening can avert less breast cancer deaths compared to a situation with a higher mortality. Second, we conclude that the incidence in **chapter 5** is rather high, compared with **chapter 3**. In the absence of screening, the modelled breast cancer incidence was 30.6% in **chapter 5**, whereas in **chapter 3**, average modelled incidence in the absence of screening among carriers of ATM, CHEK2 and PALB2 pathogenic variants was 21% (range across models 18-24%), 28% (23-33%), and 38% (36-40%) respectively. In case of a high breast cancer incidence, a larger mortality reduction can be expected when screening is implemented. A possible explanation for the relatively high risk of our modelled population in **chapter 5** relative to **chapter 3,** could be the fact that the overall age-adjusted breast cancer incidence is higher in the Netherlands (151 per 100,000 women)[37] compared to the US (126 per 100,000 women).[38] However, we do not know whether this difference is also seen in women at increased risk. Unfortunately due to a lack of data on breast cancer mortality in women at familial risk and carriers of ATM, CHEK2 and PALB2 pathogenic variants, we are not able to externally validate our findings on this.

The fact that outcomes can differ substantially between different models, highlights the importance of comparative modelling such as we did in **chapter 3**. We showed that modelling the same population with two models, could also result in different outcomes. For example, with MISCAN-Fadia lifetime breast cancer incidence among CHEK2 carriers was estimated to be 33%, whereas model W-H estimated this to be 23%. Estimations on breast cancer deaths also differed substantially between the models: lifetime breast cancer mortality in the absence of screening in CHEK2 carriers ranged from 3.1% (Model W-H) to 6.1% (MISCAN-Fadia).

Table 1. Most important differences between chapter 3 (carriers of ATM, CHEK2, and PALB2 pathogenic variants) and chapter 5 (FaMRIsc trial)

	Chapter 3	**Chapter 5**
Population	ATM, CHEK2, PALB carriers	Family history (FaMRIsc trial)
Microsimulation models used	Model W-H, MISCAN-Fadia	MISCAN-breast
Country	United States	Netherlands
Treatment assumptions	Optimal treatment effects	Actual treatment effects
Life-time breast cancer risk in the absence of screening	ATM: 21% (18-24%) CHEK2: 28% (23-33%) PALB2: 38% (36-40%)	31%
Life-time breast cancer mortality in the absence of screening	ATM: 3.4% (2.4-4.5%) CHEK2: 4.6% (3.1-6.1%) PALB2: 7.7% (6.4-9.1%)	14%
Breast cancer mortality reduction mammography relative to no screening (%)	Annual mammography 40-74: ATM: 38.5% (37.8-39.2%) CHEK2: 38.4% (38.0-38.8%) PALB2: 36.4% (34.6-38.2%)	Annual mammography 40-60, biennial mammography 60-74: 61%
Breast cancer mortality reduction MRI relative to no screening (%)	Annual mammography 40-74, annual MRI 35-74: ATM: 57.6% (57.2-58.0%) CHEK2: 57.0% (56.3-57.7%) PALB2: 54.4% (54.2-54.7%)	Annual MRI + CBE, and biennial mammography 35-60, biennial mammography 60-74: 78%

Future developments

Several developments within the field of radiology and breast cancer screening are likely to change current screening policies. The most important developments will be discussed in this paragraph.

Tomosynthesis

Digital breast tomosynthesis (DBT) has been proposed to replace digital mammography. By DBT a 3D image of the breast is made, whereas mammography takes 2D images. A meta-analysis of several prospective studies showed that DBT increases the cancer detection rate (difference of 2.4 cancers per 1,000 screens) compared with mammography.[39] A pooled reduction of the recall rate by 2.9% was shown based on US studies, whereas in European trials recall rates of DBT were similar to mammography.[39] In women with extremely dense breasts, DBT does not seem to result in a higher screen-detection rate compared with mammography.[40] A disadvantage of DBT is its higher radiation dose compared with mammography.[41] Furthermore, two large trials did not demonstrate a significant decrease in interval cancers.[42,43]

Most countries with national breast cancer screening programs still offer mammography screening. In the Netherlands where screening for women at increased risk is performed within hospitals, mammography is increasingly being replaced by DBT.

Contrast enhanced spectral mammography

Another development in the field of breast imaging is contrast enhanced spectral mammography (CESM): mammography with the use of iodinated contrast material and dual-energy energy exposure. With CESM, paired low-energy and high-energy images are obtained. The low-energy images are diagnostically similar to digital mammography.[44] Combining the low- and high-energy images highlights tumor neovascularity, similar to MRI. Most studies on CESM are in the diagnostic setting but showing promising results with comparable performance to MRI.[45] Within the screening setting, only three studies have been performed so far.[46-48] One of these studies showed detection rates of 15.5 per 1,000 for CESM, and 8.8 per 1,000 for low-energy images.[46] A study by Sorin et al. in women with dense breast tissue showed that CESM had a relative sensitivity of 91% compared with 52% of standard digital mammography.[47] There is growing interest towards using CESM for women with dense breasts because the sensitivity is not affected by breast density. Disadvantages of CESM are the use of contrast material and the higher radiation dose compared with standard digital mammography and DBT.[45] More data is needed to evaluate whether CESM can be used as a screening tool. Recently, the Contrast Enhanced Mammography Imaging Screening Trial (CMIST) was set up to provide more evidence on the use of CESM for women with dense breasts in a screening setting in the United States.[49]

Abbreviated MRI

In 2014, an abbreviated protocol of breast MRI was introduced. Since full protocol MRI (which is currently the standard MRI protocol) has long scan acquisition time (1,024-1,440 seconds) and reading time (192-396 seconds),[50] it is quite time consuming and thereby costly. With abbreviated MRI, both the acquisition time and the reading time are reduced (to 180-264 and 42-114 seconds, respectively).[50] It is thought that abbreviated MRI maintains the diagnostic accuracy of full protocol MRI.[51] However, a systematic review critically reviewing published evidence stated that the overall quality of evidence is currently low, mainly due to incomplete or short follow-up data.[50]

Artificial intelligence

The introduction of digitizing screen-film mammography resulted in an increasing interest in computer-aided interpretation of mammograms. Also in the field of MRI, computer algorithms are developed to improve screening and diagnostic performance. Over time, several developments have taken place in artificial intelligence (AI) in breast imaging, with the use of computer-aided detection (CAD) but more specifically, deep learning convolutional neural networks (CNNs). In short, with CAD and CNN an algorithm is taught what a breast cancer looks like by providing the model many examples of images with and without breast cancer present.[52] Recent studies have shown promising

results, and it is expected that AI will change how breast cancer screening is performed in the future.[52] For example, a recent study showed that artificial intelligence can be used to predict which MRI-detected lesions are malignant or benign, to reduce false positive rates.[53] Another study concluded that mammography-detection can be improved when radiologists are supported by a CAD system.[54] Moreover, recent evidence has also shown that the second reader of mammography could be replaced by artificial intelligence resulting in similar sensitivity, slightly higher specificity (5.3% higher) and a reduced workload by 44%.[55] This is in line with another large study, suggesting that the second reader could be omitted when the decision of the AI system agrees with that of the first reader.[56] More prospective studies within the real screening setting are needed before AI will actually be implemented in breast cancer screening.[52]

Non-invasive biomarkers

Breast cancer cannot only be detected using imaging, but it could also be detected using samples of blood, urine, sweat, nipple aspirate fluid, tears and breath.[57] Cancer cells, or other tissues in response to cancer cells, often release specific markers such as proteins, nucleic acids, tumor DNA, miRNAs and extracellular vesicles. These markers have the potential to supplement current approaches for the early detection of breast cancer. Among the biomarkers, those released into the blood, breath and nipple aspirate fluid seem most promising to be used for breast cancer screening.[57] To date, biomarkers are not used in clinical practice for cancer screening, and most of the biomarker studies are still at the discovery phase. However, it is thought that they will be used for (breast cancer) screening in the future, either as a supplement to imaging or to replace imaging.

Tailored screening

Another development in the field of (breast cancer) screening is tailored screening. Since knowledge of breast cancer risk factors and ways to identify different risk groups have improved, tailored screening could become an alternative to uniform screening. This way, the balance between benefits of harms of screening could be optimized. The work published in this thesis can be seen as a step towards personalized screening: using breast density and family history to identify subgroups for (different) intensified screening to optimize screening outcomes for them.

Within the DENSE trial, screening was tailored to breast density. When using risk estimation models such as the Tyrer-Cuzick model or the Breast Cancer Surveillance Consortium (BCSC) model,[58,59] this subgroup of women could be divided in more subgroups by breast cancer risk. It was shown that within the MRI participants, the highest risk groups had higher cancer detection rates and a higher positive predictive value (internal data, not yet published).

Besides the well-known pathogenic variants in genes such as BRCA1/2, PALB2, ATM and CHEK2, screening could also be tailored to polygenic risk scores.[60,61] Polygenic risk scores are based on information of multiple single nucleotide polymorphisms (SNPs), resulting in a possible discrimination of several risk groups. SNPs are common variations in the DNA sequence.

Several studies on tailoring screening to breast cancer risk have been set up, using polygenic risk scores, breast density and age for identifying risk groups. The MyPeBS (My Personal Breast Screening) trial is a large randomized controlled trial within in six countries (Belgium, France, Israel, Italy and Spain). In this trial, women aged 40-70 are divided in four risk groups based on questionnaires and genetic analysis of a saliva sample. Each risk group receives different screening, ranging from mammography every four years (lowest risk group) to mammography plus MRI every single year (highest risk group).[62] Women with dense breasts will receive ultrasound if not being appointed to the MRI screening strategy. In the US, the WISDOM (Women Informed to Screen Depending on Measures of Risk) trial has been set up to evaluate risk-based screening with screening strategies ranging from biennial mammography to annual mammography plus MRI.[63] In the Netherlands, the PRISMA study was set up to collect data on risk distributions among the Dutch screening population using blood and saliva samples, to eventually evaluate the effect of tailed screening strategies.[64]

Evidence from those trials, long-term estimations and cost-effectiveness analyses are needed for decisions on future implementation of tailored screening. A previous study evaluating whether women are willing to undergo tailored screening showed that women were positive towards tailoring the frequency of mammography screening according to their personal genetic risk.[65] However, lengthening screening intervals for women with low-risk scores may be challenging.[31]

Final conclusions

- MRI screening results in a higher sensitivity than mammography screening. The specificity of MRI is lower than the specificity of mammography during the prevalent screening round but improves in incident screening rounds.
- The additional benefit of CBE is limited, even for high risk women, and mainly results in more false positives.
- MRI screening results in both more overdiagnosis and a larger breast cancer mortality reduction in women at increased breast cancer risk, when compared with mammography screening.
- In carriers of pathogenic variants in ATM, CHEK2 and PALB2 genes who are screened with MRI, mammography before age 40 has a limited effect on breast cancer mortality reductions, while increasing the number of false positive screens.

- Women with extremely dense breast tissue who are on an MRI screening scheme, and whose breast density decreases to a lower breast density category in a subsequent screening round, still benefit from MRI screening with regards to cancer detection.
- In women with an elevated risk due to a family history of breast cancer, MRI screening only at an interval of 18 months between age 35-60, followed by biennial mammography within the national screening program is the most cost-effective screening strategy (given the NICE threshold).
- In women with extremely dense breast tissue, MRI only at an interval of four years is the most cost-effective screening strategy (given the NICE threshold).
- Women with a breast cancer family history participating in a randomized controlled trial who were undergoing MRI plus mammography screening, more often preferred screening consisting of MRI only and MRI plus mammography than women undergoing mammography only.

Recommendations

- Future randomized controlled breast cancer screening trials should be powered to evaluate performance outcomes such as interval cancer rates and the sensitivity of a screening test across all four density categories.
- Future randomized controlled trials should be powered to evaluate whether a screening strategy results in a decrease in detection of late-stage tumors.
- In addition to our cost-effectiveness analysis on MRI screening for women with extremely dense breast tissue, it is needed to evaluate whether MRI screening would still be cost-effective when offered to women whose breast density decreased.
- To optimize cost-effectiveness analyses, an estimation is needed of future performance and unit costs of MRI screening when an abbreviated protocol and artificial intelligence or other new technologies are used.
- More research should be performed on evaluating whether women with extremely dense breast tissue would accept a longer screening interval when MRI is offered to them.

REFERENCES

1. Phi XA, Houssami N, Hooning MJ, et al. Accuracy of screening in women at familial risk of breast cancer without a known gene mutation: individual patient data meta-analysis. *European Journal of Cancer.* 2017;85:31-38.
2. Wanders JO, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat.* 2017;162(1):95-103.
3. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.
4. Veenhuizen SGA, de Lange SV, Bakker MF, et al. Supplemental Breast MRI for Women with Extremely Dense Breasts: Results of the Second Screening Round of the DENSE Trial. *Radiology.* 2021:203633.
5. Kriege M, Brekelmans CTM, Boetes C, et al. Efficacy of MRI and mammography for breast-cancer screening in women with familial or genetic predisposition. *New England Journal of Medicine.* 2004;351:427-437.
6. Warner E, Hill K, Causer P, et al. Prospective study of breast cancer incidence in women with a BRCA1 or BRCA2 mutation under surveillance with and without magnetic resonance imaging. *J Clin Oncol.* 2011;29(13):1664-1669.
7. Saadatmand S, Geuzinge HA, Rutgers EJT, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol.* 2019;20(8):1136-1147.
8. Saadatmand S, Bretveld R, Siesling S, Tilanus-Linthorst MMA. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients. *BMJ.* 2015;351:h4901.
9. Saadatmand S, Obdeijn IM, Rutgers EJ, et al. Survival benefit in women with BRCA1 mutation or familial risk in the MRI screening study (MRISC). *Int J Cancer.* 2015;137(7):1729-1738.
10. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer Journal for Clinicians.* 2007;57(3):75-89.
11. National Institute for Health and Care Excellence. *Familial breast cancer (breast cancer in the family).* 2019.
12. NABON. Richtlijn Borstkanker. 2018; https://richtlijnendatabase.nl/richtlijn/borstkanker/screening.html. Accessed March 2, 2021.
13. Vreemann S, van Zelst JCM, Schlooz-Vries M, et al. The added value of mammography in different age-groups of women with and without BRCA mutation screened with breast MRI. *Breast Cancer Res.* 2018;20(1):84.
14. Obdeijn IM, Mann RM, Loo CCE, et al. The supplemental value of mammographic screening over breast MRI alone in BRCA2 mutation carriers. *Breast Cancer Res Treat.* 2020;181(3):581-588.
15. Kuhl CK, Weigel S, Schrading S, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *Journal of Clinical Oncology.* 2010;28(9):1450-1457.
16. Riedl CC, Luft N, Bernhart C, et al. Triple-modality screening trial for familial breast cancer underlines the importance of magnetic resonance imaging and questions the role of mammography and ultrasound regardless of patient mutation status, age, and breast density. *Journal of Clinical Oncology.* 2015;33(10):1128-1135.

17. Sardanelli F, Podo F, Santoro F, et al. Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. *Invest Radiol.* 2011;46(2):94-105.
18. Warner E. Screening BRCA1 and BRCA2 Mutation Carriers for Breast Cancer. *Cancers (Basel).* 2018;10(12).
19. Kriege M, Brekelmans CT, Peterse H, et al. Tumor characteristics and detection method in the MRISC screening program for the early detection of hereditary breast cancer. *Breast Cancer Res Treat.* 2007;102(3):357-363.
20. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, et al. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst.* 2013;105(17):1314-1321.
21. Kaiser CG, Dietzel M, Vag T, Froelich MF. Cost-effectiveness of MR-mammography vs. conventional mammography in screening patients at intermediate risk of breast cancer - A model-based economic evaluation. *Eur J Radiol.* 2021;136:109355.
22. O'Mahony JF, Normand C. HIQA's CEA of Breast Screening: Pragmatic Policy Recommendations are Welcome, but ACERs Reported as ICERs are Not. *Value Health.* 2015;18(8):941-945.
23. 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. *Ann Intern Med.* 2011;155(1):10-20.
24. Vilaprinyo E, Forne C, Carles M, et al. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. *PLoS One.* 2014;9(2):e86858.
25. Trentham-Dietz A, Kerlikowske K, Stout NK, et al. Tailoring Breast Cancer Screening Intervals by Breast Density and Risk for Women Aged 50 Years or Older: Collaborative Modeling of Screening Outcomes. *Ann Intern Med.* 2016;165(10):700-712.
26. Essink-Bot ML, Rijnsburger AJ, van Dooren S, de Koning HJ, Seynaeve C. Women's acceptance of MRI in breast cancer surveillance because of a familial or genetic predisposition. *Breast.* 2006;15(5):673-676.
27. Erdogan E, Tuzcu A. Comparison of mammography behaviors, health beliefs, and fear levels of women with and without familial breast cancer history. *Women Health.* 2020;60(7):776-791.
28. Dillard AJ, Couper MP, Zikmund-Fisher BJ. Perceived risk of cancer and patient reports of participation in decisions about screening: the DECISIONS study. *Med Decis Making.* 2010;30(5 Suppl):96S-105S.
29. de Lange SV, Bakker MF, Monninkhof EM, et al. Reasons for (non)participation in supplemental population-based MRI breast screening for women with extremely dense breasts. *Clin Radiol.* 2018;73(8):759 e751-759 e759.
30. Berg WA, Blume JD, Adams AM, et al. Reasons women at elevated risk of breast cancer refuse breast MR imaging screening: ACRIN 6666. *Radiology.* 2010;254(1):79-87.
31. Rainey L, van der Waal D, Broeders MJM. Dutch women's intended participation in a risk-based breast cancer screening and prevention programme: a survey study identifying preferences, facilitators and barriers. *BMC Cancer.* 2020;20(1):965.
32. van Agt H, Fracheboud J, van der Steen A, de Koning H. Do women make an informed choice about participating in breast cancer screening? A survey among women invited for a first mammography screening examination. *Patient Educ Couns.* 2012;89(2):353-359.
33. Hu C, Hart SN, Gnanaolivu R, et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med.* 2021;384(5):440-451.

34. Chiarelli AM, Blackmore KM, Muradali D, et al. Performance Measures of Magnetic Resonance Imaging Plus Mammography in the High Risk Ontario Breast Screening Program. *J Natl Cancer Inst*. 2020;112(2):136-144.
35. Breast Cancer Association C, Dorling L, Carvalho S, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med*. 2021;384(5):428-439.
36. McCabe C, Paulden M, Awotwe I, Sutton A, Hall P. One-Way Sensitivity Analysis for Probabilistic Cost-Effectiveness Analysis: Conditional Expected Incremental Net Benefit. *Pharmacoeconomics*. 2020;38(2):135-141.
37. Netherlands Comprehensive Cancer Organisation (IKNL). NKR cijfers. 2020; www.iknl.nl/nkr-cijfers. Accessed February 18, 2021.
38. American Cancer Society - Cancer Statistics Center. Incidence and Death Rates. 2020; https://cancerstatisticscenter.cancer.org/#!/. Accessed April 21, 2021.
39. Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall. *J Natl Cancer Inst*. 2018;110(9):942-949.
40. Rafferty EA, Durand MA, Conant EF, et al. Breast Cancer Screening Using Tomosynthesis and Digital Mammography in Dense and Nondense Breasts. *JAMA*. 2016;315(16):1784-1786.
41. Gennaro G, Bernardi D, Houssami N. Radiation dose with digital breast tomosynthesis compared to digital mammography: per-view analysis. *Eur Radiol*. 2018;28(2):573-581.
42. Skaane P, Sebuodegard S, Bandos AI, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. *Breast Cancer Res Treat*. 2018;169(3):489-496.
43. Houssami N, Bernardi D, Caumo F, et al. Interval breast cancers in the'screening with tomosynthesis or standard mammography' (STORM) population-based trial. *Breast*. 2018;38:150-153.
44. Francescone MA, Jochelson MS, Dershaw DD, et al. Low energy mammogram obtained in contrast-enhanced digital mammography (CEDM) is comparable to routine full-field digital mammography (FFDM). *Eur J Radiol*. 2014;83(8):1350-1355.
45. Sogani J, Mango VL, Keating D, Sung JS, Jochelson MS. Contrast-enhanced mammography: past, present, and future. *Clin Imaging*. 2021;69:269-279.
46. Sung JS, Lebron L, Keating D, et al. Performance of Dual-Energy Contrast-enhanced Digital Mammography for Screening Women at Increased Risk of Breast Cancer. *Radiology*. 2019;293(1):81-88.
47. Sorin V, Yagil Y, Yosepovich A, et al. Contrast-Enhanced Spectral Mammography in Women With Intermediate Breast Cancer Risk and Dense Breasts. *AJR Am J Roentgenol*. 2018;211(5):W267-W274.
48. Jochelson MS, Pinker K, Dershaw DD, et al. Comparison of screening CEDM and MRI for women at increased risk for breast cancer: A pilot study. *Eur J Radiol*. 2017;97:37-43.
49. American College of Radiology. Contrast Enhanced Mammography Imaging Screening Trial (CMIST). 2019; https://www.acr.org/Research/Clinical-Research/CMIST. Accessed April 8, 2021.
50. Geach R, Jones LI, Harding SA, et al. The potential utility of abbreviated breast MRI (FAST MRI) as a tool for breast cancer screening: a systematic review and meta-analysis. *Clin Radiol*. 2021;76(2):154 e111-154 e122.
51. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers RD, Bieling HB. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection-a novel approach to breast cancer screening with MRI. *J Clin Oncol*. 2014;32(22):2304-2310.

52. Sechopoulos I, Teuwen J, Mann R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Semin Cancer Biol.* 2021;72:214-225.
53. Verburg E, van Gils CH, Bakker MF, et al. Computer-Aided Diagnosis in Multiparametric Magnetic Resonance Imaging Screening of Women With Extremely Dense Breasts to Reduce False-Positive Diagnoses. *Invest Radiol.* 2020;55(7):438-444.
54. Rodriguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. *Radiology.* 2019;290(2):305-314.
55. Rodríguez-Ruiz A, Lång K, Gubern-Merida A, et al. Can AI serve as an independent second reader of mammograms? a simulation study. Vol 11513: SPIE; 2020.
56. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. *Nature.* 2020;577(7788):89-94.
57. Li J, Guan X, Fan Z, et al. Non-Invasive Biomarkers for Early Detection of Breast Cancer. *Cancers (Basel)*. 2020;12(10).
58. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med.* 2004;23(7):1111-1130.
59. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast Density and Benign Breast Disease: Risk Assessment to Identify Women at High Risk of Breast Cancer. *J Clin Oncol.* 2015;33(28):3137-3143.
60. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med.* 2008;358(26):2796-2803.
61. Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and Benefit-to-Harm Ratio of Risk-Stratified Screening for Breast Cancer: A Life-Table Model. *JAMA Oncol.* 2018;4(11):1504-1510.
62. MyPeBS. https://www.mypebs.eu/.
63. Esserman LJ, Study W, Athena I. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer.* 2017;3:34.
64. PRISMA. https://www.prisma-studie.nl/.
65. Meisel SF, Pashayan N, Rahman B, et al. Adjusting the frequency of mammography screening on the basis of genetic risk: Attitudes among women in the UK. *Breast.* 2015;24(3):237-241.

Summary (EN)

Samenvatting (NL)

SUMMARY

Breast cancer is the most common cancer among women. In the Netherlands, women aged 50 to 74 years are invited for biennial mammography screening within a national screening program. Women at high risk of developing breast cancer due to a BRCA1/2 or TP53 pathogenic variant are invited for MRI and mammography screening outside the national screening program. Apart from these high-risk women, more groups of women may benefit from MRI screening. For example, women with extremely dense breast tissue: in those women, performance of mammography is generally poor and they have a higher breast cancer risk than women with a lower breast density. Also, women with a pathogenic variant in ATM, CHEK2 and PALB2 genes, and women with a breast cancer family history without a known a pathogenic variant in BRCA1/2 or TP53 genes may benefit from MRI screening. The aim of this thesis was to evaluate whether MRI screening would be an acceptable screening modality for women at increased breast cancer risk. Risk groups included in this thesis were women with a breast cancer family history without a known causative gene mutation, women with a pathogenic variant in ATM, PALB2 or CHEK2 genes, and women with extremely dense breast tissue.

Part I: Effectiveness of MRI screening versus mammography

In the FaMRIsc trial, described in **chapter 2**, women with a cumulative breast cancer risk of at least 20% due to a breast cancer family history, but without a pathogenic variant in BRCA1/2 or TP53 genes, were randomly assigned to two screening groups. Women in the MRI group (n=674) received annual MRI and clinical breast examination (CBE) plus biennial mammography. Women in the mammography group (n=680) received annual mammography and CBE. In the MRI group, 40 cancers were detected, whereas 15 cancers were detected in the mammography group. Tumours in the MRI group were significantly smaller and more often node negative. The MRI protocol had a sensitivity of 97.5% with only one interval cancer. The mammography protocol had a sensitivity of 86.7% with two interval cancers. The MRI protocol was associated with more false positive screening results and thereby a lower specificity (83.8%) compared with the mammography protocol (91.0%).

In **chapter 3** we evaluated long-term outcomes of several screening strategies with mammography and MRI for women with a pathogenic variant in ATM, CHEK2 and PALB2 genes living in the United States. Two microsimulation models were used which we adjusted using data on breast cancer risk across the three pathogenic variants, and data from a high-risk screening program in Ontario, Canada. The models projected lifetime breast cancer incidence of 21% (18-24%), 28% (23-33%), and 38% (36-40%) among women with ATM, CHEK2 and PALB2 pathogenic variants respectively, in the absence of screening. Annual mammography only at age 40 until 74, resulted in a mortality

reduction of 36-39% and 291-621 life years gained (per 1,000 women) compared with no screening, but also in 2,092-2,224 false-positive screens and 279-296 benign biopsies. Starting with annual MRI only at age 30, and adding mammography from age 40 resulted in a mortality reduction of 55-60% and 501-1,025 life years gained (per 1,000 women) compared with no screening, and 5,075-5,415 false-positive screens and 1,439-1,528 benign biopsies. Starting mammography earlier than age 40 while already screening with MRI, increased the numbers of false positives and benign biopsies while adding little benefit.

In the next chapter, **chapter 4**, we focused on breast density. We used data of all three screening rounds of the DENSE trial in which women with extremely dense breast tissue were offered supplemental MRI screening. We evaluated whether women with extremely dense breast tissue who are on an MRI screening scheme, should still receive supplemental MRI screening when their breast density decreases to a lower density category. Outcomes were cancer detection rates, recall rates, false positive rates and positive predictive value (PPV) of mammography and supplemental MRI screening, by Volpara density grade (VDG). In women whose breast density remained extremely high (VDG4), mammography resulted in a slightly lower cancer detection rate (1.4 versus 3.2; p=0.17) and a lower PPV (8.1 versus 21.7; p=0.09) compared to women whose breast density decreased to a lower density category. Supplemental MRI screening resulted in a slightly higher detection rate (6.8 versus 4.9; 0.48) and a slightly higher PPV (22.7 versus 17.1; p=0.50) in women whose breasts remained extremely dense compared to those whose density decreased. MRI resulted in a higher detection rate than mammography in both density groups. Despite an increase in detection rate at mammography with decreasing density, MRI was still of beneficial value in detecting breast cancer.

Part II: Cost-effectiveness of MRI screening versus mammography

Chapter 5 addresses the cost-effectiveness over several screening strategies containing mammography and MRI for women with a cumulative breast cancer risk of at least 20% due to a breast cancer family history, but without a pathogenic variant in BRCA1/2 or TP53 genes. Using the MISCAN-breast model, we showed that both screening protocols of the FaMRIsc trial were not efficient: these were dominated by other screening strategies. This was mainly caused by the addition of CBE to the strategies, which was already proven to be inefficient in the FaMRIsc trial and beyond. When applying a threshold of €22,000, screening with MRI only every 18 months between the age of 35 and 60 years and subsequent screening in the national screening program until age 75 years was the most cost-effective strategy. Screening consisting of alternating annual MRI and mammography between age 35 and 60 years, followed by screening within the national screening program until the age of 75 years was almost on the frontier, with similar effects and more costs as the previously mentioned strategy. Our results

also indicated that the switch to the national screening program should not take place before the age of 60 years.

In **chapter 6**, the cost-effectiveness of MRI screening for women with extremely dense breast tissue was evaluated, using the MISCAN-breast model. Several screening strategies containing mammography and MRI between age 50 to 74 were simulated. Biennial mammography alone, which is the current screening strategy for these women in the Netherlands, resulted in 69 screen-detected breast cancers and 43 breast cancer deaths per 1000 women. Adding MRI every screening round resulted in 28 additional screen-detected cancers and 8 fewer breast cancer deaths. Screening every two years with mammography alone resulted in the lowest total costs and the lowest number of QALYs compared to all other screening strategies. All strategies on the efficiency frontier, i.e. the strategies considered efficient, did not contain mammography but only MRI. When applying the NICE threshold, quadrennial MRI screening had the highest acceptable ICER: €15,620 per QALY gained. When applying a higher threshold, triennial MRI is considered cost-effective too, with an ICER of €37,181 per QALY gained. In case a two-year interval is preferred, alternating mammography and MRI can be an alternative, although this strategy is not considered efficient. The ICERs were most sensitive to the price of MRI screening.

Part III: What do women want?

In **chapter 7** we assessed expectations, preferences and trust in mammography and MRI by women themselves. A 17-item questionnaire was sent to 412 high-risk women participating in de FaMRIsc study. Of these women, 62% (n=255/412) completed the questionnaire. The questions were grouped in four categories: 1) breast cancer (screening) history; 2) advantages and disadvantages; 3) expectations; and 4) preferences. Outcomes were stratified by screening protocol: the MRI group (43%; n=108/253) and the mammography-group (57%; n=145/253). Most important advantages of MRI were the high chance of early detection of breast cancers, the fact that you don't get X-radiation, and that it does not cause pain. Most important disadvantages of MRI were the fact that you have to lie in a small tunnel, the infusion of contrast fluid and the loud noise. More women in the MRI group (48%; n=52/108) thought that MRI screening has a much higher chance of detecting breast cancer in an early stage than mammography, compared to women in the mammography-group (33%; n=48/145). Also, more women of the MRI group had trust in the findings of a positive MRI (98%; n=105/108), compared to women of the mammography group (70%; n=101/145). Almost all women of the MRI group preferred screening with either MRI only or a combination of mammography and MRI, whereas half of the mammography-group preferred a screening strategy with MRI.

Conclusions

- MRI screening results in a higher sensitivity than mammography screening. The specificity of MRI is lower than the specificity of mammography during the prevalent screening round but improves in incident screening rounds.
- The additional benefit of CBE is limited, even for high risk women, and mainly results in more false positives.
- MRI screening results in both more overdiagnosis and a larger breast cancer mortality reduction in women at increased breast cancer risk, when compared with mammography screening.
- In carriers of pathogenic variants in ATM, CHEK2 and PALB2 genes who are screened with MRI, mammography before age 40 has a limited effect on breast cancer mortality reductions, while increasing the number of false positive screens.
- Women with extremely dense breast tissue who are on an MRI screening scheme, and whose breast density decreases to a lower breast density category in a subsequent screening round, still benefit from MRI screening with regards to cancer detection.
- In women with an elevated risk due to a family history of breast cancer, MRI screening only at an interval of 18 months between age 35-60, followed by biennial mammography within the national screening program is the most cost-effective screening strategy (given the NICE threshold).
- In women with extremely dense breast tissue, MRI only at an interval of four years is the most cost-effective screening strategy (given the NICE threshold).
- Women with a breast cancer family history participating in a randomized controlled trial who were undergoing MRI plus mammography screening, more often preferred screening consisting of MRI only and MRI plus mammography than women undergoing mammography only.

Recommendations

- Future randomized controlled breast cancer screening trials should be powered to evaluate performance outcomes such as interval cancer rates and the sensitivity of a screening test across all four density categories.
- Future randomized controlled trials should be powered to evaluate whether a screening strategy results in a decrease in detection of late-stage tumours.
- In addition to our cost-effectiveness analysis on MRI screening for women with extremely dense breast tissue, it is needed to evaluate whether MRI screening would still be cost-effective when offered to women whose breast density decreased.
- To optimize cost-effectiveness analyses, an estimation is needed of future performance and unit costs of MRI screening when an abbreviated protocol and artificial intelligence or other new technologies are used.

- More research should be performed on evaluating whether women with extremely dense breast tissue would accept a longer screening interval when MRI is offered to them.

SAMENVATTING

Borstkanker is de meest voorkomende kanker bij vrouwen. In Nederland worden vrouwen tussen de 50 en 74 jaar elke twee jaar uitgenodigd voor borstkankerscreening, bestaande uit digitale mammografie. Vrouwen met een verhoogd risico op borstkanker vanwege een BRCA1/2 of TP53 mutatie worden uitgenodigd voor MRI- en mammografiescreening buiten het bevolkingsonderzoek om. Naast deze vrouwen, zijn er nog meer groepen met een verhoogd risico op het ontwikkelen van borstkanker te onderscheiden, die nu door middel van mammografie gescreend worden en mogelijk baat kunnen hebben bij MRI screening. Een voorbeeld hiervan zijn vrouwen met zeer dicht klierweefsel. Bij deze vrouwen werkt mammografie namelijk minder goed en ze hebben een hoger borstkankerrisico dan vrouwen met een lagere borstdensiteit. Ook vrouwen waarbij borstkanker in de familie voorkomt maar die geen BRCA1/2 of TP53 mutatie hebben en vrouwen die een ATM, CHEK2 of PALB2 mutatie hebben, zouden misschien baat kunnen hebben bij MRI screening. Het doel van dit proefschrift was te evalueren of MRI een acceptabele screeningsmethode is voor vrouwen waarbij borstkanker in hun familie voorkomt, voor vrouwen met een ATM, CHEK2 en PALB2 mutatie, en voor vrouwen met zeer dicht klierweefsel.

Deel I: Effectiviteit van MRI screening in vergelijking met mammografie

In **hoofdstuk 2** hebben we gekeken naar de uitkomsten van de FaMRIsc studie. In de FaMRIsc studie werden vrouwen met een cumulatief risico op borstkanker van minstens 20% vanwege een familiegeschiedenis met borstkanker aselect verdeeld over twee screeningsgroepen. Vrouwen in de MRI groep (aantal: 674) kregen screening bestaande uit jaarlijkse MRI en palpatie en om het jaar mammografie. Vrouwen in de mammografiegroep (aantal: 680) kregen jaarlijks mammografie en palpatie. In de MRI groep werden 40 tumoren gedetecteerd en in de mammografiegroep werden 15 tumoren gedetecteerd. De tumoren in de MRI groep waren gemiddeld kleiner en minder vaak lymfeklierpositief. Het screeningsprotocol in de MRI groep had een sensitiviteit van 97,5% met één intervaltumor. Het screeningsprotocol in de mammografiegroep had een sensitiviteit van 86,7% met twee intervaltumoren. Het MRI protocol resulteerde in meer fout-positieve screeningsuitslagen en daardoor in een lagere specificiteit (83,8%) dan het mammografieprotocol (91,0%).

In **hoofdstuk 3** hebben we de langetermijneffecten van verschillende screeningsstrategieën bestaande uit mammografie en MRI voor vrouwen met een ATM, CHEK2 of PALB2 mutatie in de Verenigde Staten geëvalueerd. Hierbij hebben we gebruik gemaakt van twee microsimulatiemodellen die zijn aangepast op basis van data van het borstkankerrisico bij deze drie mutaties en op basis van screeningsdata van een screeningsprogramma voor vrouwen met een hoog risico op borstkanker. Onze modellen

schatten het risico op het krijgen van borstkanker gedurende de hele levensloop in een situatie zonder screening als volgt: ATM: 21% (18-24%), CHEK2: 28% (23-33%), en PALB2: 38% (36-40%). Per 1000 vrouwen resulteerde jaarlijkse mammografie vanaf leeftijd 40 tot 74 jaar in een mortaliteitsreductie van 36-39% en 291-621 gewonnen levensjaren in vergelijking met geen screening en 2.092-2.224 fout-positieve uitslagen en 279-296 benigne biopten. Wanneer deze vrouwen al op leeftijd van 30 jaar jaarlijks MRI zouden ondergaan en vanaf leeftijd 40 jaarlijks mammografie en MRI, dan resulteert dit in een mortaliteitsreductie van 55-60% en 501-1.025 gewonnen levensjaren in vergelijking met geen screening en 5.075-5.415 fout-positieve uitslagen en 1.439-1.528 benigne biopten. We hebben ook laten zien dat mammografie bij vrouwen jonger dan 40 jaar, die al gescreend worden met MRI, voornamelijk resulteert in een toename van het aantal fout-positieve uitslagen en benigne biopten en daarnaast weinig voordelen oplevert.

In het volgende hoofdstuk (**hoofdstuk 4**) hebben we gefocust op densiteit van het borstklierweefsel. Hierbij zijn data gebruikt van de drie screeningsrondes van de DENSE studie waarin vrouwen met zeer dicht klierweefsel (Volpara density grade (VDG) 4) additionele MRI screening aangeboden kregen na een negatief mammogram. We hebben geëvalueerd of vrouwen met zeer dicht klierweefsel die MRI screening ondergaan, nog steeds baat hebben bij additionele MRI screening wanneer de densiteit van het klierweefsel is afgenomen. We hebben hierbij gekeken naar detectie, doorverwijzingen, fout-positieven en de positief voorspellende waarde van zowel mammografie als MRI, uitsplitst naar vrouwen waarbij de densiteit in VDG4 bleef en vrouwen waarbij de densiteit afnam naar VDG3/2. In vrouwen waarbij het klierweefsel zeer dicht bleef (VDG4), resulteerde mammografie in een iets lagere screen-detectie (1,4 versus 3,2; p=0,17) en een lagere positief voorspellende waarde (8,1 versus 21,7; p=0,09) dan in vrouwen waarbij de densiteit afnam. Additionele MRI (na een negatief mammogram) resulteerde in een iets hogere screen-detectie (6,8 versus 4,9; p=0,48) en een iets hogere positief voorspellende waarde (22,7 versus 17,1; p=0,50) in vrouwen waarbij het klierweefsel zeer dicht bleef in vergelijking met vrouwen waarbij de densiteit afnam. In beide densiteitsgroepen resulteerde MRI in een hogere screen-detectie dan mammografie. Ondanks dat de detectie door middel van mammografie licht toenam bij een lagere densiteit, concludeerden wij dat additionele MRI van toegevoegde waarde was in het detecteren van borstkanker voor de groep vrouwen waarbij de densiteit afnam.

Deel II: Kosteneffectiviteit van MRI screening in vergelijking met mammografie

In **hoofdstuk 5** hebben we gekeken naar de kosteneffectiviteit van verschillende screeningsstrategieën bestaande uit mammografie en/of MRI voor vrouwen met een cumulatief borstkankerrisico van minstens 20% vanwege een

borstkankerfamiliegeschiedenis zonder een BRCA1/2 or TP53 mutatie. We hebben hierbij gebruik gemaakt van het MISCAN-borstmodel. We hebben laten zien dat de twee screeningsstrategieën in de FaMRIsc studie niet efficiënt waren en dat ze dus gedomineerd werden door betere alternatieven. Dit werd voornamelijk veroorzaakt wegens het feit dat de twee FaMRIsc strategieën palpatie bevatten, waarvan al tijdens de FaMRIsc studie was aangetoond dat het weinig voordelen opleverde. Wanneer er een grenswaarde van 22.000 euro per gewonnen levensjaar (gecorrigeerd voor kwaliteit van leven), dan is screening bestaande uit elke 18 maanden een MRI van leeftijd 35 tot 60 met daaropvolgend screening binnen het bevolkingsonderzoek tot leeftijd 74, de meest kosteneffectieve screeningsstrategie. Screening bestaande uit jaarlijks alternerend MRI en mammografie tussen dezelfde leeftijden resulteerde in vergelijkbare kosten en effecten als de eerdergenoemde strategieën maar lag net onder de efficiëntie curve. Onze resultaten hebben ook laten zien dat de overgang naar het landelijke bevolkingsonderzoek niet voor de leeftijd van 60 jaar plaats moet vinden.

In **hoofdstuk 6** hebben we gekeken naar de kosteneffectiviteit van MRI screening voor vrouwen met zeer dicht klierweefsel. Ook in dit hoofdstuk is er gebruik gemaakt van het MISCAN-borstmodel. We hebben verschillende screeningsstrategieën gemodelleerd bestaande uit mammografie en MRI tussen de leeftijden 50 en 74 jaar. Mammografie elke twee jaar, wat de huidige screening is voor deze vrouwen in Nederland, resulteerde in 69 screen-gedetecteerde tumoren en 43 borstkankerdoden per 1000 vrouwen. Met de toevoeging van MRI hieraan, werden er 28 additionele tumoren screen-gedetecteerd en waren er 8 minder borstkankerdoden. Alle strategieën die als 'efficiënt' beschouwd werden, bestonden enkel uit MRI screening en geen mammografie. Wanneer we een grenswaarde van 22.000 euro per gewonnen voor kwaliteit gecorrigeerd levensjaar toepasten, dan was MRI met een interval van vier jaar de meest kosteneffectieve screeningsstrategie (met een ICER van €15,620 per gewonnen QALY). Wanneer een iets hogere grenswaarde werd toegepast, dan was MRI met een interval van drie jaar ook kosteneffectief (ICER: €37,181 per gewonnen QALY). Wanneer men liever wil vasthouden aan een screeningsinterval van twee jaar, dan zou alternerend mammografie en MRI een goed alternatief zijn, hoewel deze strategie niet als efficiënt werd beschouwd. De kosteneffectiviteitsratio's werden het meeste beïnvloed door de prijs van MRI screening.

Deel III: Wat willen de vrouwen zelf?

In het laatste deel van dit proefschrift (in **hoofdstuk 7**), hebben we gekeken naar de verwachtingen, voorkeuren en vertrouwen in mammografie en MRI van de vrouwen zelf. Een vragenlijst bestaande uit 17 vragen was verstuurd naar 412 vrouwen met een verhoogd risico op borstkanker die meededen aan de FaMRIsc studie. Van deze 412 vrouwen vulde 62% (n=255/412) de vragenlijst in. De vragen waren in vier categorieën ingedeeld: 1) borstkanker(screenings)geschiedenis; 2) voordelen en

nadelen; 3) verwachtingen; en 4) voorkeur. De uitkomsten werden uitsplitst naar de twee screeningsstrategieën van de FaMRIsc studie die de vrouwen ondergingen. De meest belangrijke voordelen van MRI waren de hoge kans op vroege detectie van borstkanker, het feit dat er geen gebruik wordt gemaakt röntgenstraling en dat het geen pijn doet. De meest belangrijke nadelen van MRI waren het feit dat je in een smalle buis moet liggen, dat er contrastvloeistof nodig is en dat het veel lawaai maakt. Meer vrouwen in de MRI groep (48%; n=52/108) hadden het idee dat MRI een veel hogere kans heeft op het detecteren van borstkanker in een vroeg stadium dan mammografie, dan vrouwen in de mammografiegroep (33%; n=48/145). Daarnaast hadden ook meer vrouwen in de MRI groep (98%; n=105/108), vertrouwen in de bevindingen van MRI bij een positieve uitslag dan vrouwen in de mammografiegroep (70%; n=101/145). Bijna alle vrouwen in de MRI groep hadden voorkeur voor screening bestaande uit enkel MRI of een combinatie van mammografie en MRI, terwijl maar de helft van de vrouwen in de mammografiegroep dat prefereerde.

Conclusies

- MRI screening heeft een hogere sensitiviteit dan mammografie screening. De specificiteit van MRI is lager dan de specificiteit van mammografie tijdens prevalente screeningrondes maar dit verbetert in incidente screening rondes.
- Het additionele voordeel van palpatie is beperkt, zelfs bij vrouwen met een verhoogd risico op borstkanker. Het resulteert voornamelijk in meer fout-positieve uitslagen.
- In vergelijking met mammografie screening resulteert MRI screening in meer overdiagnose maar ook in een hogere borstkankersterftereductie in vrouwen met een verhoogd risico op borstkanker.
- Bij draagsters van een mutatie in het ATM, CHEK2 of PALB2 gen die gescreend worden door middel van MRI, heeft mammografie voor de leeftijd van 40 jaar een beperkt effect op de borstkankersterftereductie en zorgt het voor een toename in het aantal fout-positieve uitslagen.
- Vrouwen met zeer dicht klierweefsel die door middel van MRI gescreend worden en wiens borstdensiteit afneemt naar een lagere categorie afneemt, hebben ook dan nog steeds voordeel bij MRI screening wat betreft de borstkankerdetectie.
- Voor vrouwen met een verhoogd borstkankerrisico vanwege een borstkankerfamiliegeschiedenis is MRI screening met een interval van 18 maanden tussen de leeftijden 35 en 60 jaar, gevolgd door tweejaarlijkse mammografie binnen het landelijke screeningsprogramma, de meest kosteneffectieve screeningsstrategie (uitgaande van de NICE drempelwaarde).
- Voor vrouwen met zeer dicht klierweefsel is MRI met een vierjaarlijks interval de meest kosteneffectieve screeningsstrategie (uitgaande van de NICE drempelwaarde).

- Vrouwen met een borstkankerfamiliegeschiedenis die zowel MRI als mammografiescreening in een gerandomiseerde studie ondergingen, hadden vaker de voorkeur voor screening bestaande uit enkel MRI of MRI plus mammografie dan vrouwen die alleen mammografie screening ondergingen.

Aanbevelingen

- Toekomstige, gerandomiseerde, gecontroleerde studies zouden dusdanig gepowered moeten zijn om de uitkomsten, zoals intervalkankers en de sensitiviteit van de screeningstest, in alle densiteitscategorieën goed te kunnen evalueren.
- Toekomstige, gerandomiseerde, gecontroleerde studies zouden gepowered moeten zijn om een afname in gevorderde tumoren te kunnen evalueren.
- Ter aanvulling op onze kosteneffectiviteitsanalyse van MRI screening voor vrouwen met zeer dicht klierweefsel, is het nodig om te onderzoeken of MRI screening ook kosteneffectief is wanneer het nog steeds wordt aangeboden wanneer de densiteit is afgenomen.
- Er is een schatting nodig van de toekomstige screeningsprestaties en kosten van MRI bij gebruik van een verkort protocol en artificiële intelligentie, om zo de kosteneffectiviteitsanalyses te optimaliseren.
- Er is onderzoek nodig om inzicht te krijgen in welke mate vrouwen met zeer dicht klierweefsel een langer screeningsinterval (dan de huidige twee jaar) zouden accepteren wanneer zij MRI screening aangehouden zouden krijgen.

List of publications

PhD portfolio

Dankwoord

About the author

LIST OF PUBLICATIONS

Publications in this thesis

Saadatmand S*, **Geuzinge HA***, Rutgers EJT, Mann RM, de Roy van Zuidewijn DBW, Zonderland HM, Tollenaar RAEM, Lobbes MBI, Ausems MGEM, van 't Riet M, Hooning MJ, Mares-Engelberts I, Luiten EJT, Heijnsdijk EAM, Verhoef C, Karssemeijer N, Oosterwijk JC, Obdeijn IM, de Koning HJ, Tilanus-Linthorst MMA, on behalf of the FaMRIsc study group. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol*. 2019;20(8):1136-1147.
**Shared first authorship*

Geuzinge HA, Obdeijn IM, Rutgers EJT, Saadatmand S, Mann RM, Oosterwijk JC, Tollenaar REAM, de Roy van Zuidewijn DBW, Lobbes MBI, van 't Riet M, Hooning MJ, Ausems MGEM, Loo CE, Wesseling J, Luiten EJT, Zonderland HM, Verhoef C, Heijnsdijk EAM, Tilanus-Linthorst MMA, de Koning HJ, on behalf of the Familial MRI Screening (FaMRIsc) Study group. Cost-effectiveness of Breast Cancer Screening With Magnetic Resonance Imaging for Women at Familial Risk. *JAMA Oncol*. 2020;6(9):1381-1389

Geuzinge HA, Heijnsdijk EAM, Obdeijn IM, de Koning HJ, Tilanus-Linthorst MMA; FaMRIsc study group. Experiences, expectations and preferences regarding MRI and mammography as breast cancer screening tools in women at familial risk. *Breast*. 2021;56:1-6.

Geuzinge HA, Bakker MF, Heijnsdijk EAM, van Ravesteyn NT, Veldhuis WB, Pijnappel RM, de Lange SV, Emaus MJ, Mann RM, Monninkhof EM, de Koekkoek-Doll PK, van Gils CH, de Koning HJ, on behalf of the DENSE trial study group. Cost-Effectiveness of Magnetic Resonance Imaging Screening for Women With Extremely Dense Breast Tissue. *J Natl Cancer Inst*. 2021; 113(11):1476-1483.

Lowry KP, **Geuzinge HA**, Stout NK, Alagoz O, Hampton J, Kerlikowske K, de Koning HJ, Miglioretti DL, van Ravesteyn NT, Schechter C, Sprague BL, Tosteson ANA, Trentham-Dietz A, Weaver D, Yaffe MJ, Yeh JM, Couch FJ, Hu C, Kraft P, Polley EC, Mandelblatt JS, Kurian AW, Robson ME, on behalf of the Breast Working Group of the Cancer Intervention and Surveillance Modeling Network (CISNET), Breast Cancer Surveillance Consortium (BCSC), and the Cancer Risk Estimates Related to Susceptibility consortium (CARRIERS). Breast cancer screening strategies for women with ATM, CHEK2, and PALB2 pathogenic variants: A comparative modeling analysis. *JAMA Oncol*. 2022; online ahead of print

Geuzinge HA, Bakker MF, Mann RM, Veenhuizen SGA, Monninkhof EM, Loo CE, van Diest PJ, Lobbes MBI, Karssemeijer N, van der Zwaag J, de Koning HJ, Duvivier KM, Pijnappel RM, Veldhuis WB, van Gils CH, on behalf of the DENSE trial study group. Decreasing breast density over time and the effect on performance of mammographic and supplemental MRI screening - Results of the DENSE trial. To be submitted.

Publications not in this thesis

Sankatsing VDV, **Geuzinge HA**, Fracheboud F, van Ravesteyn NT, Heijnsdijk EAM, Kregting LM, Broeders MJM, Otten JDM, Verbeek ALM, Pijnappel RM, de Bruijn AE, de Koning HJ: Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker (LETB). Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 2004-2014. Het veertiende evaluatierapport. LETB XIV 2019.

Kregting LM, van Ravesteyn NT, Spijker W, Dierks T, Aitken CA, **Geuzinge HA**, Korfage IJ. Effects of a leaflet on Breast Cancer Screening Knowledge, Explicit Attitudes, and Implicit Associations. *Patient Educ Couns*. 2020(12):103;2499-2507

Zielonke N, **Geuzinge HA**, Heijnsdijk EAM, Heinävaara S, Senore C, Jarm K, de Koning HJ, van Ravesteyn NT, on behalf of the EU-TOPIA Consortium. Extending Age Ranges in Breast Cancer Screening in Four European Countries: Model Estimations of Harm-to-Benefit Ratios. *Cancers*. 2021;13(13):3360.

PORTFOLIO

Summary of PhD training, teaching and other relevant activities

1. PhD training	Year	ECTS
Courses		
Didactic skills in teaching, Erasmus MC	2018	0.2
Planning and Evaluation of Screening, NIHES, Erasmus MC	2018	1.4
Advanced Topics in Decision-making in Medicine, NIHES, Erasmus MC	2019	2.4
Systematic Literature Retrieval in PubMed, Erasmus MC	2019	0.3
Scientific Integrity, Erasmus MC	2019	0.3
Biomedical English Writing and Communication, MolMed, Erasmus MC	2020	2.0
Using R for Decision Modeling in Health Technology Assessment, NIHES, Erasmus MC	2020	1.1
Data analysis in Python Basic, MolMed, Erasmus MC	2020	1.5
Presentations		
Poster presentation, International Cancer Screening Network Conference (ICSN), Rotterdam, the Netherlands	2019	0.5
Oral presentation, Breast cancer screening symposium, Medical Imaging Center, UMCG, Groningen, the Netherlands	2020	1.0
Oral presentation, European Congress of Radiology, online	2021	1.0
(Inter)national conferences and meetings		
Congres 'Bevolkingsonderzoeken naar kanker – De bevolkingsonderzoeken op het spoor', Utrecht, the Netherlands	2017	0.3
Seminar 'Screening is a Choice', Antoni van Leeuwenhoek	2018	0.2
Symposium 'Model-based Decision Making: How mathematics can contribute to solutions in healthcare', Decision Modeling Center VUMC	2018	0.2
Hebon Congress, Utrecht, the Netherlands	2017, 2018	0.6
Health(y) Research Day, Rotterdam, the Netherlands	2019	0.3
International Cancer Screening Network Conference (ICSN), Rotterdam, the Netherlands	2019	0.9
Cancer Intervention and Surveillance Modeling Network (CISNET) Breast Mid-Year meeting	2020	0.4
Seminars department of Public Health, Erasmus MC	2017-2021	3.0
European Congress of Radiology, Vienna, Austria	2018, 2019, 2021	2.4

2. Teaching		
Teaching 'Hoe houden we de zorg betaalbaar?', third-year medical students, Erasmus MC	2019, 2020, 2021	1.0
Teaching 'Keuzen in de zorg', second-year medical students, Erasmus MC	2019, 2020	1.0

Supervising Community Projects, third-year medical students, Erasmus MC	2019, 2020, 2021	3.0
Tutoring first-year medical students, Erasmus MC	2020	1.5
Supervising a PhD student	2020-2021	2.5

3. Other

Writing a report for the Health Council of the Netherlands (Gezondheidsraad) and the National Institute for Public Health and the Environment (RIVM)	2020	2.0
Reviewing a manuscript for the Journal of Health Policy and Management	2020	0.4

DANKWOORD

Het zit erop! Na ongeveer vier jaar ploeteren kan ik zeggen dat ik trots ben op het feit dat ik het heb volbracht. Ik wil heel graag een aantal mensen bedanken die een bijdrage hebben geleverd aan de totstandkoming van dit proefschrift. In mijn eentje had ik dit namelijk nooit voor elkaar kunnen krijgen.

Allereerst **Harry**, ik realiseer mij dat ik ontzettend veel geluk heb gehad met het feit dat ik mee mocht werken aan zowel de FaMRIsc trial, de DENSE trial als ook de MyPeBS trial: drie grote trials met waarschijnlijk veel impact op het borstkankerscreeningsbeleid wereldwijd. Ik ben je dankbaar dat ik deze kansen heb gekregen. Toen ik bij MGZ kwam werken durfde ik weinig mijn mening te geven maar al snel kwam ik erachter dat die onzekerheid nergens voor nodig was en dat je open stond voor mijn input. Je nam me serieus en je was ook altijd wel in voor een grapje.

Eveline, dank voor de goede begeleiding de afgelopen jaren. Waar ik soms wat gestrest kon zijn, kon jij mij weer geruststellen. Je had altijd tijd voor me: ik heb nooit lang hoeven wachten op een reactie of feedback. Mede hierdoor heb ik mijn promotie zo snel kunnen afronden.
Daarnaast wil ik ook **Nicolien** bedanken. Officieel was je niet mijn begeleider maar ook aan jouw hulp heb ik heel veel gehad. Ik kon goed met je sparren over ingewikkelde MISCAN-analyses en je stond altijd voor me klaar.

Beste **Kees,** jou wil ik bedanken voor de feedback en het meedenken tijdens de laatste fase van mijn promotietraject. Ik schreef het met jouw nuttige opmerking in mijn achterhoofd: 'de discussie moet prikkelend zijn en nieuwsgierig maken naar de artikelen, maar je moet de artikelen niet nodig hebben om het te begrijpen'. Het had zo een stelling kunnen zijn!

Madeleine, dankzij u heb ik drie hele mooie artikelen in dit proefschrift. Ik begon vier jaar geleden met weinig tot geen medische kennis maar daar heeft u snel verandering in weten te brengen. U nam mij mee het ziekenhuis in zodat ik eens kon zien hoe het screenen in de praktijk ging. Dat heb ik erg gewaardeerd.
Inge-Marie, ik wil u graag bedanken voor de nuttige inzichten die u had tijdens onze overleggen. U bekeek de uitkomsten vanuit een ander perfectief dan ik en daar heb ik dan ook heel veel van geleerd.

Carla en Marije, ik ben jullie heel erg dankbaar voor de mooie kans die kreeg om de kosteneffectiviteitsanalyse van de DENSE trial te mogen doen. Ik vond de samenwerking

ontzettend prettig en vond het dus ook extra leuk dat ik nóg een onderzoek bij jullie kon doen. Dankzij jullie kritische vragen en feedback heb ik beide artikelen naar een hoger niveau kunnen brengen. Bedankt voor de gezellige meetings!

Uiteraard wil ik ook **Arry** bedanken. Jij staat voor iedereen klaar en ik heb dan ook altijd het gevoel gehad dat ik bij je terecht kon als er iets was. Ik ben je dankbaar voor de moeite die je nam om mij op te vangen en te helpen toen ik er even doorheen zat.

Valerie, vanaf het begin was je echt mijn maatje op de afdeling en ik heb heel veel aan je gehad. Ik vond veel dingen best wel spannend aan het begin, maar jij hebt mij geleerd dat dat helemaal niet nodig was. Als ik vastliep met MISCAN dan hielp je me. Onze veel te lange koffiepauzes waren misschien wat minder bevorderlijk voor de voortgang maar ook dat had ik niet willen missen! Ik vind het leuk dat we nog steeds vriendinnen zijn en dat je mijn paranimf wil zijn bij mijn promotie.

Lucie, nadat jij je eenmaal als kamergenoot bij Na2317 had gevoegd, hebben we veel gezellige momenten beleefd. De vele kroketten die we hebben gegeten, de (online) shopmomenten, maar ook gewoon af en toe even klagen over de sores van MISCAN (denk aan de errors, waar ik overigens vaker mee te maken had dan jij). Het was (en is) altijd erg vertrouwd!

Uiteraard wil ik ook **Carlijn en Sabine** bedanken: mijn 'kamergenoten' waarmee we het flexwerken vanaf dag 1 aan onze laars lapten. En natuurlijk ook **Lisanne, Rik, Steffie, Arthur, Caroline, Esther, Elleke, Erik, Clare, Dayna, Amir, Mirjam en Brechtje,** jullie hebben mijn tijd op MGZ onvergetelijk gemaakt. De lunchpauzes, vrijmibo's, frietvrijdagen en etentjes had ik niet willen missen!

Ook wil ik graag het borstkankerteam bedanken: **Valerie, Lindy, Jeroen, Sarocha, Nadine, Eveline en Nicolien**. Bedankt dat jullie er voor mij zijn geweest en voor de gezellige lunchmeetings en inhoudelijke discussies. Furthermore, I would like to thank the members of the **CISNET-breast team**, especially **Katy**, for the collaboration on our comparative modelling paper.

Verder wil ik graag de **GI-groep** bedanken. Allereerst **Iris**, dank dat ik deel uit mocht maken van dit team en dat je mij het vertrouwen gaf om de rol van begeleider op mij te nemen. Alle gezellige virtuele koffiemomentjes, cup-a-soup breaks en borrels met het team hebben de coronatijd een stuk gezelliger gemaakt. Bedankt iedereen voor het plezier en de gezelligheid!

En natuurlijk wil ik ook de gehele **screensectie** en de afdeling **MGZ** bedanken voor een mooie tijd. We hebben zo veel leuke dingen gedaan: een workshop curling, dansen met livemuziek op de SS Rotterdam, de jaarlijkse borrelboot, de kerstborrels, beachvolleyballen in Scheveningen, enzovoorts.

En ten slotte wil ik ook mijn ouders bedanken. Lieve **pap en mam**, ook al staat dit werk ver van jullie af, jullie hebben mij altijd het gevoel gegeven trots op me te zijn. Jullie hebben mij nooit in een bepaalde richting geduwd maar hebben mij altijd gestimuleerd om datgene te doen wat mij gelukkig maakt. Of het nou mbo, hbo of universiteit werd, het maakte niet uit. Daar ben ik jullie ontzettend dankbaar voor. Zonder jullie support was dit proefschrift er niet geweest.

ABOUT THE AUTHOR

Amarens Geuzinge was born on the 20th of February 1993 in Zwolle, the Netherlands. After completing her secondary education 'Atheneum' at the Meander College in Zwolle, she started studying 'Health Sciences' at the Vrije Universiteit in Amsterdam in 2013. During this Bacherlor's program, she got the opportunity to write her master thesis at the Julius Centre (UMC Utrecht) in the field of Health Technology Assessment. After obtaining her Bachelor's degree in 2016, she started the Master's program 'Health Economics, Policy and Law' at the Erasmus University in Rotterdam. During this Master's program, she studied at the University of Oslo for half a year as part of an exchange program. Again, she wrote her thesis in the field of Health Technology Assessment, for which she started health economic modelling for the first time. After graduating in 2017, Amarens started her PhD at the department of Public Health at the Erasmus University Medical Center in Rotterdam, in collaboration with the department of Surgical Oncology of the Erasmus University Medical Center and the department of Epidemiology of the Julius Centre, UMC Utrecht. Since June 2021, Amarens works as a pharmaco-economic advisor at the National Health Care Institute.