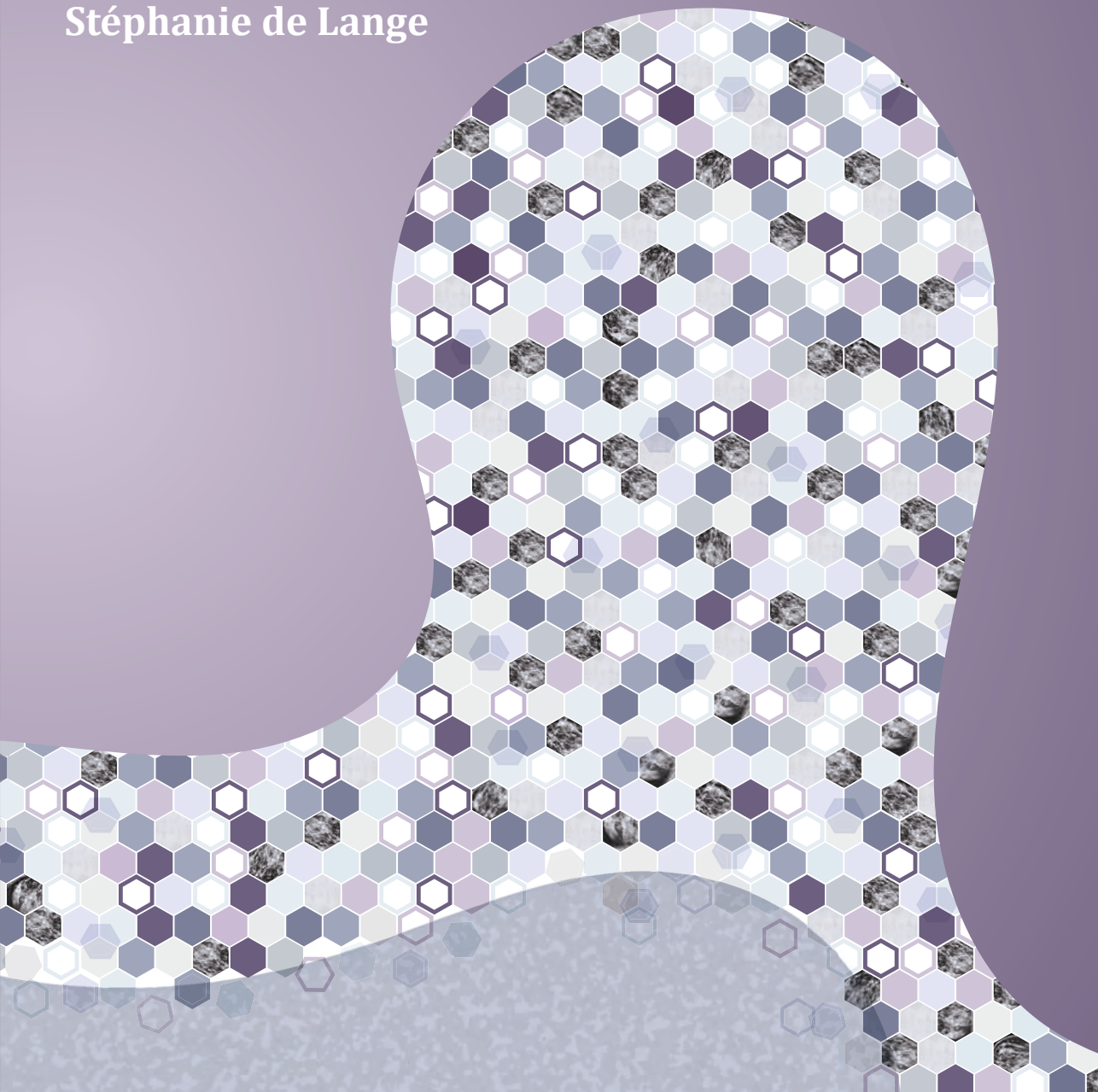
Early detection of breast cancer with MRI in women with extremely dense breasts

Stéphanie de Lange



Early detection of breast cancer with MRI in women with extremely dense breasts

Stéphanie de Lange

Early detection of breast cancer with MRI in women with extremely dense breasts

Thesis with a summary in Dutch, Utrecht University

ISBN: 978-94-6416-074-1
Cover design: Daniëlle Balk
Layout design: Daniëlle Balk | persoonlijkproefschrift.nl
Printed by: Ridderprint | www.ridderprint.nl

Early detection of breast cancer with MRI in women with extremely dense breasts

Vroege opsporing van borstkanker met MRI bij vrouwen met extreem dicht borstweefsel
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 13 oktober 2020 des middags te 12.45 uur

door

Stéphanie Vera de Lange
geboren op 30 mei 1989 te Zuidlaren

PROMOTOREN:

Prof. dr. C.H. van Gils
Prof. dr. R.M. Pijnappel

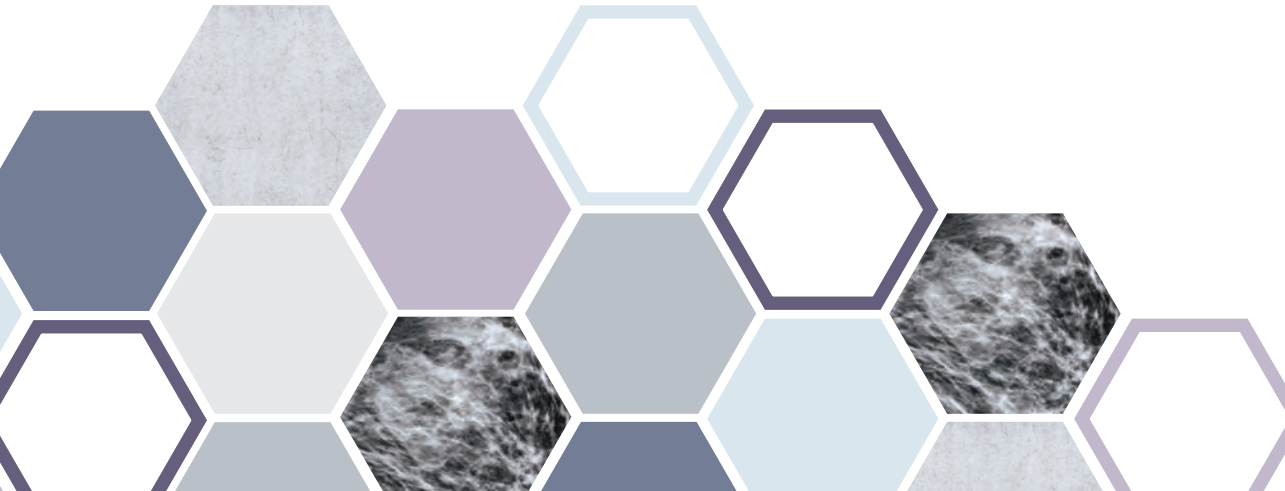
COPROMOTOREN:

Dr. W.B. Veldhuis
Dr. M.F. Bakker

TABLE OF CONTENTS

1

General introduction



Breast cancer is the most common form of cancer in women in the Netherlands: over 15,000 women are diagnosed with breast cancer yearly and approximately 3,000 die of breast cancer every year.[1,2]

Breast cancer survival differs depending on stage of disease at diagnosis: 5-year survival for advanced stage breast cancer (stage IV, metastatic disease, 5% of tumors at diagnosis) is only 28%, compared to a 99% 5-year survival for early stage breast cancers (stage I, 47% of tumors at diagnosis).[2,3]

The aims of breast cancer screening are reducing breast cancer related mortality by detecting breast cancer in an early stage, when the cancer is still asymptomatic. This often also results in a lower need for invasive therapies.

Several randomized controlled trials, conducted in the past century in Europe and the United States, have shown that mammography screening reduces breast cancer related mortality.[4–6]

In 1989, nationwide population-based mammography breast cancer screening was installed in the Netherlands, for women aged 50 to 70 years old. In 1998, the screening program was extended to women aged 70 to 74 years.[7]

The Dutch mammography screening program has, as in many European countries,[8] an age-based or "one size fits all" design. In the current Dutch screening program, all women aged between 50 to 75 years are offered mammography screening biennially. By 2014, nearly all countries in the European Union had implemented a population-based screening program with mammography,[8] although screening age ranges and screening frequencies differ between countries.

Yearly, approximately 1 million screening examinations are performed in the Dutch screening program. Of the women screened with mammography in 2017, 23.0 per 1000 women were referred for additional work-up, 6.6 per 1000 screened women were diagnosed with breast cancer, and false positive results occurred in 16.3 per 1000 screenings.[9] In total, 6,796 women were diagnosed with screen-detected breast cancer in 2017; 79% of these cancers were invasive and 21% were ductal carcinoma in situ.[9]

Since 1989, the incidence of breast cancer in the Netherlands has increased (Figure 1).[10] Part of this increase is the result of the implementation of the screening program (in 1989), however the incidence of breast cancer has also increased in younger women who are not invited for screening, and is probably explained by an increase in the occurrence of breast cancer risk factors, such as decreasing parity, increase in age at first birth, and an increase in obesity.[11–13]

In this same period, breast cancer related mortality has decreased, both in women of screening ages (50 to 75 years) and in younger women (Figure 2).[10] In women aged 50 to 74 years breast cancer mortality has decreased to 53.3 per 100.000 (European standardized rate)(2017),[9] which is a decrease in breast cancer related mortality of approximately 30% in women invited to screening, compared to 1986/1988.[14] Approximately half of the mortality reduction is estimated to be attributable to screening.[15–17] When comparing breast cancer related mortality in screening participants to that in invited women who do not participate, decrease in mortality is estimated at 58%.[18]

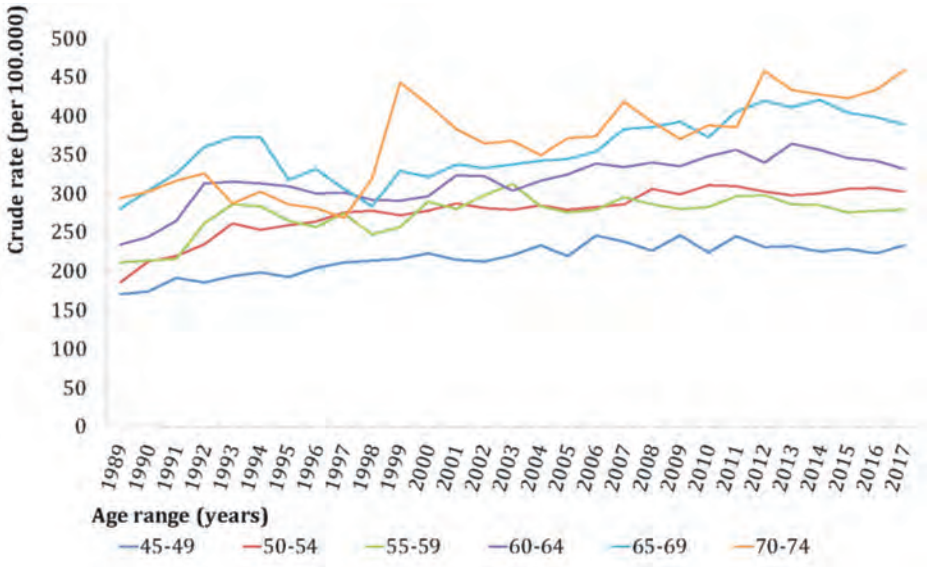


Figure 1. Breast cancer incidence rate in the Netherlands.[10]

The improvement of systemic therapies in the past decades has also contributed to the reduction in breast cancer related mortality.[16,19] Approximately 90% of all invasive breast cancers are treated with surgery[2]; this percentage has not changed since 1989. However, the proportion of women that is treated with breast sparing surgery, instead of mastectomy, has increased from 37% in 1989 to 57% in 2017.[2] Of all breast cancer patients in 2017, 45% received hormone therapy, 32% chemotherapy and 68% received radiotherapy.[20]

The reduction in breast cancer related mortality attributable to adjuvant therapies is estimated to be 14% (2008) in women aged 50 to 74 years.[16]

Besides these breast cancer therapy improvements, since the introduction of the screening program, the techniques of imaging modalities have also improved. In 2010, the Dutch breast cancer screening program completely changed from film mammography to digital mammography screening. Since the introduction of digital mammography, breast cancer detection rates increased from 5.13 per 1000 screenings (95% CI, 5.0 to 5.3) in 2004 to 6.34 per 1000 screenings (95%CI, 6.15 to 6.47) in 2011.[21] However, it is important to note that the policy regarding referral rates was changed in this period, which also resulted in an increase of referrals.[22,23]

Other European screening programs also show an increased detection after the introduction of digital mammography.[24,25] These studies reported that cancers presenting as microcalcifications increased from 1.3 per 1000 screenings to 1.9 per 1000 screenings[24] and digital mammography detects significantly more cancers depicted as microcalcifications (0.26%, compared to 0.12% with film mammography).[25]

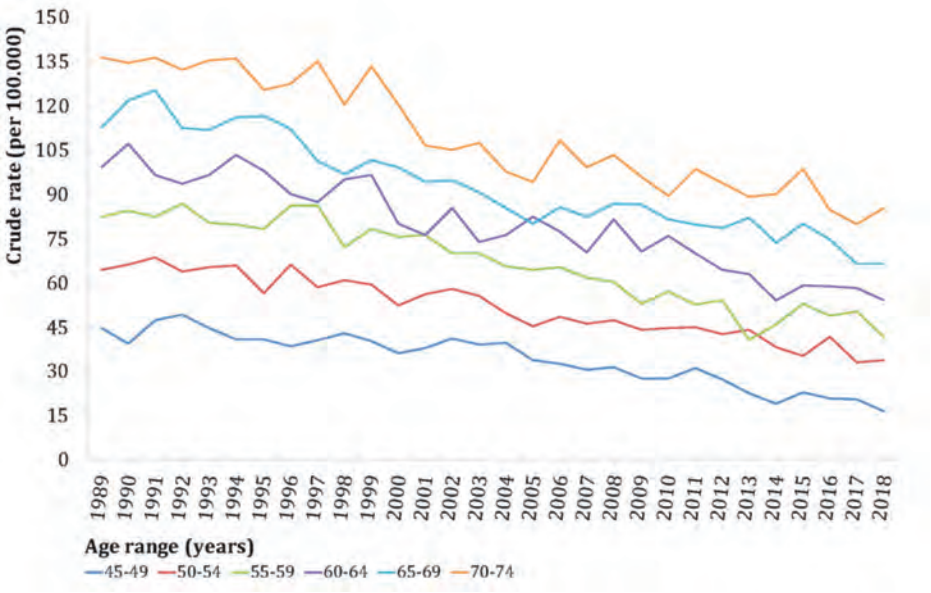


Figure 2. Breast cancer mortality in the Netherlands.[10]

Despite the reported breast cancer mortality reduction, mammography screening programs are not perfect. False positives rates in the Dutch screening program are approximately 16 per 1000 screenings,[9] false positive rates up to 57 per 1000 screenings are reported in other European screening programs,[26] and even much higher false positive rates are reported in the United States.[27]

Furthermore, among screening participants, still not all breast cancers are diagnosed with the screening program; approximately one third of cancers is not detected (2.2 per 1000 screenings). These cancers become symptomatic in the interval between screenings, before the next screening examination, after a false negative result at mammography. They are called interval cancers.

Compared to screen-detected cancers, interval cancers more often have a higher stage at diagnosis,[28,29] are more often hormone receptor (estrogen/progesterone) negative, human epidermal growth factor receptor 2 positive or triple negative and are more often associated with positive lymph nodes at diagnosis.[28]

Interval cancers can be divided into three different types. Approximately 24% of the interval cancers in the Dutch screening program are missed, despite double reading. Another 22% in hindsight showed minimal signs on the previous screening mammogram, but not enough signs to justify referral. The majority of the interval cancers (54%) were not visible at all on the previous screening mammogram, and are therefore called true interval cancers.[30] Cancers in this last category are fast growing tumors that did not exist yet or were too small to be detected on the last screening mammogram.

Tailored screening, based on individual breast cancer risk, could improve the efficacy and efficiency of breast cancer screening programs. The idea of tailored screening is that women at higher risk could be screened more frequently or with more sensitive screening methods than women at lower risk, and that women with lower risk could then be screened less frequently. This could result in a more favorable benefit–harm ratio for all women.

A woman's breast density, which is the amount of fibroglandular tissue in the breast, is an obvious candidate characteristic that could be used in the tailoring of breast cancer screening. On the one hand, high breast density is an important breast cancer risk factor,[31–33] and on the other hand it reduces sensitivity of mammography and also increases false positive results.[34–36] Breast density can be measured on mammograms, visually by the radiologist, according to the Breast Imaging-Reporting and Data system (BI-RADS) mammography density categories,[37] or semi- or fully automatically,[38–40] and is divided into four categories: a (almost entirely fatty breasts), b (scattered areas of density), c (heterogeneously dense breasts) and d (extremely dense breasts). Women in category c and d are considered to have dense breasts. Approximately 29% of the Dutch screening participants have heterogeneously dense breasts and 8% have extremely dense breasts.[34]

Breast density is, besides associated with age, also strongly associated with hormone-related factors, such as parity, age of the mother at first birth, menopausal status and use of menopausal hormone therapy.[41–43]

However, breast density is an independent risk factor for breast cancer.[31,33,42,44] Women with extremely dense breasts have a three to six-fold higher breast cancer risk than women with entirely fatty breasts,[31,33] and a two-fold higher risk than women with average breast density.[32,33] In particular, the risk of an interval cancer is increased in these women.[34,35,45]

The masking effect of tumors that breast density has on mammography, causes a reduced sensitivity of mammography.[32,36,46] Because of this reduced sensitivity of mammography, the benefits of mammography screening are less in women with dense breasts, while their risk is higher.

Supplemental imaging techniques such as ultrasound and magnetic resonance imaging (MRI) may improve cancer detection in women with dense breasts. Several studies conducted in the past decades in women with elevated breast cancer risks, showed that the addition of ultrasound or MRI to mammography increases breast cancer detection. The extra breast cancer detection of ultrasound was around 4 per 1000 screens and that of MRI over 10 to 15 per 1000 screenings.[47–51] However, false positive results were also increased and effects on outcomes such as interval cancers and breast cancer related mortality were not investigated.[47,50] It is important to realize that an increase in breast cancer detection does not necessarily lead to a decrease in breast cancer related mortality, since not all cancers diagnosed with screening would have progressed to symptomatic cancers. These cancers, that would have remained indolent in the absence of screening are overdiagnosed.

To investigate the effect of supplemental screening on breast cancer outcomes, randomized controlled trials on mortality would be preferable, but they require a very long time and a very large sample size. As an alternative, shorter-term outcomes, such as interval-cancer rates and advanced-cancer rates can be used, as a reduction in these rates is prerequisite for an eventual reduction in breast cancer mortality.[52]

The University Medical Center Utrecht initiated the Dense tissue and Early breast Neoplasm ScrEening (DENSE) trial,[53] the first randomized controlled trial to study the additional value of supplemental MRI in early detection of breast cancer in women with extremely dense breasts. The trial is embedded in the Dutch population-based screening program and is being carried out in collaboration with the National Institute for Public Health, the Dutch Expert Centre for Screening, regional screening organizations and 7 other large Dutch hospitals. Screening participants with extremely dense breasts and a negative (normal) result at screening mammography were randomized in a 1:4 ratio to a group that was invited for supplemental MRI screening (intervention arm) and a group that received mammography screening only (control arm). Only the women who were randomized to the group that received an invitation for supplemental MRI screening were informed about the trial and asked to participate.

This prerandomization design was used to prevent anxiety in the control group and to reduce the probability of self-referral induced contamination in the control arm.[54]

The primary outcome of the trial is the reduction of interval cancers in the intervention arm compared to the control arm. The trial consists of three consecutive screening rounds. Breast density is determined on all three mammograms, but changes in breast density after randomization do not affect eligibility. Women in the intervention arm are only eligible for the second or third MRI screening round after a negative result at regular screening mammography, respectively two or four years after the first screening round.

By conducting three screening rounds, important information can be obtained on the value of *ongoing* screening as compared to a *one-time-only* supplemental screening. In the first, *prevalence,* screening round, the cancer detection is always higher than the detection in subsequent, *incidence,* screening rounds. This is due to a prevalence peak that is caused by a preclinical pool of cancers that mammography does not detect and are detected because MRI is more sensitive than mammography.[55] By incorporating multiple screening rounds, the trial therefore provides information on cancer detection rates that can be expected in an ongoing screening program.

The trial started in 2011 and randomization was completed in 2015. In total, 8,061 women were randomized to the intervention arm and 32,312 women to the control arm of the trial. The third screening round of the trial is nearing completion. In this thesis we report the first results of the DENSE trial.

OUTLINE OF THE THESIS

In **chapter 2** we studied the primary outcome of the first screening round of the DENSE trial: the effect of supplemental MRI on the incidence of interval cancers in women with extremely dense breast tissue.

Not all women randomized to the intervention arm (invitation to supplemental MRI screening) of the DENSE trial, accepted the invitation to participate in supplemental MRI screening. In **chapter 3** we determined the willingness of women with extremely dense breasts to undergo breast screening with supplemental MRI in a research setting (the DENSE trial), and we examined reasons for women to participate or not.

Incidental findings are a disadvantage of nearly all cross-sectional imaging techniques, and might lead to (unnecessary) anxiety and treatment. The field of view of breast MRI does not only consist of breast tissue, surrounding tissues and organs are also visible on breast MRI. In **chapter 4** we studied the prevalence, type and clinical relevance of

extramammary incidental findings on the first screening round MRI examinations of the DENSE trial.

Two years after the first (prevalence) screening round of the DENSE trial, women were invited for a subsequent mammography screening round. In case of a negative mammography result, these women were again invited to participate in supplemental MRI screening. This second screening round shows cancer detection rates after the prevalence peak of the first screening round, and also the influence on the false positive rate of having previous MRI examinations available. The first results of the second MRI screening round (first incidence round) of the DENSE trial, are described in **chapter 5**.

Not all women with extremely dense breasts have the same breast cancer risk. Supplemental screening could be further tailored by using individual breast cancer risks, based on risk factors besides breast density. In **chapter 6** we investigated the effect of stratification by breast cancer risk on screening performance of supplemental MRI in the first and second screening round of the DENSE trial, in order to assess whether the harm-benefits balance of supplemental MRI screening is most favorable in women with highest estimated breast cancer risk. We stratified the participants by five-year breast cancer risk estimates using the Tyrer-Cuzick breast cancer risk model[56] and the Breast Cancer Surveillance Consortium breast cancer risk model.[57]

In **chapter 7** we discuss the benefits and harms of supplemental screening, the challenges of trials on (supplemental) screening, and possibilities of a transition from the current age-based or 'one-size fits all' mammography breast cancer screening program to tailored breast cancer screening.

REFERENCES

1. StatLine (Centraal bureau voor statistiek). Deaths; cause of death (extensive list), age and sex. Published December 19, 2019. Accessed June 11, 2020. https://opendata.cbs.nl/statline/#/CBS/en/dataset/7233ENG/table?dl=3B089
2. Integraal kankercentrum Nederland (IKNL). *Borstkanker in Nederland-Trends 1989-2017.*; 2018.
3. Integraal kankercentrum Nederland (IKNL). Incidentie borstkanker. Accessed June 11, 2020. https://www.iknl.nl/kankersoorten/borstkanker/registratie/incidentie
4. Shapiro S. Evidence on screening for breast cancer from a randomized trial. *Cancer*. 1977;39:2772-2782.
5. Roberts M, Alexander FE, Anderson T, et al. Edinburgh trial of screening for breast cancer: mortality at seven years. *The Lancet*. 1990;335:241-246.
6. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: Updated overview of the Swedish randomised trials. *The Lancet*. 2002;359(9310):909-919.
7. Health Council of the Netherlands. *Population Screening for Breast Cancer: Expectations and Developments.*; 2014.
8. Basu P, Ponti A, Anttila A, et al. Status of implementation and organization of cancer screening in The European Union Member States-Summary results from the second European screening report. *International Journal of Cancer*. 2018;142:44-56.
9. Integraal kankercentrum Nederland (IKNL) - Netherlands Cancer Registry. *National Evaluation of Breast Cancer Screening in the Netherlands 2017/2018.*; 2019. Accessed June 12, 2020. www.rivm.nl/monitoring-evaluatie-bevolkingsonderzoek-borstkanker
10. the Netherlands Comprehensive Cancer Organisation (IKNL). Nederlandse Kankerregistratie (NKR), IKNL. Accessed June 12, 2020. iknl.nl/nkr-cijfers
11. Statline (Centraal bureau voor statistiek). State of health of the Dutch population. Published March 25, 2010. Accessed June 13, 2020. https://opendata.cbs.nl/statline/#/CBS/en/dataset/7068eng/table?dl=3B202
12. StatLine (Centraal bureau voor statistiek). Life style and (preventive) health examination; personal characteristics. Published April 23, 2020. Accessed June 13, 2020. https://opendata.cbs.nl/statline/#/CBS/en/dataset/83021ENG/table?dl=3B200
13. StatLine (Centraal bureau voor statistiek). Birth; key figures. Published August 8, 2019. Accessed June 13, 2020. https://opendata.cbs.nl/statline/#/CBS/en/dataset/37422eng/table?dl=3B203
14. Sankatsing VDV, van Ravesteyn NT, Heijnsdijk EAM, et al. The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up. *International Journal of Cancer*. 2017;141(4):671-677.
15. Broeders M, Moss S, Nystrom L, et al. The impact of mammographic screening on breast cancer mortality in Europe: A review of observational studies. *Journal of Medical Screening*. 2012;19(SUPPL. 1):14-25.
16. de Gelder R, Heijnsdijk EAM, 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. *International Journal of Cancer*. 2015;137(1):165-172. Accessed June 11, 2020.

17. Marmot M, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: An independent review. *The Lancet*. 2012;380(9855):1778-1786.
18. Paap E, Verbeek ALM, Botterweck AAM, et al. Breast cancer screening halves the risk of breast cancer death: A case-referent study. *Breast*. 2014;23(4):439-444.
19. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *The Lancet*. 2011;378:771-784.
20. Integraal kankercentrum Nederland (IKNL). Behandeling - IKNL. Accessed June 13, 2020. https://www.iknl.nl/kankersoorten/borstkanker/registratie/behandeling
21. 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.
22. National Evaluation Team for Breast cancer screening (NETB). *National Evaluation of Breast Cancer Screening in the Netherlands 1990-2007*.; 2009.
23. Otten JDM, Karssemeijer N, Hendriks JHCL, et al. Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. *Journal of the National Cancer Institute*. 2005;97(10). https://academic.oup.com/jnci/article-abstract/97/10/748/2544017
24. Hambly NM, McNicholas MM, Phelan N, Hargaden GC, O'Doherty A, Flanagan FL. Comparison of digital mammography and screen-film mammography in breast cancer screening: A review in the Irish breast screening program. *American Journal of Roentgenology*. 2009;193:1010-1018.
25. del Turco MR, Mantellini P, Ciatto S, et al. Full-field digital versus screen-film mammography: Comparative accuracy in concurrent screening cohorts. *American Journal of Roentgenology*. 2007;189(4):860-866.
26. Sala M, Salas D, Belvis F, et al. Reduction in false-positive results after introduction of digital mammography: Analysis from four population-based breast cancer screening programs in Spain. *Radiology*. 2011;258(2):388-395.
27. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of breast cancer screening: systematic review to update the 2009 U.S. Preventive Services Task Force recommendation. *Annals of Internal Medicine*. 2016;164(4):256-267.
28. Holm J, Humphreys K, Li J, et al. Risk factors and tumor characteristics of interval cancers by mammographic density. *Journal of Clinical Oncology*. 2015;33(9):1030-1037.
29. National evaluation team for breast cancer screening (NETB). *National Evaluation of Breast Cancer Screening in the Netherlands 2004 – 2014*.; 2019.
30. Weber RJP, van Bommel RMG, Louwman MW, et al. Characteristics and prognosis of interval cancers after biennial screen-film or full-field digital screening mammography. *Breast Cancer Research and Treatment*. 2016;158(3):471-483.
31. Vachon CM, van Gils CH, Sellers TA, et al. Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Research*. 2007;9(6).
32. Boyd NF, Guo H, Martin LJ, et al. Mammographic Density and the Risk and Detection of Breast Cancer. *The New England Journal of Medicine*. 2007;356:227-236.
33. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiology Biomarkers and Prevention*. 2006;15(6):1159-1169.

34. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95-103.
35. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval-and screen-detected cancers. *Journal of the National Cancer Institute*. 2000;92:1081-1087.
36. Kerlikowske K. The mammogram that cried wolfe. *The New England Journal of Medicine*. 2007;356(3):297-300.
37. D'Orsi C, Sickles E, Medelson E, Morrison E. *ACR BI-RADS Atlas: Breast Imaging Reporting and Data System*. 5th ed. American College of Radiology; 2013.
38. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. *IEEE Transactions on Medical Imaging*. 2006;25(3):273-282.
39. Ciatto S, Bernardi D, Calabrese M, et al. A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. *Breast*. 2012;21(4):503-506.
40. Byng JW, Yaffe MJ, Jong RA, et al. Analysis of mammographic density and breast cancer risk from digitized mammograms. *Radiographics*. 1998;18(6):1587-1598.
41. Yaghjyan L, Colditz GA, Rosner B, Bertrand KA, Tamimi RM. Reproductive factors related to childbearing and mammographic breast density. *Breast Cancer Research and Treatment*. 2016;158(2):351-359.
42. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. *Journal of the National Cancer Institute*. 2007;99:1178-1187.
43. Kelsey JL, Gammon MD, John EM. Reproductive Factors and Breast Cancer. *Epidemiologic Reviews*. 1993;15(1):36-47.
44. Pettersson A, Graff RE, Ursin G, et al. Mammographic density phenotypes and risk of breast cancer: A meta-analysis. *Journal of the National Cancer Institute*. 2014;106(5).
45. Kerlikowske K, Zhu W, Tosteson ANA, et al. Identifying women with dense breasts at high risk for interval cancer a cohort study. *Annals of Internal Medicine*. 2015;162(10):673-681.
46. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *The New England Journal of Medicine*. 2005;353(17):1773-1783.
47. 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: The Journal of the American Medical Association*. 2012;307(13):1394-1404.
48. 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: The Journal of the American Medical Association*. 2008;299(18):2151-2163.
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. *Journal of Clinical Oncology*. 2010;28(9):1450-1457.

50. 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. *Journal of Clinical Oncology*. 2014;32(22):2304-2310.
51. Corsetti V, Houssami N, Ghirardi M, et al. Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: Interval breast cancers at 1 year follow-up. *European Journal of Cancer*. 2011;47(7):1021-1026.
52. Irwig L, Houssami N, Armstrong B, Glasziou P. Evaluating new screening tests for breast cancer. *BMJ*. 2006;332(7543):678-679.
53. Emaus MJ, Bakker MF, Peeters PHM, 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.
54. Zelen M. A new design for randomized clinical trials. *The New England Journal of Medicine*. 1979;300(22):1242-1245.
55. Gulati R, Feuer EJ, Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies. *American Journal of Epidemiology*. 2016;184:140-147.
56. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in Medicine*. 2004;23(7):1111-1130.
57. 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. *Journal of Clinical Oncology*. 2015;33(28):3137-3143.

2

Supplemental MRI screening for women with extremely dense breast tissue

M.F. Bakker*, S.V. de Lange*, R.M. Pijnappel, R.M. Mann, P.H.M. Peeters, E.M. Monninkhof, M.J. Emaus, C.E. Loo, R.H.C. Bisschops, M.B.I. Lobbes, M.D.F. de Jong, K.M. Duvivier, J. Veltman, N. Karssemeijer, H.J. de Koning, P.J. van Diest, W.P.T.M. Mali, M.A.A.J. van den Bosch, W.B. Veldhuis**, C.H. van Gils**, for the DENSE Trial Study Group.

* Shared first authorship
** Shared last authorship

The New England Journal of Medicine 2019; 381:2091-2102

ABSTRACT

Background

Extremely dense breast tissue is a risk factor for breast cancer and limits the detection of cancer with mammography. Data are needed on the use of supplemental magnetic resonance imaging (MRI) to improve early detection and reduce interval breast cancers in such patients.

Methods

In this multicenter, randomized, controlled trial in the Netherlands, we assigned 40,373 women between the ages of 50 and 75 years with extremely dense breast tissue and normal results on screening mammography to a group that was invited to undergo supplemental MRI or to a group that received mammography screening only. The groups were assigned in a 1:4 ratio, with 8,061 in the MRI-invitation group and 32,312 in the mammography-only group. The primary outcome was the between-group difference in the incidence of interval cancers during a 2-year screening period.

Results

The interval-cancer rate was 2.5 per 1000 screenings in the MRI-invitation group and 5.0 per 1000 screenings in the mammography-only group, for a difference of 2.5 per 1000 screenings (95% confidence interval [CI], 1.0 to 3.7; P<0.001). Of the women who were invited to undergo MRI, 59% accepted the invitation. Of the 20 interval cancers that were diagnosed in the MRI-invitation group, 4 were diagnosed in the women who actually underwent MRI (0.8 per 1000 screenings) and 16 in those who did not accept the invitation (4.9 per 1000 screenings). The MRI cancer-detection rate among the women who actually underwent MRI screening was 16.5 per 1000 screenings (95% CI, 13.3 to 20.5). The positive predictive value was 17.4% (95% CI, 14.2 to 21.2) for recall for additional testing and 26.3% (95% CI, 21.7 to 31.6) for biopsy. The false positive rate was 79.8 per 1000 screenings. Among the women who underwent MRI, 0.1% had either an adverse event or a serious adverse event during or immediately after the screening.

Conclusions

The use of supplemental MRI screening in women with extremely dense breast tissue and normal results on mammography resulted in the diagnosis of significantly fewer interval cancers than mammography alone during a 2-year screening period.

(Funded by the University Medical Center Utrecht and others; DENSE ClinicalTrials.gov number, NCT01315015.)

INTRODUCTION

Women with extremely dense breast tissue have an increased risk of breast cancer, and their cancers are also less likely to be detected on mammography.[1–3] Such patients may benefit from a tailored breast-screening strategy, supplemented with more sensitive imaging methods. The benefit of supplemental imaging is the subject of a worldwide debate. In the United States, a federal law directs breast-density reporting,[4] but supplemental screening is not recommended in American guidelines.[5] Although supplemental imaging increases the rate of cancer detection in women with dense breasts,[6] the question remains whether it improves health outcomes. The first indication for a reduction in morbidity and mortality is a reduction in the incidence of interval cancers, since such a reduction may mean that cancers that would otherwise have become symptomatic would now be detected earlier.[7,8]

The Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial is a randomized, controlled trial to study the effect of supplemental magnetic resonance imaging (MRI) on the incidence of interval cancers in women with extremely dense breast tissue. Here, we present the primary outcome of the first 2-year screening round of the DENSE trial.

METHODS

Trial design and population

The trial design has been described in detail previously.[9] In the multicenter DENSE trial, we enrolled women who were participating in the Dutch population-based digital mammography screening program, which is conducted every 2 years for women between the ages of 50 and 75 years.[10,11] From December 2011 through November 2015, we enrolled screening participants who had negative results on mammography and who had extremely dense breast tissue, which was defined as grade 4 density as measured on Volpara imaging software, version 1.5 (Volpara Health Technologies).[12] Volpara density grades range from 1 to 4 (classified as “a” to “d” in the latest version) and correspond to the four-point breast-density categories of the Breast Imaging, Reporting, and Data System (BI-RADS) of the American College of Radiology, which range from almost entirely fatty tissue to extremely dense tissue. (Details regarding the grading system are provided in the Supplementary Appendix, available with the full text of this article at NEJM.org.)[12,13]

We randomly assigned the women in a 1:4 ratio to a group that was invited to undergo supplemental MRI or to a group that received mammography screening only. The women underwent central randomization with the use of a computer-based program in varying block sizes, stratified according to hospital (among the eight participating

centers) and regional screening organization (among the four participating regions). After randomization, only the women who had been assigned to the MRI-invitation group were informed about the group assignment and were asked to participate in the trial. This process was performed according to the Zelen design[14] of randomization before informed consent was obtained. This design was used to prevent anxiety in the control group and to reduce the probability that women in the control group would arrange for MRI examination on their own initiative. For all women who had undergone randomization, data were gathered regarding breast density, age, socioeconomic status, and urbanization level.[15,16] The participants who underwent MRI screening received a travel allowance of €20 ($20.24 U.S.). The MRI examination was financed with grant money.

Trial oversight

On November 11, 2011, the trial was approved by the Dutch Minister of Health, Welfare, and Sport, under advisement from the Health Council of the Netherlands. The trial was financially supported by the University Medical Center Utrecht, the Netherlands Organization for Health Research and Development, the Dutch Cancer Society, the Dutch Pink Ribbon–A Sister's Hope organization, Stichting Kankerpreventie Midden-West, and Bayer Pharmaceuticals, with an in-kind contribution from Volpara Health Technologies. The authors designed the trial, gathered and analyzed the data, and wrote the manuscript. With the exception of the University Medical Center Utrecht, none of the funders had any role in these tasks. All the authors vouch for the accuracy and completeness of the data and analysis and for the adherence of the trial to the protocol (available at NEJM.org).

Imaging

All the trial participants had a negative result on regular mammographic screening (bilateral craniocaudal and mediolateral oblique views). A negative result was defined as a BI-RADS radiographic score of 1 or 2 on a 6-point scale on which a higher score reflects a greater cancer risk. A BI-RADS score of 6 (which indicates known biopsy-confirmed cancer) was not used for evaluation in this screening trial.[13,17] All MRI examinations were performed on 3.0 Tesla systems with the use of a dedicated bilateral breast coil. (A link to an interactive mobile app with additional data about the trial is provided in the Supplementary Appendix.)

Single-read MRI examinations were performed according to the BI-RADS MRI lexicon[18] and were conducted by breast radiologists whose experience ranged from 5 to 23 years. All the participants who had a BI-RADS score of 4 or 5 were recalled for additional workup. In participants with a BI-RADS score of 3, double reading of the MRI was performed, and if there was consensus on a score of 3, follow-up imaging with MRI after 6 months was planned. The results of the follow-up MRI had to be reported as either negative (BI-RADS score of 1 or 2, with a return to the regular screening program)

or positive (BI-RADS score of 4 or 5, with recall). Women in the mammography-only group received the standard of care, which consisted of the regular screening program with invitations to undergo mammography every 2 years.

Primary and secondary outcomes

The primary outcome was the between-group difference in the incidence of interval cancers. We collected data regarding the number of interval cancers and the tumor characteristics in the two groups through linkage with the Netherlands Cancer Registry. In the MRI-invitation group, cancers were detected either on the screening MRI or at the 6-month repeat screening, if applicable. Interval cancers included all the breast cancers that were diagnosed after negative results on mammography before the next scheduled mammography examination. If no mammography was scheduled (e.g., because of an age of >75 years), an interval cancer was defined as one diagnosed within 24 months after the negative results on mammography. This definition presumes that the interval cancer would have been detected on subsequent mammography.

Key secondary outcomes included the recall rate for additional examination, the cancer-detection rate on MRI, the false positive rate, the positive predictive value, and tumor characteristics. The recall rate was defined as the percentage of participants who had a positive result on MRI screening among all the women who had undergone MRI screening. A BI-RADS score of 3, 4, or 5 was considered to be a positive result on MRI. For women who had more than one lesion, recall was based on the lesion with the highest BI-RADS score. The cancer-detection rate on MRI was defined as the percentage of women with a positive result on MRI screening that resulted in histologically confirmed breast cancer among all the women who had undergone MRI screening. The false positive rate was defined as the percentage of women who had a positive result on screening MRI but who were later found not to have breast cancer. In calculating the positive predictive value of recall after positive results on MRI, three measures were used. The first measure (PPV1) was the percentage of women with cancers detected on MRI screening among all the participants who had positive results on MRI. For the second measure (PPV2), the denominator consisted of all the women who had an indication for biopsy (BI-RADS score of 4 or 5). For the third measure (PPV3), the denominator consisted of all the women who had undergone biopsy.[13,17]

The program sensitivity among the women who were screened with MRI is the number of women with cancers detected on MRI screening among all the women with screening-detected or interval cancers. Descriptions were provided regarding the tumor–node–metastasis (TNM) stage, grade, morphology, and receptor status. In women with more than one tumor, we described the one with the highest TNM stage. Adverse events and serious adverse events were recorded in the trial center during or immediately after the MRI examination or reported by the women within 30 days.

Statistical analysis

The trial was designed to have power of 80% to detect a between-group difference in the interval-cancer rate of 1.95 per 1000 screenings in the intention-to-screen population.[9] Interval-cancer rates were calculated as the number of interval cancers per 1000 screenings and per 1000 person-years of follow-up. Follow-up was calculated as the time from a negative result on mammography until a diagnosis of breast cancer, emigration, loss to follow-up, or death or until the next screening mammography took place or was scheduled to take place according to invitation, whichever occurred first. If the women were no longer invited to participate in the screening program because of age, a fixed follow-up of 24 months was adopted.

Complier average causal effect (CACE) analysis[19] was applied to estimate the effect of actually undergoing supplemental MRI screening in the subpopulation of women who said that they would have accepted MRI screening if it had been offered. CACE analysis was performed with the use of an instrumental-variables method in which the instrumental variable was the randomization to MRI invitation.[20–22] For this analysis, the interval-cancer rate among the MRI participants (i.e., those who actually underwent MRI examination) was compared with the rate among women who would have accepted MRI screening if it had been offered in the mammography-only group. We calculated 95% confidence intervals for differences in interval-cancer rates using a bootstrap-resampling method. (Details, formulas, and assumptions are provided in Figure S2 in the Supplementary Appendix.) The type I error rate (alpha) was set at 0.05. All the analyses were performed with the use of RStudio software, version 1.0.143.

RESULTS

Characteristics of trial population

Of the 40,373 women who underwent mammography screening, 8,061 were assigned to a group that was invited to undergo MRI and 32,312 were assigned to a group that received mammography only. Of the 8,061 women who were invited to undergo MRI, 4,783 (59%) actually underwent MRI screening (Figure 1 and Figure S1). Details regarding participation have been reported previously.[16] The MRI-invitation group and the mammography-only group were well balanced with respect to baseline characteristics (Table 1).

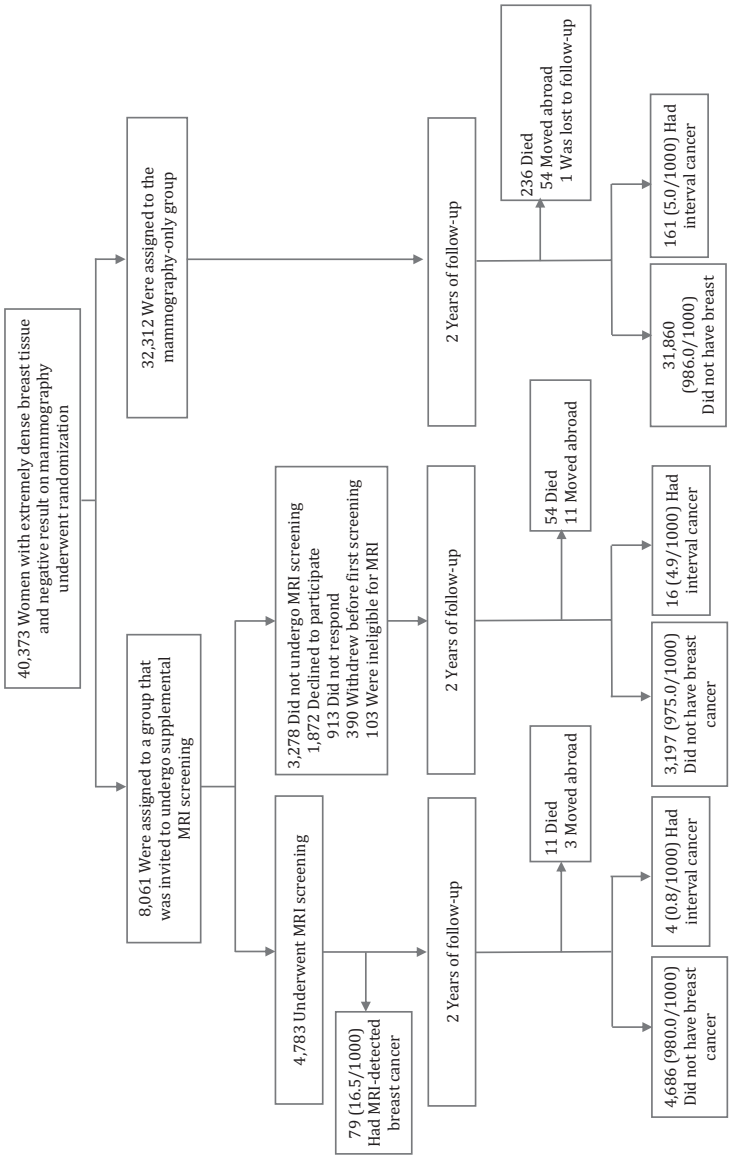


Figure 1. Enrollment and follow-up in 2-year screening program.
The women, who were between the ages of 50 and 75 years and had extremely dense breast tissue, were assigned in a 1:4 ratio to a group that was invited to undergo supplemental magnetic resonance imaging (MRI) screening or to a group that received mammography screening only. Approximately 59% of the participants in the MRI-invitation group actually underwent the supplemental procedure. Interval-cancer rates are the number of interval cancers diagnosed per 1000 screenings.

Table 1. Characteristics of the women at baseline*

	MRI-invitation group†			Mammography-only group
Characteristic	Participants (N=4,783)	Nonparticipants (N=3,278)	Total (N=8,061)	(N=32,312)
Median age (IQR), years	54 (51-59)	56	55 (51-61)	54 (51-61)
Median time between mammography and MRI (IQR), weeks	10 (8-14)		10 (8-14)	
Screening region, no. (%) ‡				
Midwestern	1,963 (41.0)	1,365 (41.6)	3,328 (41.3)	13,344 (41.3)
Eastern	1,219 (25.5)	775 (23.6)	1,994 (24.7)	7,992 (24.7)
Southwestern	623 (13.0)	450 (13.7)	1,073 (13.3)	4,301 (13.3)
Southern	978 (20.4)	688 (21.0)	1,666 (20.7)	6,674 (20.7)
Socioeconomic status, no. (%) §				
Quartile 4: highest	1,828 (38.2)	1,114 (34.0)	2,942 (36.5)	11,646 (36.0)
Quartile 3	1,144 (23.9)	775 (23.6)	1,919 (23.8)	7,620 (23.6)
Quartile 2	1,083 (22.6)	725 (22.1)	1,808 (22.4)	7,350 (22.7)
Quartile 1: lowest	716 (15.0)	656 (20.0)	1,372 (17.0)	5,655 (17.5)
Missing data	12 (0.3)	8 (0.2)	20 (0.2)	41 (0.1)
Urbanization level, no (%) ¶				
Extremely urban	871 (18.2)	703 (21.4)	1,574 (19.5)	6,527 (20.2)
Strongly urban	1,462 (30.6)	1,080 (32.9)	2,542 (31.5)	10,357 (32.1)
Moderately urban	961 (20.1)	584 (17.8)	1,545 (19.2)	6,320 (19.6)
Slightly urban	714 (14.9)	467 (14.2)	1,181 (14.7)	4,545 (14.1)
Not urban	713 (14.9)	400 (12.2)	1,113 (13.8)	4,074 (12.6)
Missing data	62 (1.3)	44 (1.3)	106 (1.3)	489 (1.5)

* In the group that was invited to undergo magnetic resonance imaging (MRI), those who actually underwent MRI are identified as "MRI participants" and those who declined are identified as "MRI nonparticipants." Percentages may not total 100 because of rounding.
IQR denotes interquartile range, and NA not applicable.

† Of the 8,061 women who were invited to undergo MRI screening, 4,783 (59%) actually underwent the screening.

‡ Data regarding region were missing for 1 woman in the mammography-only group

§ The socioeconomic status according to quartile is presented as the distribution of the Dutch population in 2014. These data were available for postal codes in neighborhoods with more than 100 households.

¶ The urbanization level was determined as the number of addresses per square kilometer on the basis of postal codes. These numbers range from 0 to 499 for not urban, 500 to 999 for slightly urban, 1,000 to 1,499 for moderately urban, 1,500 to 2,499 for strongly urban, and 2,500 or more for extremely urban.

Primary outcome

In the MRI-invitation group, an interval cancer was diagnosed in 20 women (4 among the MRI screening participants and 16 among the nonparticipants who were invited but did not undergo screening) of 8,061. In the mammography-only group, an interval cancer was diagnosed in 161 of 32,312 women, which resulted in an interval-cancer rate of 2.5 per 1000 screenings (95% confidence interval [CI], 1.6 to 3.8) in the MRI-invitation group and 5.0 per 1000 screenings (95% CI, 4.3 to 5.8) in the mammography-only group (Table 2). In the intention-to-screen analysis, the interval-cancer rate was lower by 2.5 per 1000 screenings (95% CI, 1.0 to 3.7) in the MRI-invitation group than in the mammography-only group (P<0.001).

In an analysis based on person-years, the interval-cancer rate was lower by 1.3 per 1000 person-years (95% CI, 0.6 to 1.9) in the MRI-invitation group. The exclusion of women in whom ductal carcinoma in situ (DCIS) was diagnosed did not change this result. Table S1 shows the incidence of cancers over time and includes a sensitivity analysis of different follow-up times since the receipt of negative results on MRI.

Using CACE analysis, we estimated that supplemental MRI screening among the subgroup of women who would have accepted MRI screening if it had been offered was associated with an interval-cancer rate that was lower by 4.2 per 1000 screenings (95% CI, 2.0 to 6.4) than that associated with mammography alone (P<0.001) (Figure S2).

Table 2. Interval-cancer rates and rate difference between trial groups, according to two analysis methods *

Type of Analysis	MRI-invitation group	Mammography-only group	Rate difference (95% CI)
Intention-to-screen analysis			
Women with interval cancer, no./total no.	20/8,061	161/32,312	
Interval-cancer rate (95% CI)			
No. per 1000 screenings	2.5 (1.6-3.8)	5.0 (4.3-5.8)	2.5 (1.0-3.7)
No. per 1000 person-years	1.3 (0.7-1.8)	2.5 (2.1-2.9)	1.3 (0.6-1.9)
CACE analysis†			4.2 (2.0-6.4)
MRI participants			
Participants with interval cancer, no./total no.	4/4,783		
Interval-cancer rate per 1000 screenings	0.8		
MRI nonparticipants			
Nonparticipants with interval cancer, no./total no.	16/3,278		
Interval-cancer rate per 1000 screenings	4.9		
Mammography-only participants who would have accepted MRI screening if offered‡			
Women with interval cancer, no./total no.		97/19,172	
Interval-cancer rate per 1000 screenings		5.1	
Mammography-only participants who would not have accepted MRI screening if offered‡			
Women with interval cancer, no./total no.		64/13,140	
Interval-cancer rate per 1000 screenings		4.9	

* CI denotes confidence interval.

† CACE (complier average causal effect) analysis was performed with the use of an instrumental-variables method in which the instrumental variable was the randomization to MRI invitation. For this analysis, the interval-cancer rate among the MRI participants (i.e., those who actually underwent MRI examination) was compared with the rate among the women who would have accepted MRI screening if offered in the mammography-only group.

‡ The values in these analyses were not observed but were estimated on the assumption that the interval-cancer rate per 1000 screenings among the nonparticipants in the MRI-invitation group would be the same as that among potential nonparticipants in the mammography-only group.

Secondary outcomes

Among the 4,783 MRI participants, the recall rate was 94.9 per 1000 screenings (95% CI, 86.9 to 103.6), and the cancer-detection rate with MRI was 16.5 per 1000 screenings (95% CI, 13.3 to 20.5) (Table 3). The positive predictive value of a positive MRI result (PPV1) was 17.4% (95% CI, 14.2 to 21.2), the positive predictive value of an indication for biopsy (PPV2) was 23.9% (95% CI, 19.6 to 28.8), and the positive predictive value of a biopsy (PPV3) was 26.3% (95% CI, 21.7 to 31.6). The false positive rate was 79.8 per 1000 screenings (95% CI, 72.4 to 87.9) (specificity, 92%). As a result of the MRI screening, 300 women underwent a breast biopsy; of these women, breast cancer was diagnosed in 79 (64 with invasive breast cancer and 15 with DCIS).

The program sensitivity of MRI screening was 95.2% (95%CI, 88.1 to 98.7). Table 4 shows the characteristics of all the cancers detected on MRI screening and interval tumors. The screening-detected tumors were smaller on average than those in the other groups. The median size of invasive tumors was 9.5 mm (interquartile range, 10.5 to 17.0) among MRI participants with interval cancers, 15.0 mm (interquartile range, 12.0 to 31.0) among MRI nonparticipants, and 17.0 mm (interquartile range, 12.0 to 23.0) among women in the mammography-only group.

Among the MRI participants, the absolute incidence of invasive (ductal and lobular) cancers was higher than that in the mammography-only group, as was the absolute incidence of DCIS and tubular cancers; the latter may have an indolent disease course. The absolute incidence of node-negative and early-stage cancers was also higher among MRI participants. This finding was accompanied by a slightly lower rate of late-stage cancers, but the number of such cancers that were detected was small, and a decrease in the number of late-stage cancers was not expected until after several years of follow-up.[23,24] Cancers that were detected in the MRI-invitation group appeared to be better differentiated and more often were hormone-receptor positive than those in the mammography-only group. At the next mammography screening, the cancer-detection rate was 2.0 per 1000 mammography screenings among the MRI participants, as compared with 7.1 per 1000 screenings among the MRI nonparticipants and 6.0 per 1000 screenings among the women in the mammography-only group; these findings indicate that MRI examination advanced the time of diagnosis (Table S1). Among the MRI participants, 0.1% reported either an adverse event or a serious adverse event during or immediately after the MRI examination (Table S2). These events were related to vasovagal responses, contrast reactions, or intravenous line infiltration. In addition, Table S2 shows all adverse events and serious adverse events that were reported by the MRI participants on a 30-day questionnaire that surveyed all health problems, regardless of a connection to the MRI examination.

Table 3. Evaluation of supplemental MRI screening

Variable	Participants who underwent MRI screening (N=4,783) *no./total no.*	*(%)*	Rate (95% CI) *no./1000 screenings*	
First round of screening MRI				
Participants who were recalled for additional evaluation*	454/4,783	(9.5)		
BI-RADS 3	150/454	(33.0)		
BI-RADS 4	286/454	(63.0)		
BI-RADS 5	18/454	(4.0)		
Participants who had indication for biopsy	331/4,783	(6.9)		
BI-RADS 4 or 5 on first MRI	304/331	(91.8)		
BI-RADS 4 or 5 on 6-mo follow-up MRI after initial BI-RADS 3†	27/331	(8.2)		
Participants who underwent biopsy‡	300/4,783	(6.3)		
After BI-RADS 4 or 5 on first MRI	276/300	(92.0)		
After BI-RADS 4 or 5 on 6-mo follow-up MRI after initial BI-RADS 3	24/300	(8.0)		
Women with confirmed cancers after positive MRI screening §	79/4,783	(1.7)		
Type of cancer				
Ductal carcinoma in situ	15/79	(19.0)		
Invasive cancer	64/79	(81.0)		
Recall rate	454/4,783	(9.5)	94.9	(86.9-103.6)
Biopsy rate	300/4,783	(6.3)	62.7	(56.2-70.0)
All cancers				
Cancer-detection rate	79/4,783	(1.7)	16.5	(13.3-20.5)
False positive rate	375/4,700	(8.0)	79.8	(72.4-87.9)
Measure of positive predictive value¶				
1	79/454	(17.4)		
2	79/331	(23.9)		
3	79/300	(26.3)		
Invasive cancers				
Cancer-detection rate	64/4,783	(1.3)	13.4	(10.5-17.1)
False positive rate	390/4,715	(8.3)	82.7	(75.2-90.0)
Measure of positive predictive value¶				
1	64/454	(14.1)		
2	64/331	(19.3)		
3	64/300	(21.3)		

* The assessment categories of the Breast Imaging, Reporting, and Data System (BI-RADS) of the American College of Radiology include scores ranging from 0 to 6 as follows: incomplete examination, 0; negative, 1; benign, 2; probably benign, 3; suspicious, 4; highly suggestive of cancer, 5; and known biopsy-confirmed cancer, 6. Category 6 was not used in the evaluation of screening results in this trial.

† A BI-RADS score of 4 or 5 on the follow-up MRI was the indication for biopsy.

‡ Of the 331 participants who had an indication for biopsy, 31 did not undergo the procedure because the lesion was no longer visible on additional imaging (in 18 participants), the lesion was an intramammary lymph node or cyst (in 8), biopsy of the lesion was technically not possible (in 2), or the lesion was already histologically proven to be benign (in 3).

§ Two synchronous cancers were diagnosed in 1 participant; only the tumor with the highest stage was used in the analyses. In 4 participants, breast cancer was diagnosed on biopsy after the 6-month follow-up MRI when the women had an initial BI-RADS score of 3.

¶ The positive predictive value is the proportion of participants who had confirmed breast cancer after positive results on MRI screening. Measure 1 included those who had a positive MRI result (BI-RADS score of 3, 4, or 5); measure 2, those who had an indication for biopsy (BI-RADS score of 4 or 5, including those with a BI-RADS score of 3 on initial MRI and a score of 4 or 5 on follow-up MRI); and measure 3, those who underwent biopsy (BI-RADS score of 4 or 5, including those with a score of 4 or 5 on follow-up MRI).

DISCUSSION

In the first round of MRI screening of women between the ages of 50 and 75 years with extremely dense breast tissue (defined as Volpara grade 4 [or d]) and negative results on mammography, we observed a significantly lower interval-cancer rate than in the mammography-only group (2.5 vs. 5.0 per 1000 screenings) in the intention- to-screen analysis. Among the women who were invited to undergo MRI, 59% actually underwent the procedure. Of the 20 interval cancers diagnosed in the MRI-invitation group, 4 were diagnosed in the women who had undergone MRI and 16 in those who had not. Among the women who would have accepted MRI screening if it had been offered, the incidence of interval cancers after the use of supplemental MRI screening was estimated to be lower by 4.2 per 1000 screenings than the incidence after mammography alone on CACE analysis, which resulted in an interval-cancer rate similar to that observed on mammography in women with very fatty breasts (Volpara grade 1 [or a]).[3] Depending on the proportion of women who would accept MRI screening in clinical practice, the effect at the population level could be closer to the effect in either the intention-to-screen analysis or the CACE analysis.

Undergoing supplemental MRI was associated with a cancer-detection rate of 16.5 per 1000 screenings and resulted in a false positive rate of 8.0% (79.8 per 1000 screenings). Of the women who underwent a breast biopsy on the basis of an MRI indication, 26.3% had breast cancer and 73.7% did not.

In J-START (Japan Strategic Anti-cancer Randomized Trial),[25] in which investigators evaluated supplemental ultrasonographic breast screening among Japanese women between the ages of 40 and 49 years, 58% of the participants had heterogeneously or extremely dense breasts (similar to Volpara grade 3 or 4 [c or d]). Cancer-detection rates were 3.3 per 1000 screenings for mammography alone and 5.0 per 1000 screenings for mammography plus ultrasonography. The addition of ultrasonographic screening resulted in an interval-cancer rate of 0.5 per 1000 screenings, as compared with 1.0 per 1000 screenings with mammography alone, and an increase in the false positive rate from 8.8% to 12.6%. In the Japanese trial, the baseline interval-cancer rates were much lower than those in our trial, which may be related to a lower risk of breast cancer in this population; other factors include the 1-year interval between screenings and the lack of preselection of women with extremely dense breasts.

In several paired studies of MRI screening and mammography involving women with dense breast tissue, all the women underwent MRI, so investigators could not evaluate the effect of such screening on the interval-cancer rate. In a study involving 612 women between the ages of 25 and 91 years who had breast density similar to Volpara grade 3 or 4 (c or d) and at least one other risk factor for breast cancer, Berg et al.[26] observed a cancer-detection rate with MRI and mammography together that was

Table 4. Prognostic characteristics of cancers detected on MRI screening and interval cancers*

Characteristic	MRI participants (N=4,783)				MRI nonparticipants (N=3,278)		Mammography-only group (N=32,312)	
	Cancers detected on MRI screening		Interval cancers		Interval cancers		Interval cancers	
	no.	*no./1000 screenings*	*no.*	*no./1000 screenings*	*no.*	*no./1000 screenings*	*no.*	*no./1000 screenings*
Women with diagnosed cancer†	79		4		16		161	
Histologic type								
DCIS‡	15	3.1	0		2	0.6	9	0.3
Invasive ductal carcinoma	35	7.3	2	0.4	10	3.1	113	3.5
Invasive lobular carcinoma	9	1.9	2	0.4	4	1.2	20	0.6
Mixed invasive ductal and lobular carcinoma	8	1.7	0		0		3	0.1
Tubular carcinoma	7	1.5	0		0		2	0.1
Other invasive carcinoma	5	1.0	0		0		14	0.4
Status for axillary lymph nodes§								
Negative	70	14.6	2	0.4	9	2.7	89	2.6
Positive	9	1.9	2	0.4	7	2.1	72	2.2
Tumor stage¶‖								
Early (0 or I)	72	15.1	2	0.4	8	2.4	67	2.1
Late (II, III, or IV)	7	1.5	2	0.4	8	2.4	94	2.9
Tumor grade								
DCIS								
I, well-differentiated	6	1.3	0		0		3	0.1
II, moderately differentiated	6	1.3	0		1	0.3	1	<0.1
III, poorly differentiated	3	0.6	0		1	0.3	4	0.1
Missing data or could not be assessed	0		0		0		1	

Table 4. Continued.

Characteristic	MRI participants (N=4,783)				MRI nonparticipants (N=3,278)		Mammography-only group (N=32,312)	
	Cancers detected on MRI screening		Interval cancers		Interval cancers		Interval cancers	
Invasive								
I, well-differentiated	31	6.5	0		0		29	0.9
II, moderately differentiated	24	5.0	2	0.4	7	2.1	70	2.2
III, poorly differentiated	4	0.8	1	0.2	4	1.2	39	1.2
Missing data or could not be assessed	5		1		3		14	
Receptor status**								
Positive for estrogen receptor, progesterone receptor, or both	56	11.7	3	0.6	10	3.1	119	3.7
HER2 enriched	2	0.4	1	0.2	2	0.6	15	0.5
Triple negative	4	0.8	0		1	0.3	16	0.5
Missing data	2		0		1		2	

* DCIS denotes ductal carcinoma in situ, and HER2 human epidermal growth factor receptor 2.

† If women were found to have two synchronous breast cancers, only the tumor with the highest stage was used for all the analyses.

‡ DCIS also includes comedo type, intraductal papillary carcinoma with DCIS, and Paget's disease of the nipple (noninvasive with or without DCIS).

§ For women who underwent neoadjuvant chemotherapy, lymph-node status was determined on the basis of the clinical node stage.

¶ Tumor stage is based on the pathological tumor–node–metastasis (TNM) classification for women who did not receive neoadjuvant therapy. The clinical TNM classification is presented for women with invasive cancer who received neoadjuvant therapy and for those with no available data regarding the pathological TNM stage.

|| With respect to invasive tumors only, the median tumor size was 9.5 mm (interquartile range, 6.8 to 12.0) among MRI participants with screening-detected tumors, 13.0 mm (interquartile range, 10.5 to 17.0) among MRI participants with interval cancers, 15.0 mm (interquartile range, 12.0 to 31.0) among MRI nonparticipants, and 17.0 mm (interquartile range, 12.0 to 23.0) among women in the mammography-only group. Data are not provided for women who underwent neoadjuvant chemotherapy (5 MRI participants, 4 MRI nonparticipants, and 30 women in the mammography-only group).

** Data regarding receptor status are provided only for women with invasive cancers.

higher by 18 per 1000 screenings than the rate with mammography alone; this higher detection rate was accompanied by an increase of more than 20% in the incidence of false positive results. In a trial involving 478 Chinese women between the ages of 30 and 71 years who had dense breast tissue and normal results on mammography, Chen et al.[27] observed a cancer-detection rate of 33 per 1000 screenings with MRI, which was accompanied by a false positive rate of 5.2% (with only BI-RADS scores of 4 or 5 considered to be positive). The indication for the breast examination (e.g., screening or evaluation because of symptoms) in this study is not entirely clear. In a trial involving 2120 women between the ages of 40 and 70 years that included 60% who had dense breast tissue (and approximately 20% with extremely dense breast tissue), Kuhl et al.[28] found that a first round of MRI in participants with normal findings on mammography, ultrasonography, or both resulted in a cancer-detection rate of 22.6 per 1000 screenings and was accompanied by a false positive rate of 9.7% (with a BI-RADS score of 3, 4, or 5 considered to be a positive outcome). In our trial, a cancer-detection rate of 16.5 per 1000 screenings for the first round of MRI screening and a false positive rate of 8.0% appear to be roughly in line with these findings. A direct comparison of exact numbers is difficult because of differences in populations but also because of the limited sample size of most of the other studies, which creates wide confidence intervals around estimates. The main strength of our trial is the randomized design, which allowed us to study the effect of MRI screening on the interval-cancer rate as the primary outcome. Other strengths of our trial are its embedding in a population-based screening program, the multicenter design with different centers using standardized MRI protocols from different vendors, the use of fully automated and volumetric measurements of mammographic density, and the completeness of the data collection. The women were selected solely because they had extremely dense breast tissue and not because of other risk factors. In the upcoming years, the two incident screening rounds will provide important information on the value of ongoing supplemental MRI screening as compared with a one-time-only supplemental screening. In the prevalent screening round described here, we may have detected an increased number of slow-growing cancers that were less aggressive and that had been present for a long time. This hypothesis is also indicated by the relatively large number of well-differentiated and hormone-receptor–positive cancers among the MRI participants. It is unclear how many of the cancers detected in our trial were life-threatening and what fraction, if any, represents overdiagnosis.

A limitation of our trial is that it is not large enough to look at the effect of MRI screening on breast cancer–specific or overall mortality. This outcome would require a much larger sample size and longer follow-up. The lower rate of interval cancers that we found among participants who underwent MRI is indicative of and prerequisite for an effect on mortality.[7] After that, a reduction in the number of advanced cancers would also be required to show a mortality benefit, which would require several years of follow-up.[23,24]

Thus, we are now using our results in a simulation study to evaluate the reduction in mortality and the extent of overdiagnosis, together with the effects on costs and quality of life.[29–31] The recall rate of 94.9 per 1000 screenings on MRI is a concern for potential implementation of supplemental screening. Therefore, we are now evaluating methods for minimizing false positive outcomes (e.g., by computer-aided diagnosis, radiomics, and deep-learning methods). The issue of reducing the costs of MRI screening will be addressed by validating the use of abbreviated MRI protocols[28] in this population.

We found that supplemental screening with MRI in women with extremely dense breast tissue resulted in the diagnosis of significantly fewer interval cancers than the use of mammography alone. The data from incident screening rounds and longer follow-up are needed in combination with simulation studies to assess the effect on the rate of advanced cancers and, eventually, on mortality.

ACKNOWLEDGMENTS

Supported by the University Medical Center Utrecht (project number, UMCU DENSE), the Netherlands Organization for Health Research and Development (project numbers, ZonMW- 200320002-UMCU and ZonMW Preventie 50-53125-98-014), the Dutch Cancer Society (project numbers, DCS-UU-2009-4348, UU-2014-6859, and UU-2014-7151), the Dutch Pink Ribbon–A Sister's Hope (project number, Pink Ribbon-10074), Bayer Pharmaceuticals (project number, BSP-DENSE), and Stichting Kankerpreventie Midden-West. For research purposes, Volpara Health Technologies provided Volpara Imaging Software, version 1.5, for installation on servers in the screening units. Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

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; the data managers and project support personnel for their assistance during the trial; and R.K. Stellato, M.Sc., for her assistance with the statistical analyses and F.J. Gilbert, M.D., Ph.D., for her critical review of a previous version of the manuscript.

REFERENCES

1. Boyd NF, Guo H, Martin LJ, et al. Mammographic Density and the Risk and Detection of Breast Cancer. *The New England Journal of Medicine*. 2007;356:227-236.
2. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval-and screen-detected cancers. *Journal of the National Cancer Institute*. 2000;92:1081-1087.
3. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95-103.
4. DenseBreast-Info. Is there a federal law for breast density? Published March 28, 2019. Accessed June 10, 2020. https://densebreast-info.org/is-there-a-federal-law.aspx
5. Siu AL. Screening for breast cancer: U.S. Preventive services task force recommendation statement. *Annals of Internal Medicine*. 2016;164(4):279-296.
6. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental screening for breast cancer in women with dense breasts: A systematic review for the U.S. Preventive services task force. *Annals of Internal Medicine*. 2016;164(4):268-278.
7. Irwig L, Houssami N, Armstrong B, Glasziou P. Evaluating new screening tests for breast cancer. *BMJ*. 2006;332(7543):678-679.
8. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? *Annals of Internal Medicine*. 2006;144(11):850-855.
9. Emaus MJ, Bakker MF, Peeters PHM, 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.
10. National Evaluation Team for Breast Cancer Screening. *National Evaluation of Breast Cancer Screening in the Netherlands 1990-2011/2012*.; 2014.
11. Timmers JM, den Heeten GJ, Adang EM, Otten JD, Verbeek AL, Broeders MJ. Dutch digital breast cancer screening: Implications for breast cancer care. *European Journal of Public Health*. 2012;22(6):925-929.
12. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. *IEEE Transactions on Medical Imaging*. 2006;25(3):273-282.
13. D'Orsi C, Sickles E, Medelson E, Morrison E. *ACR BI-RADS Atlas: Breast Imaging Reporting and Data System*. 5th ed. American College of Radiology; 2013.
14. Zelen M. A new design for randomized clinical trials. *The New England Journal of Medicine*. 1979;300(22):1242-1245.
15. Statistics Netherlands. Kerncijfers wijken en buurten 2014. Accessed December 19, 2016. https://www.cbs.nl/nl-nl/maatwerk/2015/48/kerncijfers-wijken-en-buurten-2014
16. 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. *Clinical Radiology*. 2018;73(8):759.e1-759.e9.
17. D'Orsi C, Mendelson E, Ikeda D. *Breast Imaging Reporting and Data System: ACR BI-RADS Breast Imaging Atlas*. American College of Radiology; 2003.
18. Ikeda D, Hylton N, Kuhl CK. *BI-RADS: Magnetic Resonance Imaging*. American College of Radiology; 2003.

19. Cuzick J, Edwards R, Segnan N. Adjusting for Non-Compliance and Contamination in Randomized Clinical Trials. *Statistics in medicine*. 1997;16(9):1017-1029.
20. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*. 1996;91(434):444.
21. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: The role of instrumental variable methods. *Statistical Methods in Medical Research*. 2005;14(4):369-395.
22. Stuart EA, Perry DF, Le HN, Ialongo NS. Estimating intervention effects of prevention programs: Accounting for noncompliance. *Prevention Science*. 2008;9(4):288-298.
23. Foca F, Mancini S, Bucchi L, et al. Decreasing incidence of late-stage breast cancer after the introduction of organized mammography screening in Italy. *Cancer*. 2013;119(11):2022-2028.
24. Tabár L, Yen AMF, Wu WYY, et al. Insights from the breast cancer screening trials: How screening affects the natural history of breast cancer and implications for evaluating service screening programs. *Breast Journal*. 2015;21(1):13-20.
25. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): A randomised controlled trial. *The Lancet*. 2016;387(10016):341-348.
26. 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: The Journal of the American Medical Association*. 2012;307(13):1394-1404.
27. Chen SQ, Huang M, Shen YY, Liu CL, Xu CX. Application of abbreviated protocol of magnetic resonance imaging for breast cancer screening in dense breast tissue. *Academic Radiology*. 2017;24(3):316-320.
28. 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):361-370.
29. Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Computer Methods and Programs in Biomedicine*. 1985;20(1):79-93.
30. Mandelblatt JS, Stout NK, Schechter CB, et al. Collaborative modeling of the benefits and harms associated with different U.S. breast cancer screening strategies. *Annals of Internal Medicine*. 2016;164(4):215.
31. 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.

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, 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, R Mus MD, N Karssemeijer PhD.

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.

SUPPLEMENTARY DATA

Interactive supplementary data on the Breast Cancer MRI protocol used in the DENSE Trial is available as an accompanying mobile app with this publication: https://itunes.apple.com/app/dense-trial/id1452898975.

VOLPARA BREAST DENSITY CLASSIFICATION

Volpara density grades range from 1-4 and mimic the four-point BI-RADS ACR breast density[12] categories which range from almost entirely fatty to extremely dense.

Category	
a	Almost entirely fatty breasts
b	Scattered areas of fibroglandular density
c	Heterogeneously dense breasts
d	Extremely dense breasts

BI-RADS ASSESSMENT CATEGORIES

BI-RADS assessment categories is a seven-point system where higher numbers reflect increasing cancer risk.[13] Category 6* (known biopsy-proven malignancy) was not used for the screening mammography and screening MRIs in the DENSE trial

Category	
0	Incomplete examination
1	Negative
2	Benign
3	Probably benign
4	Suspicious
5	Highly suggestive of malignancy
6*	Known biopsy proven malignancy

SUPPLEMENTARY FIGURES AND TABLES

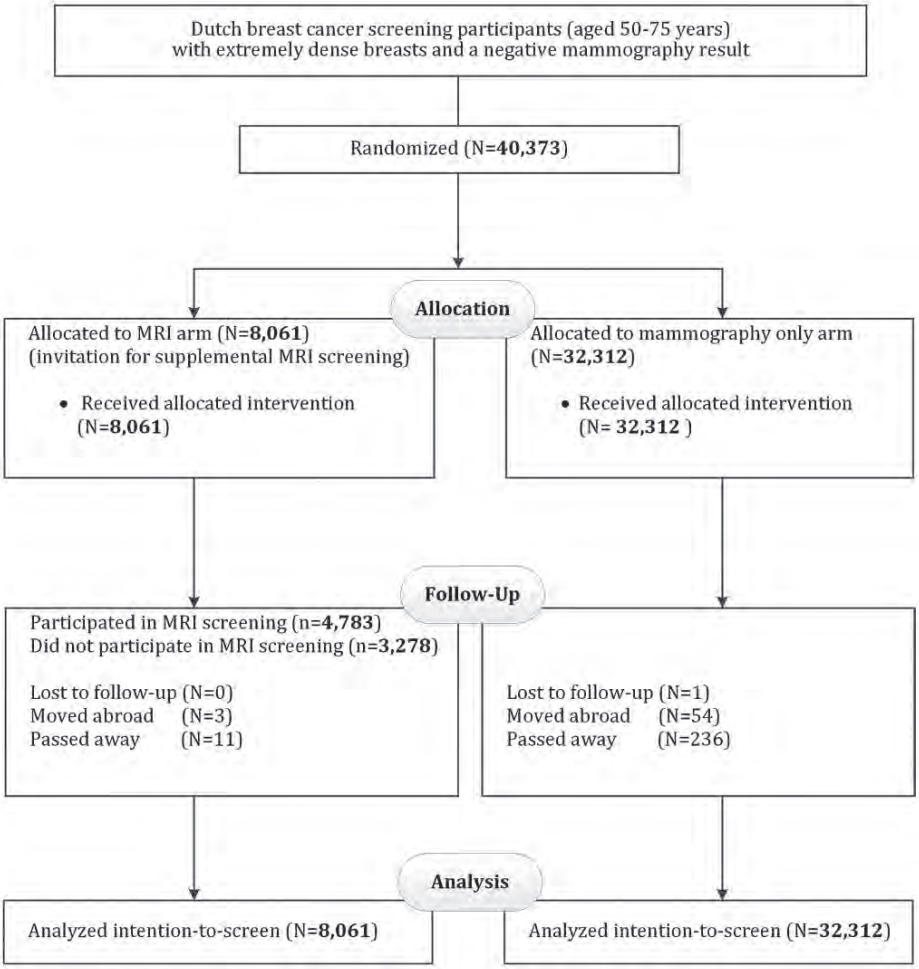


Figure S1. Consort diagram of the DENSE trial.

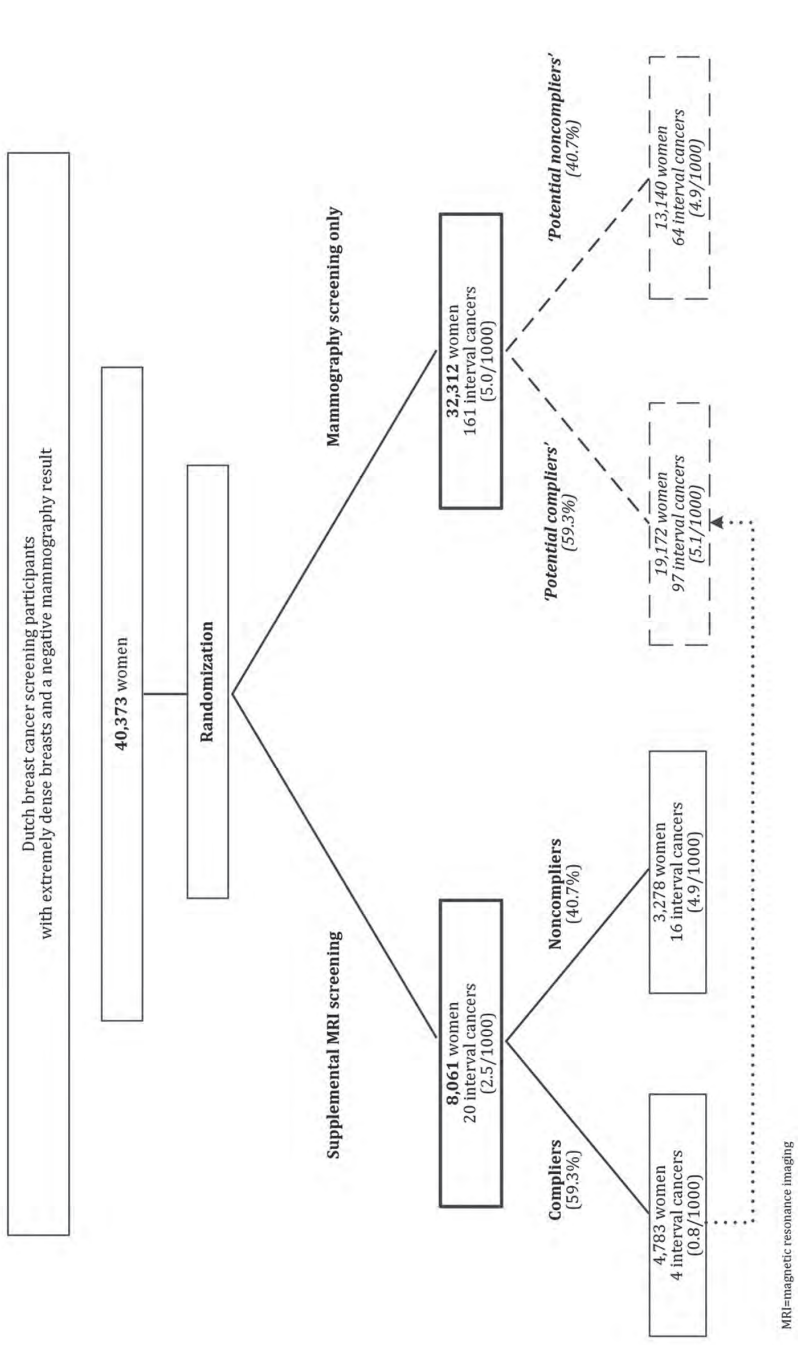


Figure S2. Flowchart of the CACE analysis (Cuzick model).
Flowchart showing the effect of additional MRI screening on the interval cancer rate adjusted for nonparticipation adapted from Cuzick et al.[19]

2

FORMULAS USED FOR CACE ANALYSIS

We used the method described in Section 3 ("The simplest case – binary outcome and no contamination") of Cuzick et al.[19] Using the notation in Cuzick et al., the CACE is $p_{11} - p^*_{11}$, where p_{11} = Pr (cancer | received MRI) and p^*_{11} is a hypothetical quantity equal to Pr (cancer | received MRI) under the scenario that receiving MRI is completely ineffective. As shown in Cuzick et al., this corresponds to the cancer proportion among the subpopulation of women who would undergo MRI if offered, but who did not in fact receive MRI.

Cuzick et al. show that $p^*_{11} = \frac{p_0 - (1-\alpha)\cdot p_{10}}{\alpha}$, where α is the proportion of women who would accept MRI screening if offered and were p_0 and p_{10} are defined as:
p_0 = Pr (cancer | randomized to mammography-only arm)
p_{10} = Pr (cancer |randomized to MRI arm, but did not receive MRI).

Owing to randomization, we can estimate each of these quantities from observed data. Specifically, the estimate of p_{11} is the observed proportion of cancer cases among those who received MRI in the MRI arm; the estimate of p_0 is the proportion of cancer cases on the mammography-only arm; and the estimate of p_{10} is the proportion of cancer cases among those who were randomized to the MRI arm, but did not receive MRI.

These quantities can all be estimated from the original and each bootstrap sample. We took 2000 bootstrap samples of the MRI arm and mammography-only arm separately, estimated p_{11}, α, p_0, and p_{10} from those samples, and estimated the difference $p_{11} - p^*_{11}$ for each sample. The percentile confidence interval was used.

ASSUMPTIONS FOR CACE ANALYSIS

The following assumptions were used, these are derived from the article of Stuart et al.[22] and can be read in more detail in this article.

In short:

1. Outcomes of each individual are not affected by treatment assignments of any other individuals
2. Opportunity to participate was assigned randomly (proportion of always-takers should be the same in the treatment and control group)
3. Opportunity to participate induces some individuals to actually participate
4. There are no defiers (being assigned to treatment group can only increase participation; also known as 'monotonicity')
5. There is no effect of assignment for the always-takers or the never-takers (to benefit from the intervention you must actually participate in it; the effect for the individual who did not take the treatment (did not participate) is 0; also known as 'exclusion restriction')

The first 4 assumptions hold for the DENSE trial. Assumption 5 is perhaps the most complicated to achieve, since this requires that there is no effect on the primary outcome (difference in interval cancer rate) of merely inviting the women (even if they decided not to participate).

However, it could be expected that some of the nonparticipants later decided to undergo supplemental imaging (MRI, ultrasound) resulting in detection of a breast cancer. (We do not have any information on this.) By design of the DENSE trials these cancers would be calculated as interval cancers. If this would happen, this would artificially increase the interval cancer rate in the nonparticipants of the MRI arm, and by that also in the 'women who would not accept MRI screening if offered' in the 'mammography-only' arm. In the CACE analysis this will automatically lead to an underestimation of the interval cancer rate in the 'women who would accept MRI screening if offered' of the mammography-only arm (see Figure S2). Thus, if assumption 5 does not hold, the reduction of the interval cancer rate may have been underestimated, if anything.

More details and background information regarding the CACE analysis (a type of instrumental variable (IV) analysis, with the IV being randomization to MRI invitation) can be found in articles written by Angrist et al.[20], Dunn et al.[21], and Stuart et al.[22]

Table S1. Cancer detection in screening interval after negative MRI or negative mammography

	MRI arm			Mammography only arm	Rate difference/ 1000 screens (95% CI)	Rate difference/ 1000 person-years (95% CI)
	Participants (N=4,783)	Non-participants (N=3,278)	Total (N=8,061)	(N=32,312)		
Initial mammogram†	Negative	Negative	Negative	Negative		
Additional detection by MRI	16.5 (79)	NA	NA	NA		
Interval cancers, time since negative MRI/Mx						
1 year	0.2 (1)	2.4 (8)	1.1 (0.6 – 2.1)	1.9 (1.5 – 2.5) (63)	0.8 (-0.3 – 1.6)	0.8 (-0.1 – 1.7)
1 year and 8 months‡	0.8 (4)	3.7 (12)	2.0 (1.2 – 3.2)	3.9 (3.3 – 4.7) (128)	2.0 (0.6 – 3.0)	1.2 (0.4 – 1.9)
1 year and 10 months §	0.8 (4)	4.6 (15)	2.4 (1.5 – 3.7)	4.3 (3.7 – 5.1) (140)	2.0 (0.5 – 3.1)	1.1 (0.3 – 1.8)
Interval cancers in total interval before subsequent mammogram	0.8 (4)	4.9 (16)	2.5 (1.6 – 3.8)	5.0 (4.3 – 5.8) (161)	2.5 (1.0 – 3.7)	1.2 (0.4 – 1.9)
Subsequent mammogram*	2.1 (1.1 – 3.9)	7.1 (5.0 – 11.8)	4.1 (2.9 – 6.0)	6.0 (5.2 – 7.0)		

NA= not applicable; Mx= mammography; MRI= magnetic resonance imaging. Unless otherwise noted, data are rates per 1000 screens, with absolute numbers within parenthesis.

Follow-up time for participants in the intervention arm (supplemental MRI screening) was calculated from negative MRI. For women with BI-RADS 3 result for the first screening MRI, follow-up time was calculated from negative follow-up MRI.

† Negative initial mammogram was a randomization criterion.

‡ Time point chosen where 75% of the participants in the MRI arm were still before subsequent mammography examination.

§ Time point chosen where 50% of the participants in the MRI arm were still before subsequent mammography examination.

* Only women who actually underwent second screening mammography were used to calculate rate/1000 screens.

Table S2. (Serious) Adverse events among women in the MRI arm who underwent the first screening MRI

	Participants MRI arm (N=4,783)	
No. of adverse events	N	(%)
Registered during or immediately after screening MRI†		
Serious adverse events‡	**5**	**(0.1)**
Vasovagal reaction	2	
Allergic reaction to contrast agent	3	
Adverse events	**3**	**(0.1)**
Extravasation of contrast agent	2	
Subluxation shoulder	1	
Health problems reported by women in 30-day questionnaire regardless of relatedness to screening MRI§		
Serious adverse events ‡	**27**	**(0.6)**
Nervous system disorders	5	
Gastro-intestinal disorders	2	
Skin and subcutaneous tissue disorders	2	
Cardiovascular disorders	2	
Musculoskeletal disorders	7	
Ear, nose, throat and eye disorders	3	
General disorders	1	
Respiratory disorders	1	
Urological disorders	1	
Reproductive system and breast disorders/complaints	1	
Medical procedures (complications during/after biopsy procedure)*	2	
Adverse events	**1,233**	**(25.7)**
Nervous system disorders	151	
Gastro-intestinal disorders	131	
Psychosocial and psychiatric disorders	142	
Skin and subcutaneous tissue disorders	53	
Cardiovascular disorders	39	
Musculoskeletal disorders	297	
Ear, nose, throat and eye disorders	84	
Neoplasms	41	
General disorders	52	
Respiratory disorders	147	
Urological disorders	8	
Endocrine disorders	6	
Reproductive system and breast disorders/complaints	53	
Malaise (including headache, nausea)	29	

MRI = magnetic resonance imaging. Unless otherwise noted, data are numbers of women, with vertically calculated percentages in parenthesis.

† Data are numbers of (serious) adverse events that occurred during or immediately after the MRI examination.

‡ A serious adverse event has been defined as an experienced adverse event that required emergency department visit and/or (unplanned) hospital admission.

§ Data are numbers of (serious) adverse events that were reported by women in a questionnaire sent 30 days after the MRI examination. The question was: "Have you experienced one or more health problems during or in the period after the MRI examination up until now? (The health problems don't necessarily have to be related to the MRI examination.)". The events were classified by organ system. Vertically calculated percentages are in parenthesis. More than one event could be registered for each woman.

* The biopsy procedure after/during which these events occurred took place after the initial screening MRI examination.

3

Reasons for (non)participation in supplemental population-based MRI breast screening for women with extremely dense breasts

S.V. de Lange, M.F. Bakker, E.M. Monninkhof, P.H.M. Peeters,
P.K. de Koekkoek-Doll, R.M. Mann, M.J.C.M. Rutten, R.H.C. Bisschops,
J. Veltman, K.M. Duvivier, M.B.I. Lobbes, H.J. de Koning, N. Karssemeijer,
R.M. Pijnappel, W.B. Veldhuis, C.H. van Gils

Clinical Radiology. 2018; 73(8), 759.e1-759.e9.

ABSTRACT

Aim

To determine the willingness of women with extremely dense breasts to undergo breast cancer screening with magnetic resonance imaging (MRI) in a research setting, and to examine reasons for women to participate or not.

Materials and methods

Between 2011 and 2015, 8,061 women (50 to 75 years) were invited for supplemental MRI as part of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial (ClinicalTrials.gov Identifier: NCT01315015), after a negative screening mammography in the national population-based mammography screening program. Demographics of participants and nonparticipants were compared. All invitees were asked to report reasons for (non)participation. Ethical approval was obtained. Participants provided written informed consent.

Results

Of the 8,061 invitees, 66% answered that they were interested, and 59% eventually participated. Participants were on average 54-years old (interquartile range 51 to 59 years), comparable to women with extremely dense breasts in the population-based screening program (55 years). Women with higher socio-economic status (SES) were more often interested in participation than women with lower SES (68% versus 59%, $p<0.001$). The most frequently stated reasons for non-participation were "MRI-related inconveniences and/or self-reported contraindications to MRI" (27%) and "anxiety regarding the result of supplemental screening" (21%). "Expected personal health benefit" (68%) and "contribution to science" (43%) were the most frequent reasons for participation.

Conclusion

Of women invited for MRI because of extremely dense breasts, 59% participated. Common reasons for nonparticipation were "MRI-related inconveniences" and "anxiety regarding the result of supplemental screening". In case of future implementation, availability of precise evidence on benefits and harms might reduce this anxiety.

INTRODUCTION

The sensitivity of mammography, including digital mammography, is reduced in women with dense breasts.[1–3] In addition, women with extremely dense breasts have a 3 to 6 times increased risk of breast cancer compared to women with almost entirely fatty breasts, and two times compared to average women.[4–7] Around 8% of the women participating in the Dutch breast cancer screening program have extremely dense breasts.[8] This is comparable to breast cancer screening participants in other western countries.[2,9]

Thus far, 30 states in the USA have enacted breast density notification laws.[10] These laws mandate that women with mammographically dense breasts are notified about their breast density and the reduced sensitivity of mammography. In several states, women should also be informed about the possible need for supplemental screening, and in some states, insurance is mandated to cover supplemental screening.[11] Since the enactment of these laws, the use of ultrasound and magnetic resonance imaging (MRI) for supplemental screening has increased in various states.[12,13]

In 2011, the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial (ClinicalTrials.gov Identifier: NCT01315015) was initiated to investigate the additional value of MRI examination to detect breast cancer in women with extremely dense breasts participating in the Dutch breast cancer screening program.[14] The first screening round of the trial was completed in 2015, and results of the primary outcome of the trial, the occurrence of interval cancers in the supplemental MRI arm compared to the mammography only arm, are expected in 2018.

Little is known about the willingness of women, who are selected on the basis of breast density only, to undergo a supplemental screening MRI, or about reasons for these women to accept or refuse this screening. To explore a possible future implementation of this technique, it is important to know the acceptability of supplemental screening with MRI.

In a study among women with elevated breast cancer risk, based on breast density (heterogeneously or extremely dense) and at least one other breast cancer risk factor, 58% of eligible women were willing to participate and 51% actually participated.[15] The aim of the current study was to determine the willingness of women with extremely dense breasts and a negative mammography screening examination, to undergo additional breast cancer screening with MRI in the DENSE trial. The influence of age, travel distance to the MRI center, socioeconomic status (SES), and degree of urbanization on participation in the study was investigated, including what reasons women reported to accept or refuse this screening option.

MATERIALS AND METHODS

Design and study population

The DENSE trial is a multicenter randomized controlled trial investigating the additional value of full breast MRI screening in women with extremely dense breasts. It has a single-consent prerandomization design and is carried out in the Dutch population based-screening program where all participants between 50 and 75 years old are screened biennially with digital mammography. Breast density is not routinely assessed, and consequently, screening participants are not informed about their mammographic density.

For the purpose of the trial, software for automatic density evaluation was installed in the participating screening centers (Volpara Imaging Software version 1.5; VolparaSolutions, Wellington, New Zealand). Women with extremely dense breasts (Volpara density grade 4), and a negative mammography result after double reading (Breast Imaging Reporting and Data System classification 1 or 2) were eligible for randomization. After randomization, only those women randomized to the intervention group (supplemental MRI examination) were informed and invited for the trial. Further details of the design have been described elsewhere.[14]

The trial has been approved by the Dutch Minister of Health, Welfare and Sport (2011/19 WBO, the Hague, the Netherlands). Ethical approval was obtained on November 11, 2011.

Between December 2011 and November 2015, 40,372 women were randomized in a 1:4 ratio to the intervention (N=8,061) or control arm (N=32,311) of the trial. All participants provided written informed consent before the MRI examination.

Invitation procedure

An invitation letter and an extensive information brochure (in Dutch) were sent by mail by the regional screening organizations, to the women randomized to the intervention arm. The invited women were informed about their breast density, breast density-related increased breast cancer risk, and reduced sensitivity of mammography in dense breasts. Information was provided on the possible advantage (early detection and treatment of breast cancer) and disadvantages (e.g., false positive result and possible side effects of MRI) of participation in the trial. Invitations were sent and participation took place before the recent concerns regarding gadolinium retention.[16,17]

Invited women could indicate (online or by regular mail) whether they were interested in participation or not. If women did not respond, the regional screening organization sent a reminder letter 3 weeks after initial invitation. Subsequently, all interested women were called by the research team, to provide more detailed information, check

eligibility, to obtain informed consent, and to schedule an appointment for the MRI examination.

(Socio)demographic information of invitees

For the women not willing to participate in the trial, only information on breast density, the result of the screening mammography, and four-digit postal codes are available. Additional characteristics are available for women who were interested in participation. The postal codes were used to calculate the travel distance between residential address and the allocated MRI center, to determine SES, and the degree of urbanization. SES was approximated, using a neighborhood social status score, provided by the Netherlands Institute for Social Research.[18] This score is based on average income, percentage of people with a low income, percentage of people with a low education level, and the percentage of unemployment in a neighborhood.[18] Available scores for 2014 were used. Invitees were categorized into quartiles based on SES, from low to high, with cut-off points based on the distribution of the score for the Netherlands in 2014. The degree of urbanization, provided by Statistics Netherlands, is based on the number of addresses per square kilometer (km^2).[19] Urbanization levels for 2014 were used. In case more than one level of urbanization was available for a postal code, the highest level of urbanization was used.

Classification of reasons for (non)participation

Invitees were asked to state their reason(s) for (non)participation in a blank space at the online registration form or on the reply card. All stated reasons were used for analyses. Reasons were coded and categorized by at least two authors. In the case of disagreement, consensus was reached by discussion with a third author. For non-participation, six categories of reasons (Table 3) were created during coding. For participation, four categories were created (Table 4). For women with interest in participation who did not complete the MRI examination, the radiologist or radiographer registered the reason for incomplete MRI in an online database.

Statistical analysis

All invitees were categorized according to “interested in participation” versus “not interested in participation”, the latter group contains both invitees who declined the invitation and invitees who did not respond (nonrespondents). Subsequently, those interested in participation were further categorized into: “actual participants” and those who were “interested, but eventually not participating”. The median age of participants was compared to the median age of all women with extremely dense breasts participating in the same regions of the Dutch breast cancer screening program in 2013.

Differences in participation for SES quartiles and per level of urbanization were analyzed using Pearson’s X^2 test. The median travel distance for women who were

interested in participation was compared to those for women who were not interested and analyzed using the Mann–Whitney *U*-test. Reported reasons for (non)participation were summarized using descriptive statistics. $p<0.05$ was considered statistically significant. All statistical analyses were performed using R, version 3.2.2.

RESULTS

Of the invited women, 5,276 (65.5%) were interested in participation, 1,872 (23.2%) declined the invitation and 913 (11.3%) were nonrespondents (Figure 1). Ultimately, 4,783 women participated in the trial and completed the MRI examination, resulting in an actual participation rate of 59.3%. Table 1 lists the (socio)demographics of all invitees. The median age of women interested in participation and also the median age of women who actually participated was 54 years (interquartile range [IQR], 51 to 59; see Table 2), compared to 55 years (IQR, 51 to 61) for women with extremely dense breasts participating in the same regions of the Dutch breast cancer screening program in 2013 (subgroup of previously published data).[8]

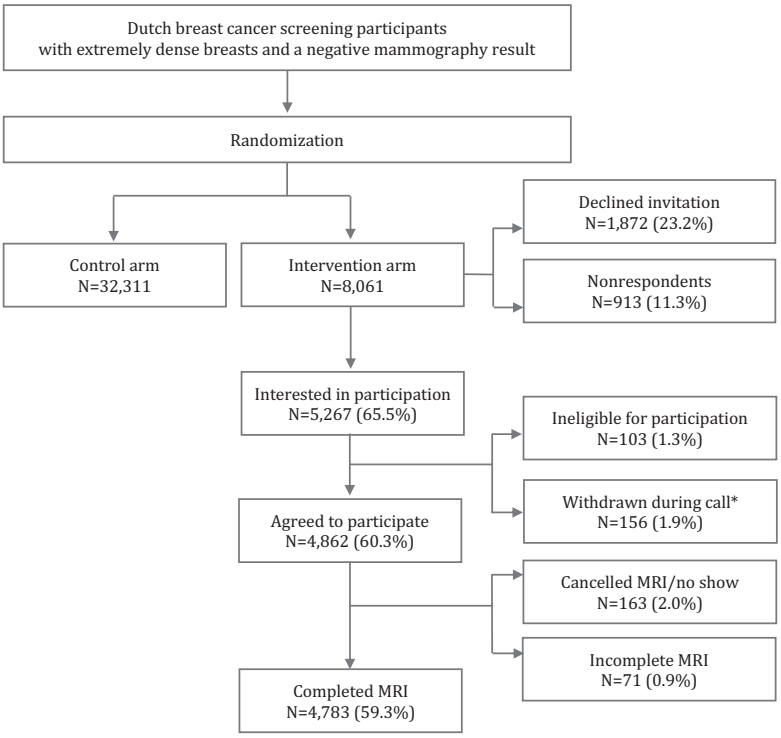


Figure 1. Flowchart of participation in the DENSE trial.
*These women were interested in participation, but withdrew during the inclusion phone call.

Table 1. Characteristics of invitees

	All invitees		Interested in participation		Not interested in participation*	
Number of women (%)	8,061	(100.0)	5,276	(65.5)	2,785	(34.5)
Age at randomization, years†	49-76‡		54	(51-59)	NA	
Missing§	NA		14	(0.3)	NA	
Travel distance, km†¶	24.8	(15.7-32.4)	24.5	(15.4-32.3)	25.5	(16.1-32.7)
SES						
Q4 (highest SES)	2,942	(100.0)	2,010	(68.3)	567	(31.7)
Q3	1,919	(100.0)	1,259	(65.6)	619	(34.4)
Q2	1,808	(100.0)	1,189	(65.8)	660	(34.2)
Q1 (lowest SES)	1,372	(100.0)	805	(58.7)	932	(41.3)
Missing§	20	(0.2)	13	(0.2)	7	(0.3)
Urbanization level (addresses/km2)**						
Extremely	1,574	(100.0)	980	(62.3)	594	(37.7)
Strongly	2,542	(100.0)	1,619	(63.7)	923	(36.3)
Moderately	1,545	(100.0)	1,062	(68.7)	483	(31.3)
Hardly	1,181	(100.0)	782	(66.2)	399	(33.8)
Not	1,113	(100.0)	766	(68.8)	347	(31.2)
Missing§	106	(1.3)	67	(1.3)	39	(1.4)

Unless otherwise noted, data are numbers of women, with horizontally calculated percentages in parenthesis. Total of percentages may not equal 100% due to rounding. SES was available for postal codes in neighborhoods of >100 households.
NA, not applicable.
* Not interested in participation: this category consists of 1,872 women who declined the invitation and 913 nonrespondents.
† Median (interquartile range).
‡ Numbers are ranges, since only birthdates of women who indicate interest in participation are known.
§ Numbers in parenthesis are percentages of column total.
¶ p=0.041 for the difference between the groups of women who are interested and those who are not.
** p<0.001 for the difference between the groups of women who are interested and those who are not.

Even though the inclusion of participants was organized centrally and not by the participating MRI centers located throughout the Netherlands, participation differed per MRI center, with rates of invitees interested in participation ranging from 63.8% to 66.6%, and actual participation ranging from 56.1% to 61.9% (see Electronic Supplementary Material Table S1).

Median travel distance between residence and MRI center was 24.5 km (IQR, 15.4 to 32.3) for women who were interested in participation, compared to 25.5 km (IQR, 16.1 to 32.7) for women who were not interested in participation (p=0.04).

Of the women with a high SES, 68.3% were interested in participation, compared to 58.7% of women with low SES ($p<0.001$). Eventually 62.1% of the women with a high SES, compared to 52.2% of the women with a low SES, actually participated (see Table 2). Interest in participation was higher among women living in a low urbanized area, i.e., 68.8%, compared to 62.3% of women living in extremely urbanized area ($p<0.001$).

After stating an interest in participation, a total of 493 women (6.1% of invitees) eventually did not participate (see Table 2). Of these women, 156 (1.9% of invitees) withdrew from participating during the inclusion phone call, and 163 women (2% of invitees) cancelled or did not show up at the MRI appointment. The median age of these women was 57 years (IQR, 52 to 64.5) and 53 years (IQR, 51 to 60), respectively.

Reason for withdrawal or cancelling was unknown for 67.3% and 25.2%, respectively. The most frequently stated reason for cancelling the MRI examination, was due to other health problems (28/163 women; see Table 3).

In total, 103 women (1.3% of invitees) with an interest in participation appeared ineligible during the inclusion phone call. The median age of these women was 55 years (IQR, 52 to 62). For 57 women, an "MRI-related inconvenience and/or self-reported contraindication" was registered as reason for ineligibility with intracorporeal metal as most frequently reported reason (36/57 women). Twenty-one women were ineligible because scheduling an MRI examination appeared impossible, i.e., 6 months after mammography still no MRI examination was scheduled (Table 3).

Seventy-one women (0.9% of invitees, 1.3% of the women with interest in participation) did not complete the MRI examination. For 70 of them an "MRI-related inconvenience and/or contraindication" was registered as reason for incomplete MRI, with claustrophobia as most frequently reported reason (N=41). Three women could not complete the MRI examination because of side effects of the contrast agent (e.g., nausea; Table 3). All 71 women refused to repeat the MRI examination or were ineligible to undergo another MRI examination.

Table 2. Characteristics of women interested in participation

	Actual participants	Interested, but eventually not participating			
		Ineligible	Withdrawn during call	Cancelled MRI / no show	Incomplete MRI
Number of women (%)	4,783 (59.3)	103 (1.3)	156 (1.9)	163 (2.0)	71 (0.9)
Age at randomization, years*	54 (51-59)	55.0 (52-62)	57 (52.0-64.5)	53 (51.0-60.0)	52.0 (50-56)
Missing†	0 (0.0)	13 (12.6)	1 (0.6)	0 (0.0)	0 (0.0)
Travel distance, km*	24.5 (15.1-32.2)	24.7 (16.9-32.3)	24.8 (17.0-32.4)	25.6 (18.4-33.2)	23.9 (16.9-30.3)
SES					
Q4 (highest SES)	1,828 (62.1)	36 (1.2)	66 (2.2)	56 (1.9)	24 (0.8)
Q3	1,144 (59.6)	26 (1.4)	31 (1.6)	36 (1.9)	22 (1.1)
Q2	1,083 (59.9)	17 (0.9)	28 (1.5)	45 (2.5)	16 (0.9)
Q1 (lowest SES)	716 (52.2)	24 (1.7)	30 (2.2)	26 (1.9)	9 (0.7)
Missing†	12 (0.3)	0 (0.0)	1 (0.6)	0 (0.0)	0 (0.0)
Urbanization level (addresses/km2)					
Extremely	871 (55.3)	23 (1.5)	37 (2.4)	38 (2.4)	11 (0.7)
Strongly	1,462 (57.5)	36 (1.4)	44 (1.7)	54 (2.1)	23 (0.9)
Moderately	961 (62.2)	19 (1.2)	34 (2.2)	28 (1.8)	20 (1.3)
Hardly	714 (60.5)	13 (1.1)	22 (1.9)	21 (1.8)	12 (1.0)
Not	713 (64.1)	10 (0.9)	17 (1.5)	21 (1.9)	5 (0.4)
Missing†	62 (1.3)	2 (1.9)	2 (1.3)	1 (0.6)	0 (0.0)

Unless otherwise noted, data are numbers of women, with percentages calculated of total number of invitees in that category in parenthesis (see Table 1). Total of percentages may not equal 65.5% due to rounding.
SES was available for postal codes in neighborhoods of >100 households.
* Median (interquartile range).
† Numbers in parenthesis are percentages of column total.

Table 3. Reasons for nonparticipation stated by women who declined the invitation and interested women who did not participated

	Declined invitation	Interested, but eventually not participating			
		Ineligible	Withdrawn during call	Cancelled MRI / no show	Incomplete MRI
Number of women	1,872	103	156	163	71
MRI-related inconveniences and/or self-reported contraindications	**503 (26.9)**	**57 (55.3)**	**14 (9.0)**	**24 (14.7)**	**70 (98.6)**
Claustrophobia	219 (11.7)	1 (1.0)	2 (1.3)	10	41 (57.7)
Refusing/fear for needles	22 (1.2)	0 (0.0)	0 (0.0)	0	1 (1.4)
Refusing MRI	74 (4.0)	NA	0 (0.0)	0	0 (0.0)
Refusing contrast agent or (self-reported) contraindications for contrast agent	133 (7.1)	18 (17.5)	9 (5.7)	10	11 (15.5)*
Physical inability to adopt to the right positioning for MRI	80 (4.3)	2 (1.9)	2 (1.3)	2	10 (14.1)
Other (self-reported) contraindication for MRI (e.g., intracorporeal metal)	61 (3.3)	36 (35.0)	2 (1.3)	3	2 (2.8)†
Inadequate MRI or MRI did not function/renal function test failed	NA	NA	NA	NA	6 (8.5)
Anxiety regarding the result of supplemental screening	**390 (20.8)**	**NA**	**5 (3.2)**	**16**	**0 (0.0)**
Emotional burden is too high	288 (15.4)	NA	3 (1.9)	11	0 (0.0)
Concerns about false-positives results or overdiagnosis	37 (2.0)	NA	1 (0.6)	3	NA
Aversion to hospitals or refusing any medical procedures	84 (4.5)	NA	2 (1.3)	3	NA
Personal reasons	**392 (20.9)**	**7 (6.8)**	**11 (7.1)**	**42**	**0 (0.0)**
Other disease/health concerns	164 (8.8)	0 (0.0)	10 (6.4)	28	0 (0.0)
Low estimate of own risk (“The regular screening program is sufficient”)	130 (6.9)	NA	0 (0.0)	0	NA
Inability to oversee consequences/aims of the study/language barrier	10 (0.5)	7 (6.8)	NA	NA	NA
Personal reasons without further explanation	54 (2.9)	NA	0 (0.0)	13	NA
Age or “I’m too old”	52 (2.8)	NA	1 (0.6)	1	NA

Table 3. Continued.

	Declined invitation	Interested, but eventually not participating			
		Ineligible	Withdrawn during call	Cancelled MRI / no show	Incomplete MRI
Practical reasons	**267 (14.3)**	**34 (33.0)**	**9 (5.7)**	**29**	**NA**
Time constraints	130 (6.9)	0 (0.0)	4 (2.6)	14	NA
Travel-related inconveniences	92 (4.9)	0 (0.0)	1 (0.6)	1	NA
Other priorities	45 (2.4)	NA	3 (1.9)	3	NA
Financial concerns (costs possible additional interventions/diagnostics)	14 (0.7)	NA	1 (0.6)	1	NA
Impossible to schedule MRI appointment/no show	NA	21 (20.4)	NA	10	NA
Incomplete registration	NA	13 (12.6)	NA	NA	NA
Burden too high (without further explanation)	219 (11.7)	NA	13 (8.3)	9	0 (0.0)
Reasons related to surveillance	**88 (4.7)**	**1 (1.0)**	**3 (1.9)**	**4**	**NA**
Already under active surveillance/referred before MRI	25 (1.3)	0 (0.0)	0 (0.0)	2	NA
Made an appointment with GP for additional screening/discontinued screening	8 (0.4)	NA	2 (1.3)	1	NA
Participation was discouraged by physician/others	18 (1.0)	1 (1.0)	1 (0.6)	1	NA
Aversion to extra prevention	20 (1.1)	NA	NA	0	NA
Do not want to participate in research/in a trial	18 (1.0)	NA	NA	0	NA
Reason could not be classified	**21 (1.1)**	**2 (1.9)**	**2 (1.3)**	**7**	**1 (1.4)**
No reason mentioned (empty or no interest without specification)	**394 (21.0)**	**4 (3.9)**	**105 (67.3)**	**41 (25.2)**	**0 (0.0)**

Data are number of women with vertically calculated percentages in parenthesis. Percentages do not add up to 100% because women could state more than 1 reason. Main categories (in bold) are not the sum of subcategories, because women could state several subcategories within one main category.

NA, not applicable; GP, general practitioner.

* The MRI was cancelled for following reasons: renal failure (n=3) determined before the MRI, inadequate administration of contrast agent (n=2) or intravenous needle (n=2), side effects of contrast agent (e.g. nausea) (n=3) and last-minute cancellation because of known allergies (n=1).

† Intracorporeal metal was registered by the MRI radiographer before the MRI, subsequently the MRI was cancelled.

Of the women who declined the invitation, 394 (21%) did not specify their reason(s). The remaining 1,478 women stated 2,018 reasons for non-participation (see Table 3). Most frequently stated reasons were "MRI-related inconveniences and/or self-reported contraindications", stated by 503 women (26.9%). Among them claustrophobia and refusing the contrast agent or self-reported contraindication to the contrast agent were the most frequently stated MRI-related reasons (stated by 219/503 women and 133/503 women, respectively). Due to the design of the trial, it could not be verified whether reported contraindications were true contraindications. Personal reasons (e.g., health problems and low estimate of own risk), stated by 392 women (20.9%), and anxiety regarding the result of supplemental screening, stated by 390 women (20.8%), were the most frequently stated reasons that were not specifically related to the MRI but to supplemental screening in general. The majority of the women who stated anxiety regarding the result of supplemental screening, stated that the emotional burden of participation was too high (288/390 women). Reported reasons for non-participation were similar for the SES groups (data not shown).

Of the women interested in participation, 4,853 (92%) stated one or more reasons for participation (7,061 reasons in total; see Table 4). The most frequently stated reasons were "expected personal health benefit" stated by 3,607 women (68.4%) and "contribution to science", stated by 2,245 women (42.6%). A positive family history for (breast) cancer was reported by 443 women (8.4%) as reason for participation. In case no further information was stated, it remained unclear whether these women also wanted to participate for expected personal health benefit or to contribute to science. The percentage of women stating "expected personal health benefit" decreased with increasing age, from 71.6% of women aged <55 years to 42.9% of all women aged ≥70 years. The percentage of women stating "contribution to science" was higher for those with higher SES, from 38.6% for Q1 (lowest SES) to 45.8% for Q4 (highest SES; see Electronic Supplementary Material Table S2).

DISCUSSION

For the DENSE trial, women were selected on the basis of breast density, without using other risk factors. Of the women randomized to the intervention arm and thereby invited for supplemental screening with MRI, 59.3% participated. The most frequently stated reasons for non-participation were "MRI-related inconveniences (e.g., claustrophobia, refusing contrast agent) and/or self-reported contraindications" (27%) and "anxiety regarding the result of supplemental screening" (21%). Women with interest in participation stated "expected personal health benefit" (68%) and "contribution to science" (43%) most frequently as reasons for participation.

Previously conducted MRI screening studies in women with genetic or familial predisposition generally showed higher participation rates varying from 58% to almost 90%.[20,21] This is expected, as women with very high risk are likely to be willing to undergo more extensive screening procedures.

MRI screening has been studied before in women with an elevated breast cancer risk without genetic predisposition,[22,23] and in low-risk women.[24] One of these studies (ACRIN6666 trial) reported on participation rates and reasons for non-participation, and showed a participation rate of 51.6%.[15] Although reasons for non-participation in the present study and ACRIN6666 trial cannot be compared directly because of different methodology, there are some notable differences. First, financial concerns as reason for non-participation was stated by 12.1% in the ACRIN6666 trial, compared to only 0.7% in the DENSE trial.[15] In the ACRIN6666 trial, in principle, insurance was billed for the MRI, but not all insurers covered these costs. In the DENSE trial, the costs of the MRI were provided for by DENSE. Only in case women were referred for further assessment, their insurer was billed for the expenses. Deductibles of health insurances are increasing in the Netherlands; hence financial concerns may become a larger barrier for screening participation in the future. Second, claustrophobia was stated by 11.7% in the DENSE trial, which is lower than the 25.4% reported in the ACRIN6666 trial.[15] On the contrary, other MRI-related reasons (e.g., contrast-related reasons) were stated more frequently in the DENSE trial. In total, MRI-related reasons were reported by 26.9% of the women in the present study who declined participation. In case of future implementation of MRI screening, these barriers for participation will probably remain. Other screening techniques, e.g., ultrasound and tomography, might be offered to these women; however, sensitivity of these techniques is much lower than that of MRI.[25,26]

Remarkably, "anxiety regarding the result of supplemental screening" was another frequently stated reason (21%) in this population of women who are (regular) participants of the population-based mammographic screening program. Mostly, these women mentioned anxiety to be diagnosed with breast cancer or anxiety while waiting for results ("emotional burden is too high"). Other screening studies have mentioned anxiety as well.[27,28] A study conducted within a British colorectal cancer screening trial (UK FS Trial) showed that cancer fear can both facilitate and deter participation in cancer screening.[27] In addition, a study on the effects of education about screening mammography, showed that >40% of the women reported to experience anxiety before, at, or after mammographic screening. This anxiety was attributed to not knowing the result of mammography (stated by 56.4%), general uncertainty (12.7%), and waiting for results (9.1%).[28]

The possible side effects of participation, as mentioned in the patient information brochure, could have been expected to influence the women's decision to participate in

the trial. It was mentioned that false-positive results could be expected in 15% of the participants. This is almost tenfold higher than in the regular Dutch mammographic screening (1.8%).[29] Of the women who declined the invitation, 2% actually explicitly stated concerns about false-positive results or overdiagnosis.

Thus far, research on acceptability of MRI by women without genetic or familial predisposition is scarce, but important to have when considering implementing supplemental MRI screening for women with dense breasts. The current study had a large sample size. Reasons for (non)participation were collect without prejudice by offering invitees a blank space to state their reason(s).

In the DENSE trial a prerandomization design was used to prevent contamination of the control group. As a result, all nonparticipants in the intervention group and the control group did not provide informed consent. Therefore, comparisons between participants and nonparticipants were limited, because only four-digit postal codes were available for these women.

In summary, when offering supplemental MRI screening to women with extremely dense breasts and a negative mammography in a trial, the participation rate is close to 60%. If effectiveness is proven and supplemental screening in women with extremely dense breasts is embedded in a regular screening program, this proportion could be higher. As “anxiety regarding the result of supplemental screening” is an important reason for non-participation, offering both techniques on the same day, delivering one combined evaluation result, and providing clear and evidence-based information of benefits and harms, might further reduce barriers to participation.

DECLARATION OF INTEREST

N.K. is a cofounder and shareholder of Volpara Health Technologies Solutions and Qview Medical (Los Altos, CA, USA) and CEO of ScreenPoint Medical (Nijmegen, NL). R.M.M. disclosed grants and non-financial support from Siemens Healthcare, grants and non-financial support from Seno Medical, grants from Identification Solutions, grants from Bayer Healthcare, grants from Elswood, and grants from Medtronic. The other authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

ACKNOWLEDGMENTS

The authors acknowledge the study participants for their contributions. The authors thank the regional screening organizations, Volpara Solutions, the Dutch Expert Centre

for Screening, and the National Institute for Public Health and the Environment for their advice and in-kind contributions for organizing the breast density assessments and the invitation of potential participants. The DENSE trial is financially supported by the University Medical Center Utrecht (Project number: UMCU DENSE), the Netherlands Organization for Health Research and Development (ZonMw, Project numbers: ZONMW-200320002-UMCU and ZonMW Preventie 50-53125-98-014), the Dutch Cancer Society (KWF Kankerbestrijding, Project numbers: DCS-UU-2009-4348, UU-2014-6859 and UU2014-7151), the Dutch Pink Ribbon/A Sister's Hope (Project number: Pink Ribbon-10074), Bayer AG Pharmaceuticals, Radiology (Project number: BSP-DENSE), and Stichting Kankerpreventie Midden-West. For research purposes, Volpara Solutions (Wellington, New Zealand) provided Volpara Imaging Software 1.5 for installation on servers in the screening units of the Dutch screening program.

REFERENCES

1. Kerlikowske K, Zhu W, Tosteson ANA, et al. Identifying women with dense breasts at high risk for interval cancer a cohort study. *Annals of Internal Medicine*. 2015;162(10):673-681.
2. Weigel S, Heindel W, Heidrich J, Hense HW, Heidinger O. Digital mammography screening: sensitivity of the programme dependent on breast density. *European Radiology*. 2017;27(7):2744-2751.
3. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95-103.
4. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiology Biomarkers and Prevention*. 2006;15(6):1159-1169.
5. Boyd NF, Guo H, Martin LJ, et al. Mammographic Density and the Risk and Detection of Breast Cancer. *The New England Journal of Medicine*. 2007;356:227-236.
6. Pettersson A, Graff RE, Ursin G, et al. Mammographic density phenotypes and risk of breast cancer: A meta-analysis. *Journal of the National Cancer Institute*. 2014;106(5).
7. Wanders JOP, Holland K, Karssemeijer N, et al. The effect of volumetric breast density on the risk of screen-detected and interval breast cancers: A cohort study. *Breast Cancer Research*. 2017;19(1):67.
8. van der Waal D, Emaus MJ, Bakker MF, et al. Geographic variation in volumetric breast density between screening regions in the Netherlands. *European Radiology*. 2015;25(11):3328-3337.
9. Sprague BL, Gangnon RE, Burt V, et al. Prevalence of mammographically dense breasts in the United States. *Journal of the National Cancer Institute*. 2014;106.
10. DenseBreast-Info. Legislation and regulations - what is required. Published February 1, 2018. Accessed February 1, 2018. http://densebreastinfo.org/
11. Dehkordy SF, Carlos RC. Dense breast legislation in the United States: State of the States. *Journal of the American College of Radiology*. 2016;13(11):R53-R57.
12. Chau SL, Alabaster A, Luikart K, Brenman LM, Habel LA. The effect of California's breast density notification legislation on breast cancer screening. *Journal of Primary Care and Community Health*. 2017;8(2):55-62.
13. Sanders LM, King AB, Goodman KS. Impact of the New Jersey breast density law on imaging and intervention volumes and breast cancer diagnosis. *Journal of the American College of Radiology*. 2016;13(10):1189-1194.
14. Emaus MJ, Bakker MF, Peeters PHM, 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.
15. 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.
16. Kanda T, Fukusato T, Matsuda M, et al. Gadolinium-based contrast agent accumulates in the brain even in subjects without severe renal dysfunction: Evaluation of autopsy brain specimens with inductively coupled plasma mass spectroscopy. *Radiology*. 2015;276(1):228-232.

17. McDonald RJ, McDonald JS, Kallmes DF, et al. Intracranial gadolinium deposition after contrast-enhanced MR imaging. *Radiology*. 2015;275(3):772-782.
18. Netherlands Institute for Social Research. Statusscores. Published 2014. Accessed March 3, 2015. https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores
19. Statistics Netherlands. Kerncijfers wijken en buurten 2014. Accessed December 19, 2016. https://www.cbs.nl/nl-nl/maatwerk/2015/48/kerncijfers-wijken-en-buurten-2014
20. Kriege M, Brekelmans CTM, Boetes C, et al. Differences between first and subsequent rounds of the MRISC breast cancer screening program for women with a familial or genetic predisposition. *Cancer*. 2006;106(11):2318-2326.
21. Evans DGR, Lalloo F, Shenton A, Boggis C, Howell A. Uptake of screening and prevention in women at very high risk of breast cancer. *The Lancet*. 2001;358(9285):889-890.
22. 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: The Journal of the American Medical Association*. 2012;307(13):1394-1404.
23. 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. *Journal of Clinical Oncology*. 2014;32(22):2304-2310.
24. 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):361-370.
25. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental screening for breast cancer in women with dense breasts: A systematic review for the U.S. Preventive services task force. *Annals of Internal Medicine*. 2016;164(4):268-278.
26. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): A randomised controlled trial. *The Lancet*. 2016;387(10016):341-348.
27. Vrinten C, Waller J, von Wagner C, Wardle J. Cancer fear: facilitator and deterrent to participation in colorectal cancer screening. *Cancer Epidemiology Biomarkers and Prevention*. 2015;24(2):400-405.
28. Lee J, Hardesty LA, Kunzler NM, Rosenkrantz AB. Direct interactive public education by breast radiologists about screening mammography: impact on anxiety and empowerment. *Journal of the American College of Radiology*. 2016;13(11):R89-R97.
29. National Evaluation Team for Breast cancer screening in the Netherlands (NETB). *National Evaluation of Breast Cancer Screening in the Netherlands, Monitor 2014*.; 2016.

SUPPLEMENTARY TABLES

Table S1. Participation by MRI center

	All invitees	Interested in participation					Nonparticipants	
MRI center		Actual participants	Ineligible	Withdrawn during call	Cancelled MRI / no show	Incomplete MRI	Declined invitation	Non-respondents
1	2,693 (100.0)	1,615 (60.0)	38 (1.4)	44 (1.6)	45 (1.7)	28 (1.0)	612 (22.7)	311 (11.5)
2	709 (100.0)	424 (59.8)	5 (0.7)	16 (2.3)	18 (2.5)	9 (1.3)	162 (22.8)	75 (10.6)
3	432 (100.0)	244 (57.7)	11 (2.6)	9 (2.1)	9 (2.1)	2 (0.5)	90 (21.3)	58 (13.7)
4	862 (100.0)	500 (58.0)	12 (1.4)	26 (1.2)	26 (3.0)	5 (0.6)	214 (24.8)	95 (11.0)
5	555 (100.0)	317 (57.1)	4 (0.7)	13 (2.3)	16 (2.9)	4 (0.7)	136 (24.5)	65 (11.7)
6	898 (100.0)	544 (60.6)	12 (1.3)	28 (3.1)	6 (0.7)	5 (0.6)	218 (24.3)	85 (9.5)
7	871 (100.0)	489 (56.1)	10 (1.1)	20 (3.2)	28 (3.2)	11 (2.0)	211 (24.2)	102 (11.7)
8	1,050 (100.0)	650 (61.9)	11 (1.1)	16 (1.5)	15 (1.4)	7 (0.7)	229 (21.8)	122 (11.6)

Data are numbers of women, with horizontally calculated percentages in parenthesis.

Table S2. Reasons for participation stated by women interested in participation

	Overall	Socioeconomic status†			
		Q1 (lowest)	Q2	Q3	Q4 (highest)
Number of women	5,276	805	1,189	1,259	2,010
Stated reasons per category					
Expected personal health benefit	3,607 (68.4)	543 (67.5)	810 (68.1)	872 (69.3)	1,370 (68.2)
Contribution to science	2,245 (42.6)	311 (38.6)	473 (39.8)	537 (42.7)	920 (45.8)
(Breast) cancer in circle of family/friends	733 (13.9)	101 (12.5)	167 (14.0)	164 (13.0)	300 (14.9)
Positive family history for breast cancer	443 (8.4)	67 (8.3)	99 (8.3)	91 (7.2)	186 (9.3)
Breast cancer in personal history or breast complaints	194 (3.7)	26 (3.2)	43 (3.6)	51 (4.1)	73 (3.6)
Breast cancer in circle of acquaintances	91 (1.7)	9 (1.1)	16 (1.3)	20 (1.6)	46 (2.3)
Positive family/personal history of other cancer	78 (1.5)	9 (1.1)	20 (1.7)	20 (1.6)	29 (1.4)
Partner is diagnosed with cancer	7 (0.1)	0 (0.0)	3 (0.3)	0 (0.0)	4 (0.2)
Other reasons	392 (7.4)	54 (6.7)	93 (7.8)	89 (7.1)	156 (7.8)
Mammography is too painful	23 (0.4)	6 (0.7)	5 (0.4)	6 (0.5)	6 (0.3)
Job in health care or research	67 (1.3)	8 (1.0)	13 (1.1)	11 (0.9)	35 (1.7)
Advised to participate by others	8 (0.2)	0 (0.0)	5 (0.4)	1 (0.1)	2 (0.1)
Due to received invitation; otherwise not specified	141 (2.7)	16 (2.0)	37 (3.1)	34 (2.7)	54 (2.7)
Interested, but still in doubt about participation	22 (0.4)	1 (0.1)	6 (0.5)	5 (0.4)	10 (0.5)
Reason could not be classified	135 (2.6)	23 (2.9)	29 (2.4)	32 (2.5)	51 (2.5)
No reason mentioned (empty reply card)	422 (8.0)	69 (8.6)	110 (9.3)	95 (7.5)	149 (7.4)

Data are numbers of women with vertically calculated percentages in parenthesis. Total of percentages may not equal 100% due to rounding.
† Data on socioeconomic status is missing for 13 women.

4

Extramammary incidental findings in a population-based breast MRI screening trial for women with extremely dense breasts

S.V. de Lange, M.F. Bakker, R.M. Mann, M.J. Emaus, P.K. de Koekkoek-Doll, R.H.C. Bisschops, M.B.I. Lobbes, M.D.F. de Jong, K.M. Duvivier, J. Veltman, R.M. Pijnappel, C.H. van Gils, W.B. Veldhuis, for the DENSE Trial Study Group.

Submitted.

ABSTRACT

Background

Supplemental screening with breast magnetic resonance imaging (MRI) in women with extremely dense breasts reduces interval cancers. Little is known about the prevalence and clinical relevance of extramammary incidental findings (IFs) on breast cancer screening MRI examinations.

Purpose

To examine the frequency and clinical relevance of incidental findings that may be expected in a breast MRI population screening program, we investigated the prevalence, type and consequences of extramammary IFs on breast MRI in a population-based screening trial for women with extremely dense breasts.

Materials and Methods

Between 2011 and 2016, 4,783 women (50 to 75 years) underwent a breast screening MRI examination in the first screening round of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial. Prevalence of IFs was evaluated. Simple cysts (e.g. hepatic or pericardial), superficial skin lesions, as well as metastases and other lesions directly related to the primary breast cancer were not included. For all IFs, data on location, type, and further evaluation were collected. We investigated whether the occurrence of IFs was age dependent. Ethical approval was obtained. Participants provided written informed consent.

Results

In 61 of 4,783 first screening round MRI examinations, 63 IFs were recorded (1.3% of total number of women). Women with IFs were on average 58 years (interquartile range [IQR], 53 to 63), compared to 54 years (IQR, 51 to 59) for women without IFs ($p<0.001$). IFs detected were most frequently pulmonary (25.4%), followed by hepatic (22.2%), mediastinal and musculoskeletal IFs (both 20.6%). Further evaluation was recommended for 35 women with IFs (57% of IFs;0.7% of 4,783 women). Extramammary malignancy (hematologic, pulmonary or thymic) was histologically confirmed in 7 women (11% of IFs). Furthermore, 3 women (5% of IFs) were referred because of an aortic aneurysm.

Conclusion

IFs that required further evaluation were infrequent in this large group of MRI-screened women from the general population.

INTRODUCTION

Since the enactment of breast density notification laws, thus far in 38 states in the United States,[1] the use of magnetic resonance imaging (MRI) for supplemental screening in women with dense breasts has increased.[2,3] A federal law now requires breast density reporting to patients and their physicians.[4]

Recently, the results of the first screening round of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial showed that the use of supplemental MRI screening in women with extremely dense breast tissue and negative mammography results in an increase in breast cancer detection and in significantly fewer interval cancers.[5]

However, supplemental screening with MRI also leads to increased recall rates and to an increase of unnecessary biopsies due to false positive results.[5,6]

Different from mammography, regular breast MRI shows a large part of the thorax and the upper abdomen. Incidental findings are unavoidable when such large parts of the body are depicted. These findings lead to often unnecessary, possible invasive, further work-up, thereby increasing costs. Information on the prevalence and severity of incidental findings on screening breast MRI is therefore relevant for radiologists and policy makers. Factual knowledge of the frequency and consequences of incidental findings allows physicians to better educate participants, so that they can make more informed decisions about the benefits and potential harms of participation in a screening investigation. In addition, the capacity needed to handle the diagnostic work-up of incidental findings should be taken into account if supplemental MRI screening is implemented on a nation-wide scale. However, until now, little is known about the prevalence and clinical relevance of extramammary incidental findings on breast cancer screening MRI examinations.

To our knowledge, previous studies on the efficacy of breast cancer screening with MRI in women with average to moderate breast cancer risk, did not report on extramammary incidental findings.[7–10]

The aim of the current study was to investigate the frequency of occurrence of incidental findings in supplemental MRI screening of women with extremely dense breasts, to analyze the work-up and clinical relevance of those findings, and to assess whether a correlation with age exists.

MATERIALS AND METHODS

Study population

Data were collected from the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial.[5,11]

In short, this multicenter randomized controlled trial investigates the additional value of MRI screening in women with extremely dense breasts. The trial is carried out in the Dutch population based-screening program where all participants between 50 and 75 years old are screened biennially with digital mammography. Women with extremely dense breasts (Volpara density grade 4; Volpara Imaging Software version 1.5; Volpara Health Technologies, Wellington, New Zealand), and a negative mammography result after double reading (Breast Imaging Reporting and Data System classification [BI-RADS] 1 or 2) were eligible for randomization. Women randomized to the intervention group are invited to participate in three consecutive screening rounds with supplemental MRI.

The DENSE trial has been approved by the Dutch Minister of Health, Welfare and Sport (2011/19 WBO, The Hague, the Netherlands). Ethical approval was obtained on November 11, 2011. All participants provided written informed consent and since April 2013 all participants were additionally asked for consent, upfront, to be informed about (relevant) incidental findings on MRI.

Between December 2011 and January 2016, 4,783 women in the intervention arm participated in the first screening round of the trial, and underwent the first screening MRI examination. The current study reports on the incidental findings in this first MRI screening round.

MRI protocol

MRI examinations were performed with 3.0-T systems from Philips or Siemens, using a dedicated phased-array bilateral breast coil. The MRI protocol consisted of an optional T2-weighted sequence according to standard protocol, a diffusion-weighted sequence and a dynamic contrast-enhanced T1-weighted sequence.11 The macrocyclic gadolinium-based contrast agent gadobutrol (Gadovist; Bayer AG, Leverkusen, Germany) was used in all examinations. The MRI examinations were performed in eight centers and were assessed by breast radiologists with a minimum experience of 5 years in reading breast MRIs (range 5 to 23 years).

The Breast Imaging-Reporting and Data System MRI imaging lexicon of the American College of Radiology was used to assess the examinations, with final MRI assessment categories ranging from 1 to 5.[12] Radiologists reported incidental findings in the MRI reports, and in a free-text space in the online trial database. In case of incidental

findings for which further evaluation was advised by the radiologist, the woman's primary care physician was informed by the radiologist.

Assessment of incidental findings

All registered incidental findings were retrospectively categorized according to anatomical site. Clinical relevance was assessed: Incidental findings were considered clinically relevant in case the radiologist advised further evaluation or follow-up. Simple cysts (e.g. hepatic, pericardial) and solitary skin lesions for which no further evaluation was advised, were not included in the current study. Furthermore, in women with MRI detected breast cancer, extramammary metastases or other findings directly related to a primary breast cancer were not included.

Primary care physicians were contacted and medical records were evaluated to collect data on further evaluation, final diagnosis and treatment.

In case of multiple incidental findings in one examination, data were collected for all findings.

Statistical analysis

Prevalence of extramammary incidental findings was defined as the number of extramammary incidental findings divided by the number of MRI examinations. The number and percentage of incidental findings per anatomical site were described. Incidental findings were categorized into malignant and non-malignant findings (including benign tumors). The median age of women with incidental findings was compared to the median age of those without incidental findings and analyzed using the Mann-Whitney U test.

P values <0.05 were considered statistically significant. All statistical analyses were performed using R, version 3.2.2.

RESULTS

Between December 2011 and January 2016, 4,783 women underwent a first screening round MRI examination. Of these women, 4,480 (93.7%) gave consent to be notified in case of any incidental finding, 18 (0.4%) refused being notified of incidental findings (no incidental findings happened to occur in this 0.4% of participants). In the first 285 (5.9%) women invited for the study, this specific question was not yet part of the informed consent form.

Women who underwent first screening round MRI were on average 54 years old (interquartile range [IQR], 51 to 59]. The body mass index was normal (18.5 to 24.9 kg/m^2) in 83.1% of participants and 60.5% of the participants were postmenopausal (Table 1).

Table 1. Characteristics of first screening round participants

No. of women	4,783	
Median age, years*	54	(51-59)
Body mass index, kg/m^2		
<18.5	216	(4.9)
18.5-24.9	3,694	(83.1)
25.0-29.9	484	(10.9)
≥30	51	(1.1)
Missing	338	
Menopausal status†		
Premenopausal	473	(10.3)
Perimenopausal	1,337	(29.2)
Postmenopausal	2,768	(60.5)
Missing	205	

Unless otherwise noted, data are numbers with percentages in parentheses.
* Data in parentheses are interquartile ranges.
† Women aged ≥60 years, or reporting history of hysterectomy 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.

In total, 63 extramammary incidental findings were reported for 61 women (1.3% of 4,783 women) (Figure 1). Two women had 2 incidental findings. In 56 of 61 women with an incidental finding (91.8% of women with incidental findings), the MRI examination had an otherwise normal result (BI-RADS category 1 or 2).

Women with incidental findings were on average 58 years (IQR, 53 to 63) old at time of MRI examination, compared to 54 years (IQR, 51 to 59) for those without incidental findings ($p<0.001$).

The most frequent anatomical site of incidental findings was pulmonary or pleural (in 16 women, 25.4%). Hepatic findings were detected in 14 women (22.2%). Mediastinal and musculoskeletal findings were detected in the same frequency (13 women each, 22.2%) (Table 2). Skin lesions were registered for 3 women (3.1%), in two of these women the lesions were based on known neurofibromatosis.

For 35 women (57.4% of women with findings; 0.7% of all 4,783 women) further evaluation of the incidental findings was advised. Of these women, 34 had consented

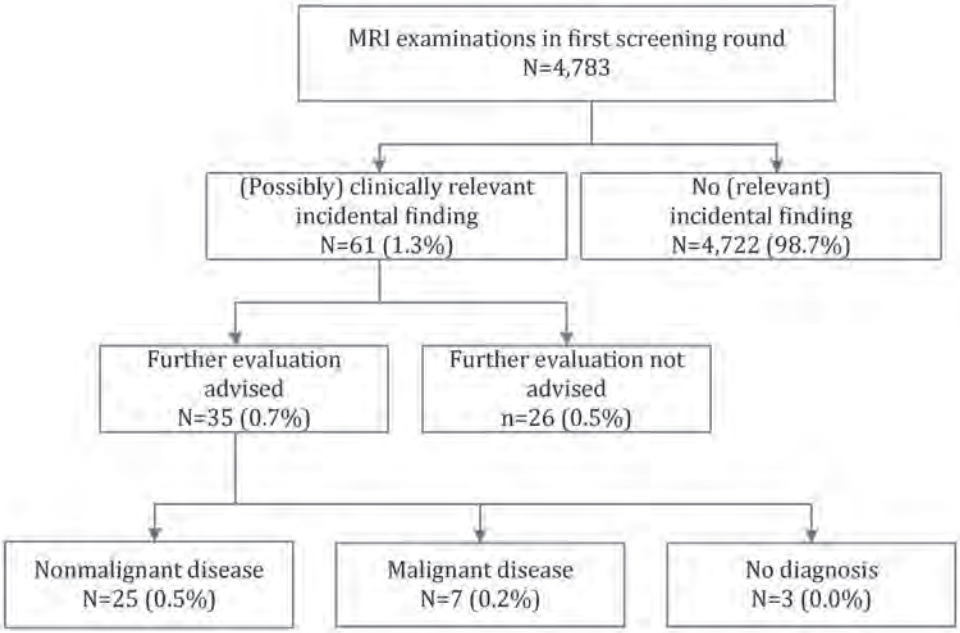


Figure 1. Flowchart of reported incidental findings in 4,783 first screening round MRI examinations.

upfront to be informed about incidental findings, one woman was included in the trial before April 2013. She was informed about the incidental finding and was referred for further evaluation.

For 26 women (41.6%) with incidental findings further evaluation was not recommended.

Of the 35 women who were referred for further evaluation of the incidental finding, 7 (20%) were diagnosed with a primary malignancy (Table 3). Four women were diagnosed with a pulmonary malignancy, of whom one died from that malignancy. Two women were diagnosed with thymoma/thymic carcinoma, one woman was diagnosed with lymphoma. Three women were diagnosed with widening or aneurysm of the ascending aorta (>4.5cm), of whom one woman died from aneurysm rupture despite immediate action by the general practitioner. Furthermore, one woman was diagnosed with a pulmonary hamartoma, and one with a paravertebral schwannoma.

Table 2. Localization of relevant extramammary incidental findings in 4,783 MRI examinations

	All incidental findings		Only incidental findings with further evaluation advised	
	N	(%)	N	(%)
No. of incidental findings‡	63	(1.3)*	37	(0.7)†
Anatomical site				
Hepatic	14	(22.2)	4	(10.8)
Pulmonary or pleural	16	(25.4)	14	(37.8)
Mediastinal	13	(20.6)	9	(24.3)
Aorta/cardiac	7		4	
Thymus	2		2	
Oesophagus	1		0	
Lymph nodes	3		3	
Musculoskeletal	13	(20.6)	6	(16.2)
Rib	5		2	
Shoulder	3		1	
Sternum	3		2	
M. Pectoralis	2		1	
Axillary lymph nodes§	4	(6.3)	3	(8.1)
Skin	3	(4.8)	1	(2.7)

MRI = magnetic resonance imaging.
Unless otherwise noted, data are numbers of incidental findings (IFs) with percentage of IFs within parenthesis. Total of percentages may not add up to exactly 100% due to rounding.

* Incidental findings were recorded in 61 MRI examinations.

† For 37 incidental findings in 35 MRI examinations further evaluation was advised. One woman had pulmonary and hepatic finding and one woman had a mediastinal and a pulmonary finding.

‡ Percentage of total MRI examinations (61/4,783 and 35/4,783).

§ Axillary enlarged lymph nodes only. One woman with both mediastinal and axillary lymph nodes was not included in this count.

Table 3. Final diagnosis of extramammary incidental findings in case of referral (N=35)

	N	(%)
Primary malignant tumors	7	(20.0)
Lymphoma	1	
(Atypical) lung carcinoid	2	(5.7)
Non-small cell lung cancer (squamous cell/adenocarcinoma)	2	
Thymoma/thymic carcinoma*	2	
Non-malignant tumors	2	
Pulmonary hamartoma	1	
Paravertebral schwannoma	1	
Aortic widening/aneurysm (>4.5 cm)†	3	(8.6)
Other	20	(57.1)
Pneumonia	4	
Atypical/benign pulmonary nodules (>1 cm)	2	
Bronchiectasis	1	
Rib fractures‡	2	
Hemangioma§	4	
Hydrops of the shoulder joint¶	1	
Intramuscular lipoma of the pectoral muscle	1	
Reactive lymphadenopathy	4	
Thymic remnant tissue	1	
No diagnosis ||	3	(8.6)

Data are numbers of incidental findings with percentage of incidental findings for which further evaluation was advised (n=35) within parentheses.

* One woman was diagnosed with thymoma type A and one with type 3B.

† Diameters of the thoracic aorta were >6 cm and 4.6 cm (ascending aorta), and 4.5 cm (aortic root).

‡ In one woman no further evaluation was performed.

§ In one woman on further imaging the liver lesions turned out to be liver cysts.

¶ Diagnosis was based on breast MRI; further imaging assessment was not performed.

|| In two women sternal abnormalities were not present on further imaging assessment; in one woman a possible pulmonary mass turned out to be a phase-encoding artifact from a large (normal) vena cava superior.

DISCUSSION

Until now, knowledge on extramammary incidental findings in supplemental screening 3.0-T breast MRI examinations in the setting of population screening for breast cancer is scarce.

We assessed all screening MRI examinations and found that incidental findings were reported in 1.3% of women. Of these women, 35/4,783 (0.7% of all participants) were referred for further evaluation of the finding. In the 35 referred women, the prevalence of malignant disease was 20% (7 women, 0.2% of all participants).

Several smaller studies retrospectively evaluated the prevalence of extramammary incidental findings on 1.5 T breast MRI performed for screening women with high breast cancer risk based on genetic or familial predisposition, pre-operative staging, surveillance in patients with known breast cancer or another clinical indication. Our study showed a lower prevalence of extramammary incidental findings on breast MRI than these previously conducted studies, which showed a prevalence of incidental findings ranging from 11% to 34%, and clinically relevant findings in up to 8% of MRI examinations.[13–20]

The differences in prevalence of incidental findings compared to ours may in part be due to differences in study population, since these studies were mostly performed in women diagnosed with breast cancer instead of in the general population. Even after exclusion of simple cysts, skin lesions and metastasis of (known) breast cancer, these studies showed a prevalence of incidental findings ranging from 7% to 25%.[13–17,20]

Some of these previous studies showed that (malignant) incidental findings were more prevalent in patients with known breast cancer than in women who underwent screening because of high breast cancer risk based on genetics or familial predisposition.[14,15,17] Niell et al. showed that further evaluation of incidental findings was recommended less frequently in women who underwent breast MRI for screening than in women who underwent a diagnostic MRI.[13]

A large meta-analysis has been performed by Gibson et al, that reported on incidental findings in population-based research imaging, screening studies (commercial or occupational) and studies among research controls. Studies in patients (diagnostic or performed because of risk factors or disease) were excluded. That meta-analysis showed a pooled prevalence of potentially serious incidental findings on thoracic MRI of 1.3% (95% confidence interval [CI]:0.2 to 8.1%) and 1.9% (95%CI: 0.3 to 12.0%) on abdominal MRI. Respectively 0.6% and 1.3% were suspected malignancies.[21]

Our study found a slightly lower percentage of potentially serious findings, with 0.7% of women being referred because of incidental findings, and among these also a lower fraction of malignancies.

The anatomic region studied is comparable to thoracic MRI. However, the positioning of the coil is different, and therefore the signal to noise ratio within the thorax is lower, particularly dorsally. Also, artefacts in the thorax caused by heart beating or unfolding of the arms commonly occur due to a choice for phase encoding directions that is optimal for breast imaging, but suboptimal for thoracic evaluation. The lack of a systematic approach for reporting of incidental findings might have resulted in underreporting of incidental findings. However, we expect that a systematic approach would have mainly resulted in more cysts and other clinically irrelevant incidental findings, while the absence of a mandated approach is unlikely to have deterred trained radiologists from reporting possibly relevant findings.

Incidental detection of an extramammary malignancy occurred in only 7 of 4,783 women in our study. Even in these women it remains uncertain whether the incidental detection of these hitherto unknown malignancies conveyed a benefit.

Likewise, the extent of potential harms, due to additional investigations and anxiety caused by incidental findings remains unclear.

What is clear is that the frequency of incidental findings that required further evaluation was low, in this large group of MRI-screened women from the general population. In addition, this study assessed first screening round results, where any prevalent lesion will be detected: the rate of detection of new incidental findings will very likely be even lower in subsequent, incidence screening rounds.

In summary, our results give an indication of how many clinically relevant incidental findings can be expected in case supplemental breast MRI screening for women with extremely dense breasts is implemented in a nation-wide breast cancer screening program.

ACKNOWLEDGMENTS

The authors thank the study participants for their contributions. The authors thank the regional screening organizations, Volpara Health Technologies, the Dutch Expert Centre for Screening, and the National Institute for Public Health and the Environment for their advice and in-kind contributions for organizing the breast density assessments and the invitation of potential participants.

FUNDING

The DENSE trial is financially supported by the University Medical Center Utrecht (UMC Utrecht, Project number: UMCU DENSE), the Netherlands Organization for Health Research and Development (ZonMw, Project numbers: ZONMW-200320002-UMCU and ZonMW Preventie 50-53125-98-014), the Dutch Cancer Society (KWF Kankerbestrijding, Project numbers: DCS-UU-2009-4348, UU-2014-6859 and UU2014-7151), the Dutch Pink Ribbon/A Sister's Hope (Project number: Pink Ribbon-10074), Bayer AG Pharmaceuticals, Radiology (Project number: BSP-DENSE), and Stichting Kankerpreventie Midden-West. For research purposes, Volpara Health Technologies (Wellington, New Zealand) provided Volpara Imaging Software 1.5 for installation on servers in the screening units of the Dutch screening program.

REFERENCES

1. DenseBreast-info. Legislation and Regulations for Dense Breast | DenseBreast-Info. Published March 28, 2019. Accessed June 13, 2020. https://densebreast-info.org/legislation.aspx
2. Chau SL, Alabaster A, Luikart K, Brenman LM, Habel LA. The effect of California's breast density notification legislation on breast cancer screening. *Journal of Primary Care and Community Health*. 2017;8(2):55-62.
3. Sanders LM, King AB, Goodman KS. Impact of the New Jersey breast density law on imaging and intervention volumes and breast cancer diagnosis. *Journal of the American College of Radiology*. 2016;13(10):1189-1194.
4. DenseBreast-Info. Is there a federal law for breast density? Published March 28, 2019. Accessed June 10, 2020. https://densebreast-info.org/is-there-a-federal-law.aspx
5. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *The New England Journal of Medicine*. 2019;381(22):2091-2102.
6. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental screening for breast cancer in women with dense breasts: A systematic review for the U.S. Preventive services task force. *Annals of Internal Medicine*. 2016;164(4):268-278.
7. 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: The Journal of the American Medical Association*. 2012;307(13):1394-1404.
8. 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. *The Lancet Oncology*. 2019;20(8):1136-1147.
9. 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: The Journal of the American Medical Association.* 2020;323(8):746-756.
10. 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):361-370.
11. Emaus MJ, Bakker MF, Peeters PHM, 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.
12. Ikeda D, Hylton N, Kuhl CK. *BI-RADS: Magnetic Resonance Imaging*. American College of Radiology; 2003.
13. Niell BL, Bennett D, Sharma A, Gazelle GS. Extramammary findings on breast MR examinations: Frequency, clinical relevance, and patient outcomes. *Radiology*. 2015;276(1):56-64.
14. Moschetta M, Telegrafo M, Rella L, Stabile Ianora AA, Angelelli G. Let's go out of the breast: Prevalence of extra-mammary findings and their characterization on breast MRI. *European Journal of Radiology*. 2014;83(6):930-934.
15. Alduk AM, Prutki M, Stern-Padovan R. Incidental extra-mammary findings in breast MRI. *Clinical Radiology*. 2015;70(5):523-527.

16. Padia SA, Freyvogel M, Dietz J, Valente S, O'Rourke C, Grobmyer SR. False-positive extra-mammary findings in breast MRI: another cause for concern. *Breast Journal*. 2016;22(1):90-95.
17. Rinaldi P, Costantini M, Belli P, et al. Extra-mammary findings in breast MRI. *European Radiology*. 2011;21(11):2268-2276.
18. Staab W, Menke J, Sohns C, et al. Prevalence and clinical significance of incidental extra-mammary findings in breast magnetic resonance imaging: A retrospective study of 1070 patients. *Clinical Imaging*. 2017;45:92-95.
19. Knox M, Slanetz P, Phillips J, et al. Incidental liver lesions seen on Breast MRI: When is additional imaging warranted? *European Journal of Radiology*. 2017;95:319-324.
20. Iodice D, di Donato O, Liccardo I, et al. Prevalence of extramammary findings on breast MRI: A large retrospective single-centre study. *Radiologia Medica*. 2013;118(7):1109-1118.
21. Gibson LM, Paul L, Chappell FM, et al. Potentially serious incidental findings on brain and body magnetic resonance imaging of apparently asymptomatic adults: Systematic review and meta-analysis. *BMJ (Online)*. 2018;363.

SUPPLEMENTARY APPENDIX

Members DENSE trial study group

University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands:
CH van Gils PhD, MF Bakker PhD, SV de Lange MD, SGH 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, R Mus MD, M Imhof-Tas MD, N Karssemeijer PhD

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 Center, 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

5

Supplemental MRI for women with extremely dense breasts – results of the second screening round of the DENSE trial

S.G.A. Veenhuizen, S.V. de Lange, M.F. Bakker, R.M. Pijnappel, R.M. Mann, E.M. Monninkhof, M.J. Emaus, P.K. de Koekkoek-Doll, R.H.C. Bisschops, M.B.I. Lobbes, M.D.F. de Jong, K.M. Duvivier, J. Veltman, N. Karssemeijer, H.J. de Koning, P.J. van Diest, W.P.T.M. Mali, M.A.A.J. van den Bosch, W.B. Veldhuis**, C.H. van Gils**, for the DENSE Trial Study Group.

** Shared last authorship

Submitted.

ABSTRACT

Background

Results of the first (prevalence) screening round of the DENSE trial have shown an MRI cancer-detection rate (CDR) of 16.5 per 1000 screenings and a false positive rate (FPR) of 79.8 per 1000, after negative mammographic screening in women with extremely dense breasts. In incidence screening rounds, a lower CDR is expected, due to a smaller pool of prevalent cancers, as well as a reduced FPR, due to availability of previous MRI examinations.

Methods

The DENSE trial is embedded within the Dutch population-based biennial digital mammography screening program, for women aged 50 to 75 years. Women were eligible for the second round when they had a negative result at regular mammography screening two years after their first MRI. The recall rate, biopsy rate, CDR, FPR, and positive predictive values (PPVs) of second round MRI screenings were determined, as well as tumor characteristics of MRI-detected cancers. This trial is registered on ClinicalTrials.gov, number NCT01315015.

Findings

Of 4,783 women in the MRI arm, 3,436 (71.8%) underwent a second MRI; 110 (3.2%) were recalled of whom 84 (2.5%) underwent a biopsy. The CDR was 5.8 per 1000 screenings (95% CI, 3.8 to 9.0) and the FPR was 26.3 per 1000 (95% CI, 21.5 to 32.3). PPV for recall was 18.2% (95% CI, 12.1 to 26.4) and for biopsy 23.8% (95% CI, 16.0 to 33.9).

Interpretation

Although lower than in the first round, the CDR was still 5.8 per 1000 screenings in the second round and the number of false-positive results sharply decreased. The findings indicate that supplemental MRI screening improves cancer detection, also in incidence rounds. Longer follow-up is needed to evaluate how this lowers the advanced cancer rate.

INTRODUCTION

Currently, screening mammography is the primary examination standard for early detection of breast cancer in women at average risk.[1]However, especially for women with dense breasts, the sensitivity of mammographic screening has proven to be limited, due to the masking effect of dense breast tissue.[2,3] Therefore, the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial investigated the effectiveness of adding a more sensitive screening tool for women with extremely dense breasts; magnetic resonance imaging (MRI).[4] Results of the first (prevalence) screening round have been published previously.[5] After a negative ('normal') screening mammography, supplemental screening with MRI led to an additional cancer-detection rate (CDR) of 16.5 per 1000 screenings, yet at the expense of a false positive rate (FPR) of 79.8 per 1000 screenings. Subsequently, compared to the mammography-only group, the interval cancer rate in the group of women invited for MRI was reduced from 5.0 to 2.5 per 1000 screenings. This 2.5 per 1000 screenings in the MRI-invitation group consisted of an interval cancer rate of 0.8 per 1000 screenings in women who actually underwent MRI (59% of invited women), and 4.9 per 1000 screenings in those who did not accept the MRI invitation.

These findings are promising for the early detection of breast cancer in women with dense breasts. However, to be able to estimate long-term benefits and harms of breast MRI screening, information from incidence screening rounds is needed as well. Here, we present the results of the second (incidence) round of the DENSE trial, focusing on the recall rate, CDR, the FPR, the positive predictive value (PPV) and tumour characteristics.

METHODS

Study design and participants

The DENSE trial design has been extensively described elsewhere.[5,6] This multicenter trial is embedded within the Dutch population-based biennial digital mammography screening program for women between 50 to 75 years of age. Women were eligible for the DENSE trial if they had a negative mammography result (Breast Imaging Reporting and Data System (BI-RADS) category 1 or 2) and extremely dense breast tissue, which was defined as density category 4 or d, measured with Volpara imaging software, version 1.5 (Volpara Health Technologies). Women were randomized to the intervention arm (invitation to undergo supplemental MRI screening, n=8,061), or to the control arm (biennial mammographic screening only, n=32,312).

The trial has been designed to consist of three consecutive screening rounds. Women who completed the first screening MRI were eligible for a second MRI round if they had responded to their next invitation of the regular mammography screening program,

and had a negative screening result. Eligibility for a second MRI round was independent of any change in breast density since the first round.

The DENSE trial has been approved by the Dutch Minister of Health, Welfare and Sport, who was advised by the Health Council of the Netherlands (2011/2019 WBO, The Hague, The Netherlands), and all participants provided written informed consent.

Breast MRI

MRIs were read by the same group of radiologists as in the prevalence round, and classified according to the BI-RADS MRI ACR lexicon.[7,8] If women had a BI-RADS 4 or 5 score, they were recalled for additional diagnostic work-up. In case of a BI-RADS 3 score, double reading was performed, and if there was consensus on a score of 3 a follow-up MRI after six months was planned. The follow-up MRI had to be reported as either negative (BI-RADS 1 or 2, with return to the regular mammography screening program) or positive (BI-RADS 4 or 5 score) after which women were recalled for additional work-up.

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. Details of the full screening MRI protocol have been described previously.[5,6]

Outcomes

The primary outcomes in this report are the recall rate, the CDR, the FPR, PPVs and tumor characteristics in the second round of MRI screening. The interval cancer rate, another important outcome of interest, is not included in this report because data collection for this end point is still ongoing. Calculations of all outcomes have been described previously, and are referred to in Table 1.[5]

We described tumor characteristics of MRI detected cancers: morphology, tumor-node-metastasis (TNM) stage, grade, and receptor status. In women with more than one tumor we described the one with the highest TNM stage. All rates were presented per 1000 screenings. A comparison was made with previously published results of the first round.[5] Differences in distributions of tumor characteristics between the first and second rounds were tested using Chi square test or Fisher's exact test in case of low counts.

Adverse events (AEs) and serious adverse events (SAEs) were recorded during or immediately after the MRI examination in the trial center or reported by the women within 30 days. An SAE was defined as an adverse event that required emergency department visit or unplanned hospital admission.

All analyses were performed with the use of RStudio software (RStudio, version 1.1.456).

Role of the funding source

The funders of the DENSE trial were not involved in study design, data collection, data analysis, data interpretation, or writing of the manuscript.

RESULTS

Of the 4,783 women who participated in the first round of the DENSE trial, 3,436 (71.8%) also participated in the second MRI screening round (Figure 1). Women who were not invited for a subsequent mammography screening round, because of age (>75 years), emigration or death, or who declined the invitation for mammography screening, or who actively withdrew from the DENSE trial before re-invitation, were not re-invited for the second MRI screening round of the DENSE trial (N=319). The MRI participation rate among those who were re-invited was 81.3%.

Of the 3,436 participants in the second MRI screening round, 110 were recalled because of a finding classified as BI-RADS 3 (N=20), BI-RADS 4 (N=87) or BI-RADS 5 (N=3) (Figure 1 and Table 1). This resulted in a recall rate of 32.0 per 1000 screenings (95% confidence interval [CI], 26.6 to 38.4).

Of the recalled women, 92 had an indication for additional work-up (BI-RADS 4 or 5). In two women this indication was based on a BI-RADS 4 or 5 score on a 6-month follow-up MRI, after a BI-RADS 3 score on the second MRI screening round. Eighty-four of these women underwent a breast biopsy, which resulted in a biopsy rate of 24.4 per 1000 screenings (95% CI, 19.8 to 30.2).

In the second MRI round, 20 women were diagnosed with an MRI screen-detected breast cancer, of which 6 were ductal carcinoma in situ (DCIS), and 14 were invasive cancers (Tables 1 and 2). For comparison we included the findings from the first round in italics in Tables 1 and 2.[5] The CDR of the second round was 5.8 per 1000 screenings (95% CI, 3.8 to 9.0), compared to a CDR of 16.5 per 1000 screenings in the first round (95% CI, 13.3 to 20.5). The FPR in the second round was 26.3 per 1000 screenings (95% CI, 21.5 to 32.3) (specificity, 97%), compared to an FPR of 79.8 per 1000 screenings in the first round (95% CI, 72.4 to 87.9). In the second round, the FPR due to BI-RADS 3 scores was 5.2 per 1000 screenings (95% CI, 3.3 to 8.3), and the FPR due to BI-RADS 4 and 5 scores was 21.1 per 1000 screenings (95% CI, 16.8 to 26.5). PPV of recall (BI-RADS 3, 4 or 5) was 18.2% (95% CI, 12.1 to 26.4) in the second round, compared to 17.4% in the first round (95% CI, 14.2 to 21.2). PPV of a BI-RADS 4 or 5 classification was 21.7% (95% CI, 14.5 to 31.2) in the second round, compared 23.9% in the first round (95% CI, 19.6 to 28.8) and PPV of biopsy was 23.8% (95% CI, 16.0 to 33.9) in the second round, compared to 26.3% in the first round (95% CI, 21.7 to 31.6).

Table 1. Screening performance of second (incidence) supplemental MRI screening round

	Women who underwent second screening MRI (N=3,436)				Prevalence screening round†	
			Rate (95% CI)		Rate (95% CI)	
	no./total no.	*(%)*	*no./1000 screenings*		*no./1000 screenings*	
Second round of MRI screening						
Women who were recalled	110/3,436	(3.2)				
BI-RADS 3	20/110					
BI-RADS 4	87/110					
BI-RADS 5	3/110					
Women with biopsy indication‡	92/3,436	(2.7)				
BI-RADS 4 or 5 on second MRI	90/92					
BI-RADS 4 or 5 on 6-mo follow-up MRI after initial BI-RADS 3†	2/92					
Women who underwent biopsy¶	84/3,436	(2.4)				
After BI-RADS 4 or 5 on second MRI	82/84					
After BI-RADS 4 or 5 on 6-mo follow-up MRI after initial BI-RADS 3	2/84					
Women with MRI screen-detected cancers	20/3,436					
Type of cancer						
Ductal carcinoma in situ	6/20	(30.0)				
Invasive cancer	14/20	(70.0)				
Recall rate [a]	110/3,436		32.0	(26.6-38.4)		
Biopsy rate [b]	84/3,436		24.4	(19.8-30.2)		
All cancers						
Cancer-detection rate [c]	20/3,436		5.8	(3.8-9.0)	*16.5*	*(13.3-20.5)*
False positive rate [d]	90/3,416		26.3	(21.5-32.3)	*79.8*	*(72.4-87.9)*
False positive rate BI-RADS 3 scores	18/3,416		5.2	(3.3-8.3)	*31.1*	*(26.5-36.4)*
False positive rate BI-RADS 4 and 5 scores	72/3,416		21.1	(16.8-26.5)	*48.7*	*(42.9-55.3)*
PPV of positive MRI result [e]	20/110		18.2	(12.1-26.4)	*17.4*	*(14.2-21.2)*
PPV of indication for biopsy [f]	20/92		21.7	(14.5-31.2)	*23.9*	*(19.6-28.8)*
PPV of biopsy [g]	20/84		23.3	(16.0-33.9)	*26.3*	*(21.7-31.6)*

Table 1. Continued.

	Women who underwent second screening MRI (N=3,436)			*Prevalence screening round*†
			Rate (95% CI)	*Rate (95% CI)*
	no./total no.	*(%)*	*no./1000 screenings*	*no./1000 screenings*
Invasive cancers				
Cancer-detection rate	14/3,436		4.1 (2.4-6.8)	*13.4 (10.5-17.1)*
False positive rate	96/3,422		28.1 (23.0-34.1)	*82.7 (72.5-90.0)*
PPV of positive MRI result	14/110		12.7 (7.7-20.2)	*14.1 (11.2-17.6)*
PPV of indication for biopsy	14/92		15.2 (9.3-23.9)	*19.3 (15.4-23.9)*
PPV of biopsy	14/84		16.7 (10.2-26.1)	*21.3 (17.1-26.3)*

MRI= magnetic resonance imaging; PPV= positive predictive value. False positive rates for the incidence round were calculated without interval cancers, since this data collection is still ongoing.

† Results of the prevalence screening round were adapted from previous publication by Bakker et al.[5]

§ 2 of 90 women initially had a BI-RADS 3 result (initial second screening round MRI) and a BI-RADS 4 or 5 at follow-up (indication for biopsy was based on the follow-up MRI).

¶ 8 of 90 women with an indication for biopsy did not undergo a biopsy because the lesion was not or no longer visible on additional imaging (N=4), the lesion turned out be an intramammary lymph node or cyst (N=2), or the lesion was already histologically proven to be benign elsewhere (N=2).

a The recall rate was calculated by dividing the number of recalled women (BI-RADS 3, 4, 5) by the total number of women who underwent a second screening MRI.

b The biopsy rate was calculated by dividing the number of women who underwent a biopsy by the total number of women who underwent a second screening MRI.

c The cancer detection rate was calculated by dividing the number of MRI-detected cancers by the total number of women who underwent a second screening MRI.

d The overall FPR was calculated by dividing the number of women who were recalled (BI-RADS 3, 4 or 5) by the number of women who underwent a second screening MRI and did not have an MRI screen-detected cancer. The FPR was also calculated separately for BI-RADS 3 scores and BI-RADS 4 and 5 scores (including BI-RADS 4 or 5 scores after a BI-RADS 3 score at the initial MRI).

e PPV of positive MRI: Women with MRI screen-detected breast cancer among women with a positive MRI result (BI-RADS 3, 4 or 5).

f PPV of an indication for biopsy: Women with MRI screen-detected breast cancer among women with an indication for biopsy (BI-RADS 4 or 5, including women with BI-RADS 3 at initial MRI and BI-RADS 4 or 5 at follow-up MRI).

g PPV of biopsy: Women with MRI screen-detected breast cancer among women who underwent biopsy (BI-RADS 4 or 5, including women with BI-RADS 4 or 5 at follow-up MRI).

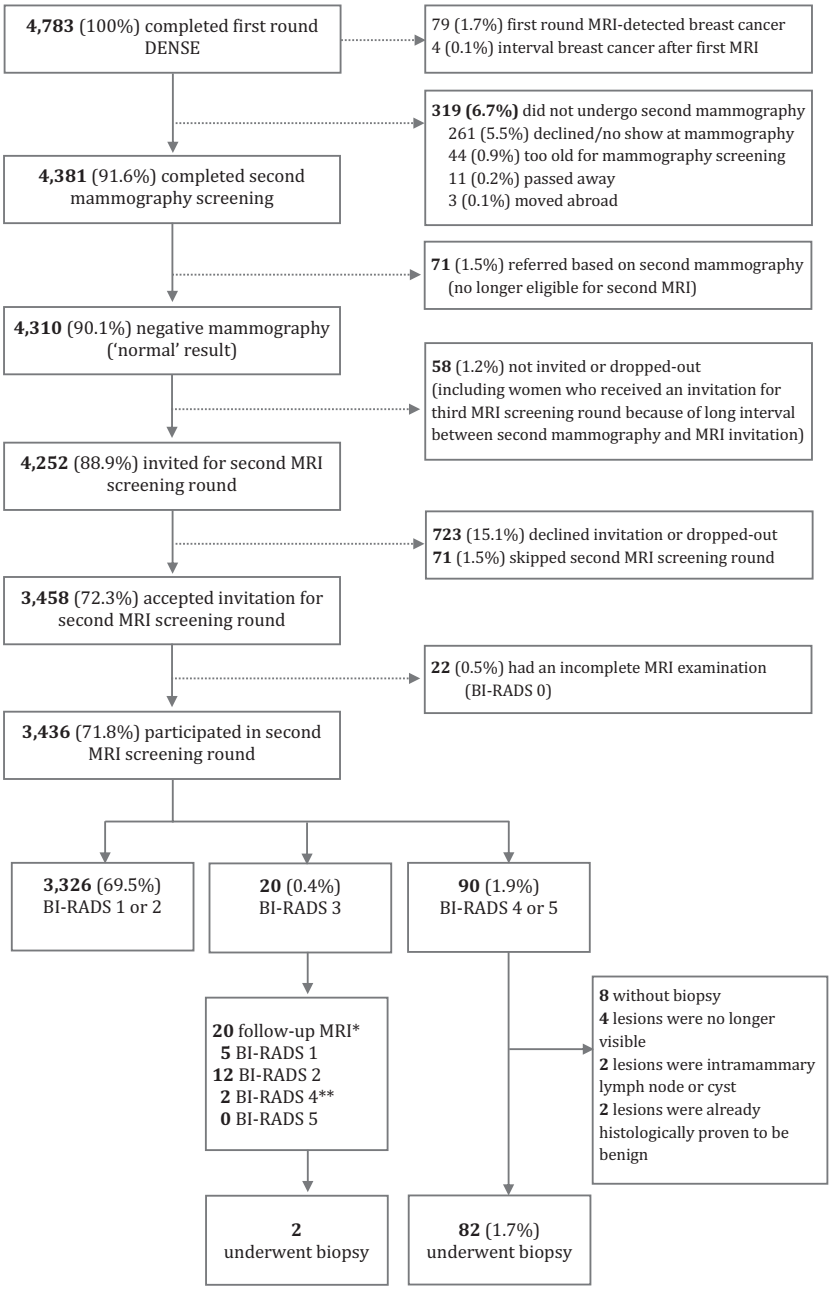


Figure 1. Participation in second MRI screening round.
*1 woman dropped out before the follow-up MRI was performed
**2 women underwent biopsy after BI-RADS 4 result at follow-up MRI

The detection rate of DCIS was 1.7 per 1000 screenings (95% CI, 0.8 to 3.8) in the second round, compared to 3.1 per 1000 screenings (95% CI, 1.9 to 5.2) in the first round. The percentage of DCIS among all cancers detected in the second round was 30%, compared to 19% in the first round, but this difference was not statistically significant (p=0.36, Fisher's exact test). All of the detected DCIS were moderately or poorly differentiated.

In the second round, none of the MRI-detected cancers were node positive (0 per 1000, 95% CI, 0 to 16.1) or late stage (stage II to IV) (0 per 1000, 95% CI, 0 to 16.1) at diagnosis, compared to 1.9 per 1000 screenings (95% CI, 1.0 to 3.6) and 1.5 per 1000 screenings (95% CI, 0.8 to 3.0), in the prevalence round respectively.

The proportion of poorly differentiated (grade III) invasive cancers was comparable for the first and second screening round (6.8% and 6.7%, respectively, p=1.00, Fisher's exact test) (Table 2). In the second screening round, all 14 invasive cancers were positive for estrogen and/or progesterone receptor, compared to 56 out of 64 invasive cancers in the first screening round (p=0.34, Fisher's exact test).

During the second round of MRI screening one SAE occurred, caused by a vasovagal reaction, leading to an emergency department visit. Additionally, two AEs were registered during or immediately after the screening MRI, both caused by extravasation of the contrast agent.

Table 2. Characteristics of MRI screen-detected cancers

Characteristic	**Incidence round (second MRI screening round) (N=3,436)**		*Prevalence round* (first MRI screening round) † (N=4,783)	
	no.	*no./1000 screenings*	*no.*	*no./1000 screenings*
Women diagnosed with cancer‡	20		79	
Histologic type				
Ductal carcinoma in situ*	6	1.7	15	3.1
Invasive ductal carcinoma	8	2.3	35	7.3
Invasive lobular carcinoma	3	0.9	9	1.9
Mixed invasive ductal and lobular carcinoma	2	0.6	8	1.7
Tubular carcinoma	1	0.3	7	1.5
Other invasive carcinoma	0	0.0	5	1.0
Tumor size, mm§¶				
Invasive tumor size, mm§	7.0	(6.0-10.0)	9.5	(6.8-12.0)
Missing	1		3	
Status for axillary lymph nodes§				
Negative	20	5.8	70	14.6
Positive	0	0.0	9	1.9
Tumor stage				
Early (0 or I)	20	5.8	72	15.1
Late (II, III, or IV)	0	0.0	7	1.5
Tumor grade ‖				
DCIS				
I, well-differentiated	0	0.0	6	1.3
II, moderately differentiated	5	1.5	6	1.3
III, poorly differentiated	1	0.3	3	0.6
Missing data or could not be assessed	0		0	
Invasive				
I, well-differentiated	6	1.7	31	6.5
II, moderately differentiated	7	2.0	24	5.0
III, poorly differentiated	1	0.3	4	0.8
Missing data or could not be assessed	0		5	
Receptor status**				
Positive for estrogen receptor, progesterone receptor, or both	14	4.1	56	11.7
HER2 enriched	0	0.0	2	0.4
Triple negative	0	0.0	4	0.8
Missing data	0		2	

Unless otherwise noted, data are numbers, with risk per 1000 screens in parenthesis. MRI= magnetic resonance imaging; ER= estrogen receptor; PR= progesterone receptor.

* Includes comedo type of ductal carcinoma in situ (DCIS), intraductal papillary carcinoma with DCIS or Paget's disease of the nipple (non-invasive with or without DCIS).

† Results of the prevalence screening round are adapted from previous publication by Bakker et al.[5]

‡ One woman was diagnosed with two synchronous breast cancers. Only the tumor with the highest stage was used for all analyses.

§ Median (interquartile range).

¶ Tumor size based on pathology result, for invasive cancers only.

‖ Risks per 1000 screens in parenthesis are calculated for the total number of DCIS and invasive cancers, respectively.

** Data regarding receptor status are provided only for invasive cancers.

DISCUSSION

In the second round of supplemental MRI screening of the DENSE trial, again following a negative mammography result, the additional CDR was 5.8 per 1000 screenings, compared to a CDR of 16.5 per 1000 screenings in the prevalence round.[5] The FPR was 26.3 per 1000 screenings, compared to 79.8 per 1000 screenings in the prevalence round.[5] No node positive or late stage cancers were observed in the incidence round.

The decrease in CDR in the incidence round as compared to the prevalence round is as expected, because in the prevalence round a larger pool of prevalent breast cancers is expected that could not be detected with mammography only.[5] This has also been observed in other prospective studies with incidence rounds of supplemental MRI screening. The MRI study that is most comparable to ours in terms of target population is that of Kuhl et al. who studied women between 40 to 70 years of age, with an average risk of breast cancer and a wide range of breast densities (approximately 20% extremely dense).[9] They found an average CDR of 6.9 per 1000 screenings during incidence screening rounds, compared to 22.6 per 1000 screenings in the prevalence round. Approximately half of their study population underwent two screening rounds, and screening intervals ranged between 12 and 36 months.

There are no other studies to date with a comparable study population and MRI results from more than one screening round. The studies mentioned below are performed in higher risk populations, but cited solely to illustrate the comparison of incidence with prevalence MRI screening rounds. In the FaMRIsc trial, women aged 30 to 55 years, with a lifetime breast cancer risk of at least 20% based on family history, but BRCA1, BRCA2 and TP53 wild-type, were examined with annual MRI, clinical breast examination and biennial mammography, compared to annual mammography and clinical breast examination.[10] The additional CDR in incidence rounds in the intervention arm was 4.1 per 1000 screenings compared to 23.8 per 1000 screenings in the prevalence round.[10] In a study performed by Vreemann et al. women at increased breast cancer risk (BRCA carriers, women with a family or personal history) were annually screened with MRI and mammography, where CDRs on MRI (irrespective of mammographic screening) of 10.6 and 22.5 per 1000 in their incidence versus prevalence rounds respectively were found.[11] The German Consortium for Hereditary breast and Ovarian Cancer report on the 10-year breast cancer annual MRI screening program for women with increased risk of breast cancer (BRCA1 mutation carriers and non-carriers with family history). Among BRCA1 mutation carriers an increase in CDR of 19.9 to 28.4 per 1000 screenings from prevalence to incidence rounds was seen, whereas among BRCA2 mutation carriers and non-carriers a decrease in CDRs from prevalence to incidence rounds was seen; 43.5 to 19.5 per 1000 and 13.6 to 5.9 per 1000 screenings, respectively.[12] An observational prospective study by Chiarelli et al. on annual MRI screening and mammography in high risk women (BRCA mutation

carriers or >25% lifetime breast cancer risk), showed a CDR for MRI (irrespective of the mammography result) in incidence rounds of 10.6 per 1000 screenings compared to 16.4 per 1000 in the prevalence round.[13]

In our study, women underwent regular mammographic screening two years after the first (prevalence) MRI screening round, and among those who particated in the first screening round, a CDR on mammography of 2.0 per 1000 screenings was observed.[5] Subsequently, among those with a negative mammographic result, an additional CDR on MRI of 5.8 per 1000 in our incidence screening round was observed.

Furthermore, we observed an important decrease of the FPR from 79.8 per 1000 screenings in the first round to 26.3 per 1000 screenings in the second round. This can partly be explained by the availability of previous MRI examinations as comparison which facilitated the reading. Additionally, it can be explained by the fact that the radiologists, although experienced in breast MRI screening of the high-risk population, went through a learning curve in reading breast screening MRIs from average risk women with dense breasts during the DENSE trial. A lower FPR had been observed in incidence rounds of other recent prospective MRI studies as well, but not to the extent that we observed here. In the study of Kuhl et al. in average risk women the rate of false positive examinations was 99.0 per 1000 screenings in the prevalence round and 47.0 per 1000 in the incidence rounds.[9] In the high-risk studies of Saadatmand et al.[10] and Chiarelli et al.[13] the false-positive results on MRI decreased from 123.0 to 51.1 per 1000 screenings and from 205.8 to 96.8 per 1000 screenings, in prevalence and incidence screening rounds respectively. All above-mentioned studies regarded MRI BI-RADS 3, 4 and 5 as positive examinations (Table 3).

In incidence rounds one would expect to find relative more tumors with an aggressive behavior than in the prevalence round.[14] Our results do not provide evidence to support this. The proportion of DCIS among the MRI-detected cancers seems to be relatively higher in the incidence than in the prevalence round (30% vs. 19%; not statistically significant), although all of the detected DCIS were intermediately or poorly differentiated. In addition, the proportion of poorly differentiated invasive cancers did not differ between the first and second screening round. In the second screening round, all 14 invasive cancers were estrogen and/or progesterone receptor positive, compared to 56 of 64 invasive cancers (87.5%) in the first screening round, which was not statistically significantly different.

In contrast, Kuhl et al. found 38% DCIS in the prevalence compared to 15% DCIS in all incidence rounds combined.[9] In general, results of their study suggested that supplemental MRI improved early diagnosis of aggressive breast cancers on MRI compared to mammography only. In the FaMRIsc trial the proportion of DCIS decreased from 60% in the prevalence round of MRI plus mammography screening to 28% in the annual incidence rounds.[10]

Table 3. Summary table of literature

Study	Study size*	Population	Screening interval MRI	CDR (per 1000 screenings)		FPR (per 1000 screenings)		Positive examinations
				Incidence rounds	Prevalence round	Incidence rounds	Prevalence round	BI-RADS
Our study	4,783 (3,436)	Women at average risk and extremely dense breasts	2 years	5.8 (95% CI, 3.8 to 9.0)	16.5 (95% CI, 13.3 to 20.5)	26.3 (95% CI, 21.5 to 32.3)	79.8 (95% CI, 72.4 to 87.9)	3, 4, 5
Kuhl et al.[9]	2,120 (1,741)	Average risk and wide range of breast densities	12 to 36 months	6.9 (95% CI, 3.6 to 12.9)	22.6 (95% CI, 17.1 to 29.9)	47.0	99.0	3, 4, 5
Saadatmand et al.[10]	674	High-risk, but BRCA1, BRCA2 and TP53 wild-type	1 year	4.1	23.8	51.1	123.0	3, 4, 5
Vreemann et al.[11]	2,463	High risk population: (BRCA carriers or family or personal history of breast cancer)	1 year	10.6 (95% CI, 8.3 to 13.1)	22.5 (95% CI, 15.5 to 27.4)	28.4	111.8	0, 3, 4, 5
Bick et al.[12]	4,573	High risk population (BRCA carriers and non-carriers with high risk)	1 year	BRCA1: 28.4 (95% CI, 21.7 to 37.1) BRCA2: 19.5 (95% CI, 12.9 to 29.4) non-carriers: 5.9 (95% CI, 4.3 to 8.0)	BRCA1: 19.9 (95% CI, 12.8 to 30.9) BRCA2: 43.5 (95% CI, 29.8 to 62.9) non-carriers: 13.6 (95% CI, 10.0 to 18.4)	89.0	154.0	0, 3, 4, 5
Chiarelli et al.[13]	8,782	High risk women (BRCA carriers or >25% lifetime breast cancer risk)	1 year	10.6 (95% CI, 8.8 to 12.6)	16.4 (95% CI, 13.9 to 19.3)	96.8	205.8	3, 4, 5

CDR = cancer-detection rate. FPR = false positive rate.
*The study size is the number of women participating in the prevalence screening round, with the number of women participating in the incidence screening round(s) within parenthesis (if numbers were available).

This is the first study that reports on an incidence screening MRI round in women with extremely dense breasts at average risk of breast cancer. Such studies are needed to adequately measure the performance of an ongoing, as opposed to a one-time-only intervention. Although the CDR is, as expected, substantially lower in the second round compared to the first round, still a considerable number of cancers are detected by supplemental MRI screening after a negative mammographic result only two years after the first MRI. It needs to be established whether the optimal frequency of MRI screening would be every two years, as we did in this trial, or that a lower frequency would be almost as effective but more cost-efficient.

The sharp decrease in false positives of MRI screening in the second round compared to the first round is reassuring and may be even further reduced through the use of computer aided diagnosis.[15] The next step, after longer follow-up and linkage with the cancer registry, is to compare the two DENSE trial arms and study the effect of multiple supplemental MRI screening rounds on the interval cancer and late stage cancer rates.

ACKNOWLEDGMENTS

The authors thank the trial participants for their contributions. The authors thank the regional screening organizations, Volpara Solutions, the Dutch Expert Centre for Screening, and the National Institute for Public Health and the Environment for their advice and in-kind contributions for organizing the breast density assessments and the invitation of potential participants.

The DENSE trial is financially supported by the University Medical Center Utrecht (UMC Utrecht, Project number: UMCU DENSE), the Netherlands Organization for Health Research and Development (ZonMw, Project numbers: ZONMW-200320002-UMCU and ZonMW Preventie 50-53125-98-014), the Dutch Cancer Society (KWF Kankerbestrijding, Project numbers: DCS-UU-2009-4348, UU-2014-6859 and UU-2014-7151), the Dutch Pink Ribbon / a Sister's hope (Project number: Pink Ribbon-10074), Bayer AG Pharmaceuticals, Radiology (Project number: BSP-DENSE), and Stichting Kankerpreventie Midden-West. For research purposes, Volpara Health Technologies (Wellington, New Zealand) has provided Volpara Imaging Software 1.5 for installation on servers in the screening units of the Dutch screening program.

REFERENCES

1. Fracheboud J, de Koning HJ, Boer R, et al. Nationwide breast cancer screening programme fully implemented in the Netherlands. *Breast*. 2001;10(1):6-11.
2. Boyd NF, Guo H, Martin LJ, et al. Mammographic Density and the Risk and Detection of Breast Cancer. *The New England Journal of Medicine*. 2007;356:227-236.
3. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95-103.
4. Niell BL, Freer PE, Weinfurtner RJ, Arleo EK, Drukteinis JS. Screening for Breast Cancer. *Radiologic Clinics of North America*. 2017;55(6):1145-1162.
5. Bakker MF, de Lange SV., Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *The New England Journal of Medicine*. 2019;381(22):2091-2102.
6. Emaus MJ, Bakker MF, Peeters PHM, 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.
7. D'Orsi C, Sickles E, Medelson E, Morrison E. *ACR BI-RADS Atlas: Breast Imaging Reporting and Data System*. 5th ed. American College of Radiology; 2013.
8. D'Orsi C, Mendelson E, Ikeda D. *Breast Imaging Reporting and Data System: ACR BI-RADS Breast Imaging Atlas*. American College of Radiology; 2003.
9. 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):361-370.
10. 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. *The Lancet Oncology*. 2019;20(8):1136-1147.
11. Vreemann S, Gubern-Mérida A, Schlooz-Vries MS, et al. Influence of risk category and screening round on the performance of an MR imaging and mammography screening program in carriers of the BRCA mutation and other women at increased risk. *Radiology*. 2018;286(2):443-451.
12. 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 Research and Treatment*. 2019;175(1):217-228.
13. Chiarelli AM, Prummel M v., Muradali D, et al. Effectiveness of screening with annual magnetic resonance imaging and mammography: Results of the initial screen from the Ontario High Risk Breast Screening Program. *Journal of Clinical Oncology*. 2014;32(21):2224-2230.
14. Anderson TJ, Alexander F, Chetty U, et al. Comparative pathology of prevalent and incident cancers detected by breast screening: Edinburgh breast screening project. *The Lancet*. 1986;327(8480):519-523.
15. 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. *Investigative Radiology*. 2020;55(7):438-444.

SUPPLEMENTARY APPENDIX

Members DENSE trial study group

University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands:
CH van Gils PhD, MF Bakker PhD, SV de Lange MD, SGH 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, R Mus MD, M Imhof-Tas MD, N Karssemeijer PhD

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 Center, 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

6

Screening performance of supplemental MRI in women with extremely dense breasts stratified by breast cancer risk

S.V. de Lange, M.F. Bakker, R.M. Pijnappel, E.M. Monninkhof,
M.J. Emaus, C.E. Loo, R.M. Mann, R.H.C. Bisschops, M.B.I. Lobbes, M.J.C.M. Rutten,
K.M. Duvivier, J. Veltman,N. Karssemeijer, H.J. de Koning, P.J. van Diest,
W.P.T.M. Mali, M.A.A.J. van den Bosch, W.B. Veldhuis, C.H. van Gils,
for the DENSE Trial Study Group.

Submitted.

ABSTRACT

Importance

Supplemental MRI after negative mammography screening increases cancer detection and significantly reduces interval cancers in women with extremely dense breasts. However, false positive rates (FPR) and costs also increase compared to mammography alone.

Objective

Here, we investigate whether the harm-benefit balance of supplemental MRI in women with extremely dense breasts is most favorable in women with the highest estimated breast cancer risk.

Design

In the DENSE trial screening participants with extremely dense breasts and negative mammography now completed two rounds of supplemental MRI screening. We categorized participants into quartiles based on Tyrer-Cuzick and Breast Cancer Surveillance Consortium (BCSC) 5-year risk. We also explored stratifications based on age and percent volumetric density.

MRI screening performance measures (rates for recall, biopsy, MRI cancer-detection, FPR and positive predictive value [PPV] of MRI) were calculated for women in different risk quartiles.

Results

In total 4,783 women underwent first screening round MRI, and 3,436 (71.8%) underwent second screening round MRI. MRI cancer-detection rates were higher in the highest quartile (Q4) of Tyrer-Cuzick risk in both screening rounds (p-trend <0.001 and 0.033). FPRs were higher in Q4 of Tyrer-Cuzick risk, without statistically significant trend. PPV was higher in Q4 of Tyrer-Cuzick risk; only statistically significant in the first and combined screening rounds (p-trend 0.002 and <0.001).

In Q4 of BCSC risk, MRI cancer-detection rates were also higher; statistically significant for the second and combined screening rounds (p-trend 0.044 and 0.014). FPRs were lower in Q4 of BCSC risks in both screening rounds (p-trend 0.018 and 0.011). PPV was higher in Q4 of BCSC risks in both screening rounds (p-trend 0.017 and 0.007)

The number of interval cancers was too small to compare between risk groups. Risk stratifications based on age and percent density within women with extremely dense breasts, did not distinguish subgroups with a more optimal harm-benefit balance as well as Tyrer-Cuzick and BCSC risks did.

Conclusion

Breast cancer risks may be used to select women with extremely dense breasts and otherwise increased breast cancer risk for supplemental MRI to improve detection and PPV.

INTRODUCTION

Breast density, the amount of fibroglandular tissue in the breast, is an important independent risk factor for breast cancer. Women with extremely dense breasts have a 3-6 times increased breast cancer risk compared to women with almost entirely fatty breasts, and a 2 times increased risk compared to average women.[1–3] In addition, sensitivity of mammography is reduced in women with dense breasts.[4–6]

In the United States, a federal law directs breast density reporting but provides little guidance on how supplemental screening should address that risk difference.[7,8]

Previous research showed that cancer detection rates increase with the use of supplemental screening modalities in women with dense breasts,[9–11] but recall rates also increase. Recently, the first screening round results of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial showed that additional magnetic resonance imaging (MRI) screening in women with extremely dense breasts leads to significantly lower interval cancer rates.[12] However, false positive results are higher and costs are substantial.

Therefore, it has been proposed that, besides breast density, estimated 5-year or 10-year breast cancer risk should be considered when deciding for supplemental screening.[6,13]

Screening performance of supplemental MRI when offered to women with extremely dense breasts and otherwise increased breast cancer risk, for example based on breast cancer risk prediction models, is unknown.

In the current study we investigate whether the balance between benefits and harms can be improved in supplemental MRI screening. For that, we examined screening performance of supplemental MRI among women with extremely dense breasts when stratified by breast cancer risk.

METHODS

Study Population

The study population of this cohort study consists of women participating in the DENSE trial. The trial design and first screening round results have been described in detail previously.[12,14] DENSE is a multicenter, randomized controlled trial, performed in the Dutch population based-screening program, investigating the additional value of MRI screening in women with extremely dense breasts. Screening participants, aged between 50 and 75 years old, with extremely dense breasts (Volpara density grade (VDG) 4, comparable to American College of Radiology category 4 or d, determined fully automatically using Volpara Imaging Software version 1.5; Volpara Health Technologies, Wellington, New Zealand) and a negative result at mammography were eligible for randomization. Women were randomized in a 1:4 ratio to an invitation for supplemental MRI (MRI arm, intervention arm) or mammography screening only (control arm, regular mammographic screening). After randomization, only women randomized to the intervention arm were invited for the trial.

The trial has been approved by the Dutch Minister of Health, Welfare and Sport (2011/19 WBO, the Hague, the Netherlands). Ethical approval was obtained on November 11, 2011. All participants provided written informed consent before the MRI examination.

Screening performance measures

All MRI examinations were performed on 3.0 Tesla systems and the macrocyclic gadolinium-based contrast agent gadobutrol (Gadovist; Bayer AG, Leverkusen, Germany) was used in all examinations. We calculated the following screening performance measures for the first (prevalence) and second (incidence) screening round and these rounds combined: recall rate, biopsy rate, false positive rate (women with positive MRI who were found not to have breast cancer), MRI cancer-detection rates and positive predictive value (PPV) of positive MRI (BI-RADS 3, 4 or 5). For the first screening round the interval-cancer rate was also calculated. The interval-cancer rate of the second screening round could not be included, because data collection for this end point is still ongoing. All rates were calculated per 1000 screenings.

Measurement of risk factors

Information on breast cancer risk factors was obtained from self-report before the first screening MRI. Women were asked to complete a questionnaire on health status and breast cancer risk factors, including information on personal and family history of (breast) cancer, parity, number of children, age at first birth, age at menarche and menopausal status.

Breast density was measured fully automatically, on the screening mammogram made just before randomization, using Volpara Imaging Software (Volpara Health

Technologies, Wellington, New Zealand). In principle, for each participant volumetric percent density was available for the right and left breast. The mean of both volumetric percent densities (i.e. left and right) was used for analyses. In case measurement for one breast was missing, the available contralateral measurement was used. A continuous Volpara density measurement was not available for the first 40 participants who were randomized between December 2011 and August 2012.

Breast cancer risk models

Breast cancer risks were estimated using the Tyrer-Cuzick (or International Breast Cancer Study [IBIS]) model and the Breast Cancer Surveillance Consortium (BCSC) model.[15,16] The Tyrer-Cuzick model has originally been developed to predict breast cancer risk in women with an elevated risk (i.e. at least a 2-fold relative risk for women aged 45 to 70 years, 4-fold risk for women aged 40 to 44 years and a ten-fold risk for women aged 35 to 39 years.[15,17] The following risk factors are included: age, parity and age at first birth, age at menarche, BMI, age at menopause, use of hormone replacement therapy, number of first, second- and third-degree relatives with breast cancer and ovarian cancer, benign breast disease and mammographic density.

The BCSC model was developed in the Breast Cancer Surveillance Consortium cohort of > 1 million U.S. women presenting for screening mammography.[16] The model predicts breast cancer risk, based on age, race/ethnicity, family history of breast cancer in a first-degree relative, benign breast disease and mammographic density (ACR BI-RADS density).

Risk estimates

We calculated Tyrer-Cuzick 5-year and BCSC 5-year breast cancer risk estimates.

Tyrer-Cuzick breast cancer risk estimates were calculated using a risk evaluation tool for batches (version 8b, UK rates, allowing for competing mortality), with Volpara volumetric density (%).[18] Polygenic risk scores, used in this version, were not available for our population. Missing data were coded according to the specifications of the risk evaluator input file descriptions. Tyrer-Cuzick risk estimates cannot be calculated for women with a previous diagnosis of breast cancer or ductal carcinoma in situ (DCIS).[19]

BCSC breast cancer risk estimates were calculated using the BCSC SAS program (version 2.0).[20] BCSC risk estimates cannot be calculated for women older than 74 years, women with previous diagnosis of breast cancer or DCIS, and for women with previous breast augmentation or mastectomy.[16] Information on ethnicity was not recorded within the DENSE trial. We calculated therefore the risk for non-Hispanic white women in the BCSC model, as in 2014, 90.2% of the women aged 50-55 years in the Netherlands had a Dutch or Western migration background.[21] Women were categorized into quartiles based on the obtained 5-year risk estimate distributions.

In addition, we explored simpler ways of risk stratification, for which less information is needed: women were also categorized into quartiles based on age at randomization, and into quartiles of percentage breast density within VDG 4.

Statistical analysis

Screening performance measures were determined overall and across quartiles of Tyrer-Cuzick 5-year and BCSC 5-year breast cancer risk estimates, quartiles of age distribution and for quartiles of percentage breast density (within VDG 4).

For the performance measures of the first screening round and the second screening round separately, we calculated Wilson's 95% confidence intervals (CIs).

For the measures of the combined screening rounds, rates and 95% CIs were calculated using generalized estimating equations (GEE) with 'independence' correlation structure to account for correlation between the MRI examinations of the same woman.

We tested for linear trends in screening measures across quartiles of different scores using chi squared test for linear trend. Statistical tests were two-sided with significance level of 0.05. Analyses were performed using RStudio (version 1.0.143).

RESULTS

Between December 2011 and January 2016, 4,783 women who were randomized to the intervention arm of the DENSE trial, participated in the first supplemental MRI screening round. Of these first-round participants, 91.6% (N=4,381) underwent consecutive mammography screening 2 years after the first screening round MRI, and 71.8% participated in the second MRI screening round (N=3,436) (Figure 1).

Table 1 lists the characteristics of the first screening round participants overall and stratified per quartile of Tyrer-Cuzick 5-year risk. Tyrer-Cuzick breast cancer risk estimates could be calculated for 94% of all eligible women. Risk estimates are missing for 274 first screening round participants who did not fill out the baseline questionnaire on breast cancer risk factors, and for 27 women who reported a previous history of breast cancer or DCIS.

The median age of women participating in the first screening round was 54 years, interquartile range 51 to 59 years. Of all first-round participants, 15.6% (n=746) reported a positive family history of breast cancer (first-degree relative: mother, sister or daughter).

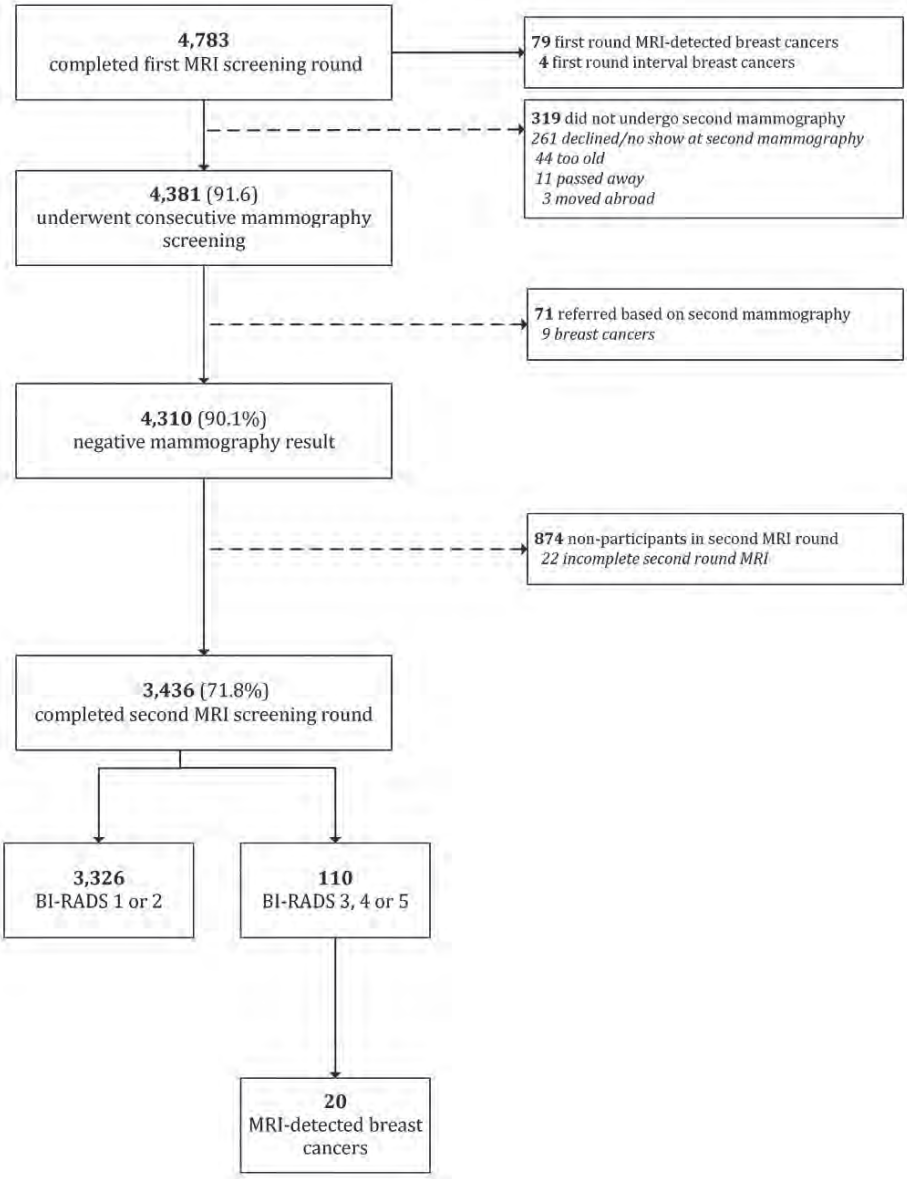


Figure 1. Participation second (incidence) MRI screening round.

Table 1. Characteristics of MRI screening participants for quartiles of Tyrer-Cuzick 5-year risk estimates

	Overall	**Q1**	**Q2**	**Q3**	**Q4**
		<1.94%	1.94-2.66%	2.66-3.83%	>3.83%
Characteristics	N=4,783	N=1,121	N=1,120	N=1,120	N=1,121
Age group					
49-54 years	2,635 (55.1)	744 (66.4)	657 (58.7)	578 (51.6)	509 (45.4)
55-59 years	993 (20.8)	212 (18.9)	207 (18.5)	234 (20.9)	279 (24.9)
60-64 years	602 (12.6)	78 (7.0)	126 (11.2)	157 (14.0)	198 (17.7)
65-69 years	410 (8.6)	60 (5.4)	92 (8.2)	114 (10.2)	111 (9.9)
70-74 years	133 (2.8)	25 (2.2)	37 (3.3)	35 (3.1)	21 (1.9)
≥75	10 (0.2)	2 (0.2)	1 (0.1)	2 (0.2)	3 (0.3)
Age at menarche, yrs.					
<12	364 (8.2)	70 (6.3)	90 (8.2)	84 (7.7)	118 (10.7)
12-13	1,988 (45.0)	411 (37.2)	508 (46.4)	533 (48.8)	525 (47.8)
≥14	2,065 (46.8)	624 (56.5)	496 (45.3)	476 (43.5)	455 (41.4)
Missing	*366*	*16*	*26*	*27*	*23*
BMI (kg/m2)					
<18.5	216 (4.9)	125 (11.3)	35 (3.2)	27 (2.5)	28 (2.6)
18.5-24.9	3,694 (83.1)	969 (86.4)	995 (89.9)	883 (80.5)	824 (75.3)
25.0-29.9	484 (10.9)	25 (2.2)	76 (6.9)	168 (15.3)	212 (19.4)
≥30	51 (1.1)	1 (0.1)	1 (0.1)	19 (1.7)	30 (2.7)
Missing	*338*	*1*	*13*	*23*	*27*
Menopausal status [a]					
Premenopausal	473 (10.3)	68 (6.1)	130 (11.6)	136 (12.2)	139 (12.4)
Perimenopausal	1,337 (29.2)	362 (32.7)	348 (31.5)	335 (29.3)	291 (25.3)
Postmenopausal	2,768 (60.5)	690 (61.1)	639 (56.6)	646 (58.6)	691 (62.9)
Missing	*205*	*1*	*3*	*3*	*0*

Table 1. Continued.

Characteristics	**Overall** N=4,783	**Q1** <1.94% N=1,121	**Q2** 1.94-2.66% N=1,120	**Q3** 2.66-3.83% N=1,120	**Q4** >3.83% N=1,121
History of biopsies					
No prior biopsy	3,738 (83.3)	958 (85.9)	959 (86.1)	927 (83.3)	891 (79.5)
Prior biopsy, result unknown	729 (16.2)	157 (14.1)	155 (13.9)	186 (16.7)	213 (19.0)
Atypical hyperplasia	23 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	17 (1.5)
Missing	*293*	*6*	*6*	*7*	*0*
First-degree relatives with breast cancer [b]					
Yes	746 (15.6)	12 (1.1)	25 (2.2)	107 (9.6)	596 (53.2)
No	2,081 (43.5)	632 (56.4)	610 (54.5)	561 (50.1)	270 (24.1)
Unknown	1,956 (40.9)	477 (42.6)	485 (43.3)	452 (40.4)	255 (22.7)
Parity					
Nulliparous	1,051 (23.4)	210 (18.7)	236 (21.1)	279 (24.9)	314 (28.0)
1 birth	540 (12.0)	132 (11.8)	125 (11.2)	140 (12.5)	142 (12.7)
≥2 births	2,908 (64.6)	778 (69.4)	756 (67.5)	695 (62.1)	665 (59.3)
Missing	*284*	*1*	*3*	*6*	*0*
HRT use [c]					
Never	3,325 (73.9)	862 (77.0)	836 (74.9)	809 (72.6)	797 (71.1)
Past use of HRT	550 (12.2)	98 (8.8)	128 (11.5)	154 (13.8)	165 (14.7)
Current use of HRT	623 (13.9)	160 (14.3)	152 (13.6)	151 (13.6)	159 (14.2)
Missing	*285*	*1*	*4*	*6*	*0*

Unless otherwise noted, data are numbers of women, with vertically calculated percentages in parenthesis. Total of percentages may not equal 100% due to rounding.

Tyrer-Cuzick 5-year risk scores were not available for 274 women who did not complete the baseline questionnaire, and 27 women who reported a previous history of breast cancer or ductal carcinoma in situ.

[a] Women aged ≥60 years, or reporting history of hysterectomy 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.

[b] First degree-relatives (mother, sister, daughter) with breast cancer.

[c] HRT includes use of hormonal contraceptives at age of 50 years and above.

The prevalence of women with a positive family history of breast cancer (first degree relative), women with prior history of breast biopsy, women with a higher BMI and younger age at menarche is highest in the highest quartile (Q4) of Tyrer-Cuzick 5-year risk (Table 1).

Screening performance measures for quartiles of risk

Screening performance measures for the first, second and combined screening rounds, overall and stratified per quartile of Tyrer-Cuzick 5-year risk estimates, are listed in Table 2. In the first and second screening round, MRI cancer-detection rates increased for increasing quartiles of Tyrer-Cuzick 5-year risk; all with statistically significant linear trends (p-trend <0.001 and 0.033 respectively). False positive rates increased for increasing quartiles of Tyrer-Cuzick risks, but the trend over quartiles was not statistically significant. PPV of MRI increased for increasing quartiles of Tyrer-Cuzick risks, however the trend was only statistically significant in the first screening round and combined screening rounds (p-trend 0.002 and <0.001 respectively). Interval cancer rates are as yet only available after the first, but not the second screening round. With only 4 interval cancers in total after supplemental MRI, this number was too low for meaningful statistical comparison.

Screening performance measures across quartiles of BCSC 5-year risks are listed in Table 3. MRI cancer-detection rates increased for increasing quartiles of BCSC risks, with significant linear trends for the second screening round and combined screening rounds (p-trend 0.044 and 0.014 respectively). In both screening rounds and combined screening rounds, false positive rates decreased for increasing quartiles of BCSC risks, all with a significant linear trend (p-trend 0.018, 0.011 and 0.002 respectively). PPV of MRI increased for increasing quartiles of BCSC risk, all with a significant linear trend (p-trend 0.017, 0.007 and <0.001 for the first, second and combined screening rounds respectively).

In search for simpler methods of stratification we also stratified by quartiles of age at randomization. Here, MRI cancer-detection rates did not show a statistically significant linear trend across increasing quartiles of age (Table 4). Recall rates, biopsy rates and false positive rates decreased with increasing quartiles of age in the first, second and combined screening rounds (Table 4); all with a significant linear trend (p-trend 0.003 for biopsy rate in the second round and p<0.001 for all other rates). In both screening rounds and combined rounds, PPV of MRI increased with increasing age quartiles; all with a significant linear trend (p-trend 0.015, 0.02 and 0.001 respectively).

Table 2. Screening performance overall and per quartile of Tyrer-Cuzick 5-year risk estimates

	Overall			Q1 <1.94%			Q2 1.94-2.66%			Q3 2.66-3.83%			Q4 >3.83%			P trend
	n	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	
First screening round		(N=4,783)			(N=1,121)			(N=1,120)			(N=1,120)			(N=1,121)		
Recall	*454*	94.9	(86.9-103.6)	*93*	83.0	(68.2-100.6)	*107*	95.5	(79.7-114.2)	*104*	92.9	(77.2-111.3)	*121*	107.9	(91.1-127.5)	0.065
Biopsies	*300*	62.7	(56.2-70.0)	*58*	51.7	(40.2-66.3)	*69*	61.6	(49.0-77.2)	*71*	63.3	(50.5-79.1)	*85*	75.9	(61.8-92.9)	**0.023**
False positives	*375*	79.8	(72.4-87.9)	*84*	75.6	(61.5-92.7)	*91*	82.5	(67.7-100.2)	*88*	79.8	(65.2-97.3)	*91*	83.5	(68.5-101.4)	0.567
MRI screen-detected cancers	*79*	16.5	(13.3-20.5)	*9*	8.0	(4.2-15.2)	*16*	14.3	(8.8-23.1)	*16*	14.3	(8.8-23.1)	*30*	26.8	(18.8-37.9)	**<0.001**
Interval cancers	*4*	0.8	(0.3-2.1)	*1*	0.9	(0.2.-5.0)	*1*	0.9	(0.2-5.0)	*1*	0.9	(0.2-5.0)	*1*	0.9	(0.2-5.0)	-
PPV of positive MRI		17.4	(14.2-21.2)		9.7	(5.2-17.4)		14.9	(9.4-22.9)		15.4	(9.7-23.5)		24.8	(17.9-33.2)	**0.004**
Second screening round		(N=3,436)			(N=808)			(N=827)			(N=809)			(N=830)		
Recall	*110*	32.0	(26.6-38.4)	*20*	24.8	(16.1-37.9)	*26*	31.4	(21.5-45.7)	*23*	28.4	(19.0-42.3)	*38*	45.8	(33.5-62.2)	**0.030**
Biopsies	*85*	24.7	(20.0-30.5)	*17*	21.0	(13.2-33.4)	*18*	21.8	(13.8-34.1)	*15*	18.5	(11.3-30.4)	*32*	38.5	(27.4-53.9)	**0.041**
False positives#	*90*	26.3	(21.5-32.3)	*18*	22.3	(14.2-35.0)	*23*	27.9	(18.7-41.5)	*17*	21.2	(13.3-33.6)	*30*	36.5	(25.7-51.6)	0.157
MRI screen-detected cancers	*20*	5.8	(3.8-9.0)	*2*	2.5	(0.7-9.0)	*3*	3.6	(1.2-10.6)	*6*	7.4	(3.4-16.1)	*8*	9.6	(4.9-18.9)	**0.033**
PPV of positive MRI		18.2	(12.1-26.4)		10.0	(2.8-30.1)		11.5	(4.0-29.0)		26.1	(12.5-46.5)		21.1	(11.1-36.3)	0.181
Combined screening rounds		(N=8,219)			(N=1,929)			(N=1,947)			(N=1,929)			(N=1,951)		
Recall	*564*	68.6	(63.4-74.3)	*113*	58.6	(48.8-70.2)	*133*	68.3	(57.9-80.5)	*127*	65.8	(55.6-77.9)	*159*	81.5	(69.9-94.9)	**0.010**
Biopsies	*384*	46.7	(42.4-51.5)	*75*	38.9	(31.0-48.7)	*87*	44.7	(36.2-55.1)	*86*	44.6	(36.2-54.8)	*117*	60.0	(50.3-71.4)	**0.003**
False positives#	*465*	57.3	(52.4-62.5)	*102*	53.2	(43.9-64.4)	*114*	59.1	(49.5-70.5)	*105*	55.1	(45.6-66.3)	*121*	63.3	(53.0-75.3)	0.272
MRI screen-detected cancers	*99*	12.0	(9.9-14.6)	*11*	5.7	(3.2 -10.3)	*19*	9.8	(6.2-15.3)	*22*	11.4	(7.5-17.3)	*38*	19.5	(14.2-26.7)	**<0.001**
PPV of positive MRI		17.6	(14.6-20.9)		9.7	(5.5-16.8)		14.3	(9.4-21.2)		17.3	(11.7-24.9)		23.9	(17.9-31.1)	**0.001**

MRI= magnetic resonance imaging; PPV= positive predictive value.
#Interval cancers are not taken into account in calculation of false positive rate for second screening round and combined screening rounds.
For 274 women Tyrer-Cuzick 5-year risk was not calculated because these women did not complete the baseline questionnaire and for another 27 women Tyrer-Cuzick 5-year risk was not calculated because of previous history of breast cancer or ductal carcinoma in situ diagnosis.

Table 3. Screening performance overall and per quartile of BCSC 5-year risk estimates

	Overall			Q1 <1.68%			Q2 1.68-2.07%			Q3 2.07-2.51%			Q4 >2.51%			P trend
	n	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	
First screening round		(N=4,783)			(N=1,090)			(N=1,136)			(N=1,109)			(N=1,133)		
Recall	*454*	94.9	(86.9-103.6)	*124*	113.8	(96.3-134.0)	*98*	86.3	(71.3-104.0)	*93*	83.9	(68.9-101.6)	*107*	94.4	(78.8-112.9)	0.132
Biopsies	*300*	62.7	(56.2-70.0)	*79*	72.5	(58.5-89.4)	*62*	54.6	(42.8-69.4)	*65*	58.6	(46.3-74.0)	*75*	66.2	(53.1-82.2)	0.670
False positives	*375*	79.8	(72.4-87.9)	*107*	99.9	(83.4-119.3)	*87*	77.4	(63.2-94.5)	*76*	69.6	(56.0-86.2)	*81*	73.2	(59.3-90.1)	**0.018**
MRI screen-detected cancers	*79*	16.5	(13.3-20.5)	*17*	15.6	(9.8-24.8)	*11*	9.7	(5.4-17.3)	*17*	15.3	(9.6-24.4)	*26*	22.9	(15.7-33.4)	0.093
Interval cancers	*4*	0.8	(0.3-2.1)	*2*	1.8	(0.5-6.7)	*1*	0.9	(0.2-5.0)	*0*	0	(0-3.5)	*1*	0.8	(0.2-5.0)	0.359
PPV of positive MRI		17.4	(14.2-21.2)		13.7	(8.7-20.9)		11.2	(6.4-19.0)		18.3	(11.7-27.3)		24.3	(17.2-33.2)	**0.017**
Second screening round		(N=3,436)			(N=797)			(N=859)			(N=824)			(N=790)		
Recall	*110*	32.0	(26.6-38.4)	*37*	46.4	(33.9-63.3)	*24*	27.9	(18.8-41.2)	*19*	23.1	(14.8-35.7)	*27*	34.2	(23.6-49.3)	0.140
Biopsies	*85*	24.7	(20.0-30.5)	*28*	35.1	(24.4-50.3)	*18*	21.0	(13.3-32.9)	*12*	14.6	(8.4-25.3)	*24*	30.4	(20.5-44.8)	0.401
False positives#	*90*	26.3	(21.5-32.3)	*32*	40.4	(28.8-56.5)	*24*	27.9	(18.8-41.2)	*15*	18.3	(11.1-30.0)	*17*	21.8	(13.7-34.6)	**0.011**
MRI screen-detected cancers	*20*	5.8	(3.8-9.0)	*5*	6.3	(2.7-14.6)	*0*	0	(0-4.5)	*4*	4.9	(1.9-12.4)	*10*	12.7	(6.9-23.1)	**0.044**
PPV of positive MRI		18.2	(12.1-26.4)		13.5	(5.9-28.0)		0	(0-13.8)		21.1	(8.5-43.3)		37.0	(21.5-55.8)	**0.007**
Combined screening rounds		(N=8,219)			(N=1,887)			(N=1,995)			(N=1,933)			(N=1,923)		
Recall	*564*	68.6	(63.4-74.3)	*161*	85.3	(73.5-98.9)	*122*	61.2	(51.2-72.9)	*112*	57.9	(48.3-69.4)	*134*	69.7	(58.9-82.2)	0.056
Biopsies	*384*	46.7	(42.4-51.5)	*107*	56.7	(46.9-68.4)	*80*	40.1	(32.3-49.7)	*77*	39.8	(31.9-49.6)	*99*	51.5	(42.4-62.3)	0.481
False positives#	*465*	57.3	(52.4-62.5)	*139*	74.5	(63.4-87.4)	*111*	55.9	(46.4-67.3)	*91*	47.6	(39.0-57.9)	*98*	51.9	(42.6-63.1)	**0.002**
MRI screen-detected cancers	*99*	12.0	(9.9-14.6)	*22*	11.7	(7.7-17.7)	*11*	5.5	(3.1-9.9)	*21*	10.9	(7.1-16.6)	*36*	18.7	(13.5-25.9)	**0.014**
PPV of positive MRI		17.6	(14.6-20.9)		13.7	(9.2-19.9)		9.0	(5.1-15.6)		18.8	(12.7-26.9)		26.9	(20.0-35.0)	**<0.001**

BCSC=Breast Cancer Surveillance Consortium; MRI= magnetic resonance imaging; PPV= positive predictive value.
#Interval cancers are not taken into account in calculation of false positive rate for second screening round and combined screening rounds.
For 274 women BCSC 5-year risk was not calculated because these women did not complete the baseline questionnaire, for an additional 41 women, BCSC 5-year risk could not be calculated; for 8 women (0.2%) because of age (>74 years), 27 women (0.6%) because of previous breast cancer or ductal carcinoma in situ diagnosis and for 6 women (0.1%) because of previous breast augmentation.

Table 4. Screening performance overall and per quartile of age at randomization

	Overall			Q1 49-50			Q2 51-53			Q3 54-58			Q4 59-76			P trend
	n	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	
First screening round		(N=4,783)			(N=1,018)			(N=1,297)			(N=1,156)			(N=1,312)		
Recall	*454*	94.9	(86.9-103.6)	*126*	123.8	(104.9-145.4)	*142*	109.5	(93.6-127.6)	*102*	88.2	(73.2-106.0)	*84*	69.3	(56.3-85.0)	**<0.001**
Biopsies	*300*	62.7	(56.2-70.0)	*84*	82.5	(67.1-101.0)	*86*	66.3	(54.0-81.2)	*69*	59.7	(47.4-74.9)	*61*	50.3	(39.4-64.1)	**<0.001**
False positives	*375*	79.8	(72.4-87.9)	*107*	107.2	(89.5-127.9)	*123*	96.3	(81.3-113.7)	*85*	74.8	(60.9-91.5)	*60*	46.6	(36.4-59.5)	**<0.001**
MRI screen-detected cancers	*79*	16.5	(13.3-20.5)	*19*	18.7	(12.0-29.0)	*19*	14.7	(9.4-22.8)	*17*	14.7	(9.2-23.4)	*24*	18.3	(12.3-27.1)	0.957
Interval cancers	*4*	0.8	(0.3-2.1)	*1*	1.0	(0.2-5.5)	*1*	0.8	(0.1-4.4)	*2*	1.7	(0.5-6.3)	*0*	0.0	(0.0-2.9)	0.556
PPV of positive MRI		17.4	(14.2-21.2)		15.1	(9.9-22.4)		13.4	(8.7-20.0)		16.7	(10.7-25.1)		19.4	(13.4-27.2)	**0.015**
Second screening round		(N=3,436)			(N=729)			(N=970)			(N=860)			(N=877)		
Recall	*110*	32.0	(26.6-38.4)	*34*	46.6	(33.6-64.5)	*36*	37.1	(26.9-51.0)	*23*	26.7	(17.9-39.8)	*17*	19.4	(12.1-30.8)	**<0.001**
Biopsies	*85*	24.7	(20.0-30.5)	*25*	34.3	(23.2-50.1)	*29*	29.9	(20.9-42.6)	*18*	20.9	(12.3-32.8)	*12*	13.7	(7.8-23.8)	**0.003**
False positives#	*90*	26.3	(21.5-32.3)	*29*	40.4	(28.0-56.9)	*32*	33.1	(23.6-46.4)	*20*	23.3	(15.2-35.8)	*9*	10.4	(5.5-19.6)	**<0.001**
MRI screen-detected cancers	*20*	5.8	(3.8-9.0)	*5*	6.9	(2.9-15.9)	*4*	4.1	(1.6-10.6)	*3*	3.5	(1.2-10.2)	*8*	9.1	(4.6-17.9)	0.532
PPV of positive MRI		18.2	(12.1-26.4)		14.7	(6.5-30.1)		11.1	(4.4-25.3)		13.0	(4.5-32.1)		47.1	(26.2-69.0)	**0.02**
Combined screening rounds		(N=8,219)			(N=1,747)			(N=2,267)			(N=2,016)			(N=2,189)		
Recall	*564*	68.6	(63.4-74.3)	*160*	91.6	(78.8-106.3)	*178*	78.5	(68.1-90.3)	*125*	62.0	(52.1-73.6)	*101*	46.1	(37.9-56.0)	**<0.001**
Biopsies	*384*	46.7	(42.4-51.5)	*109*	62.4	(51.7-60.5)	*115*	50.7	(42.4-60.5)	*87*	43.2	(35.2-52.8)	*73*	33.3	(26.4-42.0)	**<0.001**
False positives#	*465*	57.3	(52.4-62.5)	*136*	78.9	(67.1-92.7)	*155*	69.1	(59.3-80.3)	*105*	52.6	(43.5-63.6)	*69*	32.0	(25.3-40.4)	**<0.001**
MRI screen-detected cancers	*99*	12.0	(9.9-14.6)	*24*	13.7	(9.2-20.4)	*23*	10.2	(6.7-15.2)	*20*	9.9	(6.4-15.3)	*32*	14.6	(10.4-20.6)	0.711
PPV of positive MRI		17.6	(14.6-20.9)		15.0	(10.3-21.3)		12.9	(8.7-18.7)		16.0	(10.5-23.6)		31.7	(23.5-41.2)	**0.001**

MRI= magnetic resonance imaging; PPV= positive predictive value.

#Interval cancers are not taken into account in calculation of false positive rate for second screening round and combined screening rounds.

When we stratified by quartiles of percent density within Volpara density grade 4, MRI cancer-detection rates were lower in higher quartiles of percent density in the first screening round and combined rounds; linear trend was only statistically significant in the first screening round (p-trend 0.047) (Supplemental table 1). Recall rate, biopsy rate, false positive rate and PPV of MRI decreased in the first screening round with increasing density quartiles, this trend was only statistically significant for biopsy rate (p-trend for biopsy 0.02, and p-trend 0.195, 0.594 and 0.116 for recall, false positives and PPV). In the second screening round, rates increased with increasing density with significantly higher recall and false positive rates in Q4 of density (p-trend 0.039 and 0.017 respectively), and an increase of biopsy rate over quartiles of density, without statistically significant trend (p-trend 0.154). As a result, PPV of MRI decreased across quartiles of density, without a statistically significant trend (p-trend 0.823).

DISCUSSION

MRI cancer-detection rates of supplemental MRI in women with extremely dense breasts were significantly higher in women in the highest quartiles of Tyrer-Cuzick 5-year breast cancer risk than in women in the lowest quartile, in both screening rounds separately and combined. In addition, PPV of MRI was higher in Q4 of Tyrer-Cuzick 5-year risk in both screening rounds and combined screening rounds. However, the linear trend was only significant in the first round and combined screening rounds.

The BCSC model also showed higher MRI cancer-detection rates in highest risk quartiles, however the trend was only statistically significant for second and combined screening rounds. Interestingly, the false positive rate was lower in the highest quartile, which was not the case when using the TC risk estimates, and therefore, PPV of MRI was significantly higher in the highest quartile of BCSC risk in both screening rounds and combined rounds.

In case we had only offered supplemental MRI to the women in Q4 of Tyrer-Cuzick risk, we would have performed 25% of the MRI examinations that were performed in the trial, this would have resulted in the diagnosis of 38.4% (38/99) of the breast cancers against 26.0% (121/465) of the false positive results. For the BCSC model, performance of 25% of the MRI examinations (Q4) would have resulted in 36.4% (36/99) of the breast cancer diagnosis, against 21.1% (98/465) of false positive results.

To our knowledge there are no other studies yet, comparing the effect of supplemental screening in women with dense breasts across categories of breast cancer risk. In earlier, observational studies based on the BCSC registries (cohort of women who underwent yearly digital screening mammography between 2005 and 2014) however,

Kerlikowske et al.[6,13] showed that the highest interval-cancer rates were found in women with extremely dense breasts who also have a 5-year BCSC risk ≥1.67% (which is 47.5% of the women with extremely dense breasts in their population).[6] In a more recent study, they showed that advanced-cancer rates were highest in women with extremely dense breasts and 5-year BCSC risk ≥1.0% (which is 80.2% of the women with extremely dense breasts in their population).[13]

In our study we did not use the cut-off points that are regularly used in the literature for elevated (≥1.67%) and high (≥3.00%) risk.[6,13] This was because risk estimates in our population of women with extremely dense breasts were on average already higher than 1.67% with approximately three-quarters of our population at elevated risk.

Our results show that risk models might be used in further tailoring supplemental MRI screening based on breast cancer risk scores on top of breast density, to improve detection rates and positive predictive value. It should be noted, however, that the Tyrer-Cuzick model requires detailed information on breast cancer risk factors,[15] and such information is currently not available in most population-based screening programs. The BCSC model[16] is a simpler model, for which less information is required. This model did not show significantly higher MRI cancer-detection rates in the first screening round, but did so in the second, and interestingly false positive results decreased for the highest quartile of BCSC risk. Nevertheless, this model also requires information that is currently not available within organized screening programs.

Therefore, we also assessed whether age or percent breast density could be used to select women with more favorable MRI detection rates and PPVs in our study population. MRI screen-detected cancers did not show any trend for quartiles of increasing age, however. When stratifying for percent density within the extremely dense category, breast cancer detection appeared to decrease and false positives appeared to increase in the higher quartile. Thus, stratification based on age or density do not seem to be good alternatives for the risk-based scores.

A limitation of our study is that we were not able to assess the effect of risk stratification on reduction in interval cancers and advanced cancers and eventually mortality. Because information on breast cancer risk factors, besides density and age, is not available for the control arm due to the design of the trial, we could not compare interval-cancer rates stratified by breast cancer risks between the intervention and control arm. Besides that, the number of interval cancers in the first round was too low for meaningful statistical comparison between risk-based quartiles within the MRI arm, and data on interval cancers in the second screening round are not yet available.

On the basis of the results of Kerlikowkse et al.[16] who did show the highest risk of interval cancers and advanced cancers in women with high breast cancer risk scores participating in a yearly digital mammography screening program, one might argue that also in our trial the women with highest breast cancer risk scores would have the highest risk of interval cancers with mammography screening and therefore likely the largest benefit of MRI screening.

Both risk models used in this study have been externally validated,[19,22–24] but never in a Dutch population. They use different risk factors, but their discriminatory accuracy is in the same range with an area under the curve of 0.66 and 0.64 for the BCSC and Tyrer-Cuzick model (version without polygenic risk scores) respectively.[22,24] The area under the curve of these models is not perfect, and therefore the use of these models might result in over- or underestimation of the actual breast cancer risk. If the performance of these risk models could be further improved, the benefit of stratification would increase, although perhaps at the cost of collecting even more risk factor information.

The benefit of stratification by the BCSC model might be underestimated since information on ethnicity was not available in our study.

To our knowledge this is the first study to assess screening performance measures of supplemental MRI screening in women with extremely dense breasts and otherwise increased breast cancer risk. Our study showed that among women with extremely dense breasts, breast cancer risk estimates may be used to select subgroups of women who could benefit more from supplemental screening. However, the use of breast cancer risk models in a screening setting is not straightforward, and the effect on interval cancers is still unknown.

REFERENCES

1. Boyd NF, Guo H, Martin LJ, et al. Mammographic Density and the Risk and Detection of Breast Cancer. *The New England Journal of Medicine*. 2007;356:227-236.
2. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiology Biomarkers and Prevention*. 2006;15(6):1159-1169.
3. Pettersson A, Graff RE, Ursin G, et al. Mammographic density phenotypes and risk of breast cancer: A meta-analysis. *Journal of the National Cancer Institute*. 2014;106(5).
4. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95-103.
5. Weigel S, Heindel W, Heidrich J, Hense HW, Heidinger O. Digital mammography screening: sensitivity of the programme dependent on breast density. *European Radiology*. 2017;27(7):2744-2751.
6. Kerlikowske K, Zhu W, Tosteson ANA, et al. Identifying women with dense breasts at high risk for interval cancer a cohort study. *Annals of Internal Medicine*. 2015;162(10):673-681.
7. DenseBreast-Info. Is there a federal law for breast density? Published March 28, 2019. Accessed June 10, 2020. https://densebreast-info.org/is-there-a-federal-law.aspx
8. Siu AL. Screening for breast cancer: U.S. Preventive services task force recommendation statement. *Annals of Internal Medicine*. 2016;164(4):279-296.
9. 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: The Journal of the American Medical Association*. 2012;307(13):1394-1404.
10. 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: The Journal of the American Medical Association*. 2008;299(18):2151-2163.
11. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): A randomised controlled trial. *The Lancet*. 2016;387(10016):341-348.
12. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *The New England Journal of Medicine*. 2019;381(22):2091-2102.
13. Kerlikowske K, Sprague BL, Tosteson ANA, et al. Strategies to Identify Women at High Risk of Advanced Breast Cancer during Routine Screening for Discussion of Supplemental Imaging. *JAMA Internal Medicine*. 2019;179(9):1230-1239. doi:10.1001/jamainternmed.2019.1758
14. Emaus MJ, Bakker MF, Peeters PHM, 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.
15. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in Medicine*. 2004;23(7):1111-1130.

16. 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. *Journal of Clinical Oncology*. 2015;33(28):3137-3143.
17. Cuzick J, Forbes J, Edwards R, et al. First results from the International Breast Cancer Intervention Study (IBIS-I): A randomised prevention trial. *The Lancet*. 2002;360(9336):817-824. doi:10.1016/S0140-6736(02)09962-2
18. Risk Evaluator Software. Published September 17, 2017. Accessed June 14, 2020. http://www.ems-trials.org/riskevaluator/
19. Brentnall AR, Harkness EF, Astley SM, et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Research*. 2015;17(1):147.
20. NCI-funded Breast Cancer Surveillance Consortium (P01 CA154292 and HHSN261201100031C). BCSC program (Breast Cancer Surveillance Consortium). Accessed June 14, 2020. https://tools.bcsc-scc.org/BC5yearRisk/
21. StatLine (Centraal bureau voor Statistiek). Population; sex, age, migration background and generation, 1 January. Published June 24, 2019. Accessed June 14, 2020. https://opendata.cbs.nl/statline/#/CBS/en/dataset/37325eng/table?dl=3B25A
22. Vachon CM, Pankratz VS, Scott CG, et al. The Contributions of Breast Density and Common Genetic Variation to Breast Cancer Risk. *Journal of the National Cancer Institute*. 2015;107(5):397.
23. Tice JA, Bissell MCS, Miglioretti DL, et al. Validation of the breast cancer surveillance consortium model of breast cancer risk. *Breast Cancer Research and Treatment*. 2019;175(2):519-523.
24. McCarthy AM, Guan Z, Welch M, et al. Performance of breast cancer risk-assessment models in a large mammography cohort. *Journal of the National Cancer Institute*. 2020;112:489-497.

SUPPLEMENTARY TABLE

Supplemental Table 1. Screening performance overall and per quartile of Volpara percent density within density grade 4 (>15.5%)

	Overall			Q1 <16.95%			Q2 16.95-19.04%			Q3 19.05-22.14%			Q4 >22.15%			P trend
	n	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	*n*	rate	(95%CI)	
First screening round		(N=4,783)			(N=1,128)			(N=1,129)			(N=1,124)			(N=1,128)		
Recall	*454*	94.9	(86.9-103.6)	*126*	105.4	(89.3-124.1)	*105*	89.7	(74.6-107.4)	*105*	88.9	(74.0-106.5)	*107*	89.5	(74.6-107.0)	0.195
Biopsies	*300*	62.7	(56.2-70.0)	*89*	74.5	(60.9-90.8)	*73*	62.3	(49.9-77.7)	*62*	52.5	(41.2-66.7)	*69*	57.7	(45.8-72.4)	**0.02**
False positives	*375*	79.8	(72.4-87.9)	*98*	84.0	(69.5-101.4)	*86*	74.7	(60.8-91.3)	*89*	76.5	(62.6-93.2)	*91*	77.2	(63.3-93.8)	0.594
MRI screen-detected cancers	*79*	16.5	(13.3-20.5)	*28*	23.4	(16.3-33.7)	*19*	16.2	(10.4-25.2)	*16*	13.5	(8.4-21.9)	*16*	13.4	(8.3-21.6)	**0.047**
Interval cancers	*4*	0.8	(0.3-2.1)	*1*	0.8	(0.1-4.7)	*0*	0	(0.0-3.3)	*2*	1.7	(0.5-6.2)	*1*	0.8	(0.1-4.7)	0.7
PPV of positive MRI		17.4	(14.2-21.2)		22.2	(15.8-30.2)		18.1	(11.9-26.5)		15.2	(9.6-23.3)		15.0	(9.4-22.9)	0.116
Second screening round		(N=3,436)			(N=816)			(N=831)			(N=814)			(N=828)		
Recall	*110*	32.0	(26.6-38.4)	*15*	17.5	(10.7-28.7)	*28*	33.4	(23.2-47.8)	*38*	45.1	(33.1-61.3)	*28*	32.2	(22.4-46.2)	
Biopsies	*85*	24.7	(20.0-30.5)	*14*	16.4	(9.8-27.2)	*22*	26.2	(17.4-39.4)	*25*	29.7	(20.2-43.5)	*23*	26.5	(17.7-39.4)	**0.039**
False positives#	*90*	26.3	(21.5-32.3)	*9*	10.6	(5.6-20.0)	*24*	28.7	(19.4-42.4)	*33*	39.4	(28.2-54.9)	*23*	26.7	(17.8-39.7)	0.154
MRI screen-detected cancers	*20*	5.8	(3.8-9.0)	*6*	7.0	(3.2-15.2)	*4*	4.8	(1.9-12.2)	*5*	5.9	(2.5-13.8)	*5*	5.8	(2.5-13.4)	**0.017**
PPV of positive MRI		18.2	(12.1-26.4)		40.0	(19.8-64.3)		14.3	(5.7-31.5)		13.2	(5.8-27.3)		17.9	(7.9-35.6)	0.823
Combined screening rounds		(N=8,219)			(N=1,944)			(N=1,960)			(N=1,938)			(N=1,956)		
Recall	*564*	68.6	(63.4-74.3)	*141*	68.8	(58.4-80.8)	*133*	66.2	(56.0-78.0)	*143*	70.7	(60.1-83.0)	*135*	65.4	(55.4-77.0)	0.818
Biopsies	*384*	46.7	(42.4-51.5)	*103*	50.2	(41.5-60.7)	*95*	47.3	(38.7-57.6)	*87*	43.0	(34.9-53.0)	*92*	44.6	(36.5-54.4)	0.306
False positives#	*465*	57.3	(52.4-62.5)	*107*	53.1	(44.1-63.8)	*110*	55.4	(46.1-66.3)	*122*	60.9	(51.1-72.5)	*114*	55.8	(46.6-66.8)	0.552
MRI screen-detected cancers	*99*	12.0	(9.9-14.6)	*34*	16.6	(11.9-23.2)	*23*	11.4	(7.6-17.2)	*21*	10.4	(6.8-15.9)	*21*	10.2	(6.6-15.6)	0.061
PPV of positive MRI		17.6	(14.6-20.9)		24.1	(17.8-31.8)		17.3	(11.8-24.6)		14.7	(9.8-21.5)		15.6	(10.4-22.7)	**0.0496**

MRI= magnetic resonance imaging; PPV= positive predictive value.

#Interval cancers are not taken into account in calculation of false positive rate for second screening round and combined screening rounds.

SUPPLEMENTARY APPENDIX

Members DENSE trial study group

University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands:
CH van Gils PhD, MF Bakker PhD, SV de Lange MD, SGH 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, R Mus MD, M Imhof-Tas MD, N Karssemeijer PhD

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 Center, 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

7

General discussion

From age-based, or "one size fits all", to tailored breast cancer screening?

The Dense tissue and Early breast Neoplasm ScrEening (DENSE) trial is the first randomized controlled trial to study the effect of supplemental MRI screening in women with extremely dense breasts.

In the first MRI screening round, the cancer-detection rate of MRI after negative mammography was 16.5 per 1000 screenings. The results of the first round of the trial showed that supplemental MRI significantly reduces the interval-cancer rate in women with extremely dense breasts. The interval-cancer rate decreased from 5.0 per 1000 screenings in the mammography-only arm to 2.5 per 1000 screenings in the intervention arm (invitation for MRI). However, only 59% of the women for supplemental MRI screening actually participated and underwent screening MRI (chapter 2 and 3).

A complier average causal effect (CACE) analysis[1] was used to take into account the diluting effect of nonparticipation in MRI screening, on the estimate of interval cancer reduction in the intention-to-treat analysis. Using CACE analysis, we estimated that supplemental MRI screening among women who would have accepted screening if it had been offered, was associated with an interval-cancer rate that decreased to 0.8 per 1000 screenings, compared to 5.0 per 1000 screenings for mammography-only participants who would have accepted MRI screening (chapter 2).

The increased cancer detection occurred at a cost of more false positive results (chapter 2). False positive rates in the first MRI screening round were 79.8 per 1000 screenings, compared to 23.8 false positive results per 1000 screenings in women with extremely dense breasts participating in the mammography screening program.[2]

In the first, or *prevalence* screening round of supplemental screening, cancers that were already present but destined to be diagnosed later, are now detected earlier. As MRI is more sensitive than mammography it 'dives deeper' into the preclinical pool of cancers that mammography does not detect.[3] This causes a prevalence peak. In subsequent (incidence) screening rounds, detection rates are expected to decrease gradually towards the normal background incidence, since the pool of prevalent cancers will be depleted, and only new cancers come in at their regular incidence rate, although in an earlier stage than with mammography alone. This was confirmed in the second MRI screening round, which is the first *incidence* round, where supplemental breast cancer detection decreased to 5.8 per 1000 screenings (chapter 5). However, cancer detection in the second MRI screening round, after another negative mammography, is still substantial. Although populations are not entirely comparable, several studies on supplemental screening with ultrasound showed an additional detection rate of up to 4 per 1000 screenings, but then in the *prevalence* screening.[4,5]

In the second MRI screening round, the false positive rate was much lower, with 26.3 per 1000 screenings compared to 79.8 per 1000 screenings in the first screening round (chapter 5).

Effects of supplemental MRI screening on longer term, such as a reduction in the rate of advanced cancers, reduction in interval-cancer rates in subsequent screening rounds and modeling of cost-effectiveness and effects on breast cancer mortality are still under investigation.

Overdiagnosis

The use of supplemental MRI screening increases cancer detection. It is inevitable that among the cancers detected, there will also be cancers that would never have presented without MRI screening. Supplemental screening will thus also lead to overdiagnosis: the diagnosis and treatment of breast cancers that would have remained asymptomatic without screening, and would not have affected a woman's life expectancy.[6]

The proportion of breast cancers that is overdiagnosed is hard to estimate.[3] Randomized controlled trials, observational studies and modelling studies have been used to estimate overdiagnosis in mammographic screening, but all methodologies have their own challenges and disadvantages that may lead to biased estimates of overdiagnosis.[7–9]

The estimates of overdiagnosis in mammography screening range from 1% to 52% of the breast cancers detected by screening.[8,10–13] In the Dutch mammographic screening program, according to a modelling study, the extent of overdiagnosis is estimated to be around 8% of the breast cancers detected by screening.[14]

To ensure that the outcome of the DENSE trial could not be attributed to overdiagnosis only, the primary outcome was a reduction in interval cancers, as these are cancers presenting because of symptoms, and therefore are not overdiagnosed.

To be able to prove that interval-cancer rates decrease due to supplemental screening, a comparison with a population without supplemental screening is needed. This comparison group has to be similar to the group with supplemental screening in 'background' incidence of breast cancer. Randomized controlled trials ensure this comparability, because confounding by indication is no issue. This is, of course, provided that randomization is successful.[3]

Even now that we have observed that supplemental screening with MRI does indeed reduce interval-cancer rate, it remains important to estimate the overdiagnosis that is potentially associated with it.

A randomized controlled trial design, such as used for the DENSE trial, is the best design to estimate the excess cumulative incidence of breast cancer over time, because of the comparability in 'background' incidence of the intervention and control populations.[7–9]

It is important to realize that the excess cumulative incidence gives a good estimate of the number of overdiagnosed cancers, but only after a long follow-up, namely in the first year after screening is stopped, *plus* the maximum preclinical period (i.e., the period of time before which a cancer will be diagnosed clinically). A follow-up period that is too short will lead to an overestimation of the degree of overdiagnosis.[3]

Tailored screening

The DENSE trial is a first step towards tailoring the existing population-based screening program, based on a breast cancer risk factor, mammographic density, that is easily determined on mammography. Further tailoring of screening might optimize the harm-benefit balance of screening. In this line of reasoning, the benefits may be larger and perhaps the harms more acceptable or less pronounced in women who do not only have extremely dense breasts, but are also at increased breast cancer risk because of other risk factors. Ideally, one would evaluate the reduction in the interval-cancer rate and/or reduction in advanced-cancer rate through supplemental MRI screening within strata of women at different risks of breast cancer. Because of the prerandomization design of the DENSE trial, however, information on breast cancer risk factors, besides density and age, is not available for the control arm, since only the women in the MRI invitation arm were informed about the study. In the current Dutch screening program, information on breast cancer risk factors is not collected, and breast cancer risk estimates could therefore not be calculated for the control arm.

Instead, we were able to evaluate *early* screening performance measures within strata of women at different risks, *within* the MRI arm. Results of this study showed that supplemental MRI cancer detection rates were significantly higher in women in the highest quartile of 5-year breast cancer risk estimates (chapter 6).

Although we could not compare interval-cancer rates between the MRI and control arm for women in different risk strata, it seems reasonable to assume that in the highest quartile of 5-year breast cancer risk estimates (cut-off point 2.5%), the interval-cancer rate with mammography is also highest. We base this assumption on previous research that has shown that women with extremely dense breasts and 5-year breast cancer risk estimates of at least 1.67% are at highest risk of developing interval cancers in annual mammography screening.[15]

After the start of the DENSE trial, more randomized screening trials were set-up to study tailored screening strategies based on individual breast cancer risks, in order to optimize existing screening programs.[16,17]

The WISDOM (Women informed to Screen Depending on Measures of Risk) study is a multicenter study, conducted in the United States, that compares risk-based screening to annual mammography screening in women aged 40 to 74 years.[16] The study started in 2016 and aims to randomize 65,000 women who are assigned to annual screening or risk-based screening. In the risk-based screening arm, until age 50, screening intervals of 1 year, 2 years, or no screening are used, and frequency of screening is based on mammographic density, 5-year breast cancer risk estimates and genetic testing. After the age of 50, women who do not meet criteria for annual mammogram and MRI (e.g. 5-year risk of 6% or higher, potential mutation carrier) are offered annual or biennial mammography screening, based on breast cancer risk estimates: women with elevated risk (e.g. 5-year risk of developing ER-breast cancer at or above 0.75% or women in top of 2.5th percentile of risk by 1-year age category) are offered annual mammography screening and women aged 50 to 74 years without elevated risk are offered biennial mammogram instead of annual screening,[18]which is currently standard in the US.[19]

Another study aimed at tailored screening is the European Commission funded MyPeBS (My Personal Breast Cancer Screening) study, which is conducted in 5 countries (Belgium, France, Israel, Italy and the United Kingdom). The study aims to compare standard mammography screening to mammography screening with risk-based intervals and supplemental MRI. The trial aims to randomize 85,000 women in 2.5 years and has a follow-up of 4-years. Women aged 40 to 70 years without prior history of breast cancer or known very high breast cancer risk are eligible for randomization.[17]

All tailored screening options in this study are compared to the existing national screening programs, and tailoring is based on breast density, genetic profile, personal and family history of breast cancer and age. Women with dense breasts (category c and d) will receive mammography and/or tomosynthesis with or without supplemental ultrasound or MRI, and they will be randomized to regular screening frequency (depending on the national/regional guidelines) or to a screening frequency varying from 1 to 4 years, depending on breast cancer risk. In the risk-based screening arm, women with very high breast cancer risk will be offered annual screening with mammography and, until age 60, annual MRI.[17,20]

The Tomosynthesis Mammographic Imaging Screening Trial (TMIST) is a large randomized controlled trial that studies risk-based screening with tomosynthesis compared to risk-based screening with mammography. The trial started in 2017 and

aims to include approximately 165,000 women. In the trial, screening frequency will be based on several breast cancer risk factors, including age, breast density, family history of breast cancer and genetic risk. Women with dense breasts (category c and d) who are aged 45 to 69 years will be screened annually with mammography or tomosynthesis and women aged 70 to 74 will be screened biennially, with either mammography or tomosynthesis.[21]

All above mentioned studies are randomized trials with large study populations that consist of several consecutive screening rounds. The WISDOM and MyPeBS study both have a control arm that receives 'regular' screening. In the TMIST study the screening frequency is based on breast cancer risk, in both the mammography and the tomosynthesis arm.[21]

TMIST intensifies screening in women with extremely dense breasts, compared to the Dutch screening program, by changing the frequency to annual screening. It is arguable that this type of tailoring will improve screening in dense breasts, since it does not influence the masking effect on mammography or tomosynthesis that is caused by breast density. Moreover, previous research has shown that also in annual programs, women with dense breasts and otherwise increased breast cancer risk have an increased risk of interval cancers.[15]

As for the value of WISDOM and MyPeBs for density- and risk-based screening in the Netherlands, it is important to realize that these studies adhere to current local guidelines in the participating countries (commonly younger starting age, higher frequency and ultrasound for dense breasts), and as a result the tailored screening that is offered to low risk women is actually similar to the current Dutch screening program for all women. Minimizing or descaling current screening in low risk women is not part of these studies.[22]

Tailored screening with the use of breast cancer risk models is dependent on the quality of the models. The Breast Cancer Surveillance Consortium (BCSC) model[23] and Tyrer-Cuzick model,[24] which are the most frequently used models, have both been externally validated,[25–28] but also these models do not predict perfectly. The discriminatory accuracy in external validations is similar for both models, with areas under the curve of 0.66 and 0.64 for the BCSC and Tyrer-Cuzick model respectively.[26,27] Improvement is likely to be achieved by the addition of single nucleotide polymorphism information to the model[29] and by including other mammography-based information than breast density, such as texture features[30] or microcalcifications.[31] As a result of the imperfect discriminatory accuracy, there could be over- or underestimation of the actual breast cancer risks and in tailored screening, subsequently, some women could receive too much screening and others, whose risks are underestimated, possibly do not receive enough screening.

Future perspectives for tailored screening studies

All trials on supplemental or tailored screening have in common that they need to be large, consist of multiple screening rounds, require a long follow-up time and that they are very costly. To accomplish such trials is a real challenge.

It took almost 10 years from the first randomization in the DENSE trial until the publication of the results of the first screening round. This was despite the choice for an intermediate outcome, interval cancer reduction, instead of breast cancer related mortality reduction. The previously mentioned WISDOM, MyPeBs and TMIST trials too, have all chosen to study intermediate outcomes, and based the required sample size on these outcomes. The trials have not yet completed recruitment and estimated primary completion dates of these trials are between 2022 and 2025.[17,18,32] The accrual of these trials is challenging: WISDOM has included approximately half of the women since 2016, and TMIST has included 15% since 2017.[33,34] The even larger populations and longer follow-up time that are needed in studying mortality as outcome, seem to make it infeasible to conduct such trials. Firstly, screening modalities used in trials are likely to be surpassed by newer techniques during follow-up. Secondly, if women have the impression that there are potentially better screening techniques available, justified or not, they may no longer be willing to wait another 20 years for the trial results, and a rise in opportunistic supplemental screening will be the result. Besides that, the costs for such large, long-running trials will be so high that they are difficult, if not impossible to run with research grant funding.

The effects on (breast cancer) mortality will then have to be estimated using the intermediate outcome data from the randomized trials. This can be done through the use of so-called microsimulation models, for example the Microsimulation Screening Analysis (MISCAN) program.[35] Collaborations like the Cancer Intervention and Surveillance Modeling Network (CISNET),[36] that conduct joint analysis of multiple independent models, can be used to strengthen evidence that is provided by solitary models.[37]

Many supplemental screening studies use the so-called paired design, but, as argued earlier, the randomized control group is essential for validly estimating the intermediate outcomes that are the input for the microsimulation models. Can an improvement in efficiency be obtained then, by using alternative randomized trial designs, such as the prerandomized, or single-consent Zelen design, which we used, or 'trials-within-cohorts' (TwiCs) design?[38–40]

In these designs, only the trial participants randomized to the intervention are asked consent for the trial. This is likely to result in faster inclusions in the intervention arm. A disadvantage, especially of the single-consent Zelen design, is that data collection from the control arm is limited, since these participants are not actively recruited.

If regular screening data could be enriched with the routine collection of risk factor data, this could provide a possibility for evaluating risk- and density-based screening strategies in a randomized fashion using these designs.

For the sample size needed, these designs seem to have no benefit however, and for any design it is important to include incidence screening rounds and sufficient follow-up to be able to study reductions in interval-cancer rates and advanced-cancer rates.

Therefore, the gain of using these designs compared to traditional randomized trials probably is limited to more efficient inclusion, and less contamination. The concerns about overdiagnosis and thus the need for screening trials studying mortality as the end point would largely be reduced if we would be able to determinate, preferably already on the images, which are the most clinically relevant cancers to detect and treat.

An important step in this direction is being taken in the Dutch IMAGINE study that aims to develop a computer algorithm based on image features, that can improve mammography screening by distinguishing between aggressive and more indolent types of breast cancer.[41] Similar algorithms could of course also be developed for supplemental screening with MRI.

In the short term, it can be expected that artificial intelligence can already be used to predict whether lesions on breast MRI are benign or malignant.[42] Although this does not address the topic of overdiagnosis, it will help in reducing false positive rates of screening.

The first results of the DENSE trial showed a significant reduction of interval cancers in women with extremely dense breasts who were offered supplemental MRI screening. These results will be used in microsimulation models to estimate the effect on mortality and the harm-benefit ratios on the longer term. It will probably be impossible to conduct, and pay for, a trial that is large enough and running long enough to empirically investigate a (breast) cancer mortality reduction. Therefore, the evidence provided by trials such as the DENSE trial is likely to be the best available for policy makers to take decisions on the hotly debated issues of dense breasts and tailored screening.

The first results of the trial described in this thesis were reason for the Secretary of State of the Ministry of Health, Welfare and Sport to request official advice from the Dutch Health Council on the desirability of supplemental MRI screening for women with extremely dense breasts.[43] At the same time the National Institute for Public Health and the Environment was asked to investigate what would be the consequences in terms of organization, logistics and costs.[43] A formal response of the Secretary of State is expected while this thesis is being printed.

REFERENCES

1. Cuzick J, Edwards R, Segnan N. Adjusting for Non-Compliance and Contamination in Randomized Clinical Trials. *Statistics in medicine*. 1997;16(9):1017-1029.
2. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95-103.
3. Gulati R, Feuer EJ, Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies. *American Journal of Epidemiology*. 2016;184:140-147.
4. 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: The Journal of the American Medical Association*. 2012;307(13):1394-1404.
5. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): A randomised controlled trial. *The Lancet*. 2016;387(10016):341-348.
6. Brodersen J, Schwartz LM, Heneghan C, et al. Overdiagnosis: what it is and what it isn't. *BMJ Evidence-Based Medicine*. 2018;23:1-3.
7. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *The Lancet Oncology*. 2007;8:1129-1138.
8. Ripping TM, ten Haaf K, Verbeek ALM, van Ravesteyn NT, Broeders MJM. Quantifying overdiagnosis in cancer screening: A systematic review to evaluate the methodology.
9. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: A systematic review of methods. *BMJ*. 2015;350.
10. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. *The New England Journal of Medicine*. 2016;375(15):1438-1447.
11. de Gelder R, Heijnsdijk EAM, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiologic Reviews*. 2011;33(1):111-121.
12. Marmot M, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: An independent review. *The Lancet*. 2012;380(9855):1778-1786.
13. Autier P, Boniol M, Koechlin A, Pizot C, Boniol M. Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study. *BMJ*. 2017;359:j5224.
14. de Gelder R, Fracheboud J, Heijnsdijk EAM, et al. Digital mammography screening: Weighing reduced mortality against increased overdiagnosis. *Preventive Medicine*. 2011;53(3):134-140.
15. Kerlikowske K, Zhu W, Tosteson ANA, et al. Identifying women with dense breasts at high risk for interval cancer a cohort study. *Annals of Internal Medicine*. 2015;162(10):673-681.

16. Esserman LJ. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *npj Breast Cancer*. 2017;3(1):34.
17. ClinicalTrials.gov. My Personalized Breast Screening. Published May 5, 2020. Accessed June 11, 2020. https://clinicaltrials.gov/ct2/show/NCT03672331
18. ClinicalTrials.gov. Women Informed to Screen Depending on Measures of Risk (Wisdom Study). Published March 18, 2018. Accessed June 13, 2020. https://clinicaltrials.gov/ct2/show/NCT02620852
19. Siu AL. Screening for breast cancer: U.S. Preventive services task force recommendation statement. *Annals of Internal Medicine*. 2016;164(4):279-296.
20. The project - MyPeBS. Accessed June 15, 2020. https://mypebs.eu/the-project/
21. National Cancer Institute. TMIST (Tomosynthesis Mammographic Imaging Screening Trial). Published October 3, 2019. Accessed June 11, 2020. https://www.cancer.gov/about-cancer/treatment/clinical-trials/nci-supported/tmist
22. Torjesen I. Health groups criticise EU funded breast cancer screening study as a "missed opportunity." *BMJ*. 2020;368. Accessed June 12, 2020. http://group.bmj.com/group/rights-licensing/
23. 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. *Journal of Clinical Oncology*. 2015;33(28):3137-3143.
24. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in Medicine*. 2004;23(7):1111-1130.
25. Tice JA, Bissell MCS, Miglioretti DL, et al. Validation of the breast cancer surveillance consortium model of breast cancer risk. *Breast Cancer Research and Treatment*. 2019;175(2):519-523.
26. Vachon CM, Pankratz VS, Scott CG, et al. The Contributions of Breast Density and Common Genetic Variation to Breast Cancer Risk. *Journal of the National Cancer Institute*. 2015;107(5):397.
27. McCarthy AM, Guan Z, Welch M, et al. Performance of breast cancer risk-assessment models in a large mammography cohort. *Journal of the National Cancer Institute*. 2020;112:489-497.
28. Brentnall AR, Harkness EF, Astley SM, et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Research*. 2015;17(1):147.
29. Brentnall AR, Cuzick J. Risk Models for Breast Cancer and Their Validation. *Statistical Science*. 2020;35(1):14-30.
30. Wanders JOP, van Gils CH, Karssemeijer N, et al. The combined effect of mammographic texture and density on breast cancer risk: A cohort study. *Breast Cancer Research*. 2018;20(1).
31. Eriksson M, Czene K, Pawitan Y, Leifland K, Darabi H, Hall P. A clinical model for identifying the short-term risk of breast cancer. *Breast Cancer Research*. 2017;19(1).
32. ClinicalTrials.gov. Digital Tomosynthesis Mammography and Digital Mammography in Screening Patients for Breast Cancer. Published May 21, 2020. Accessed June 13, 2020. https://clinicaltrials.gov/ct2/show/NCT03233191

33. TMIST. First Quarter 2020 TMIST Study Updates. Published April 2020. Accessed June 15, 2020. https://pages.acr.org/index.php/email/emailWebview
34. WISDOM. The WISDOM study. Accessed June 15, 2020. https://www.thewisdomstudy.org/
35. Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Computer Methods and Programs in Biomedicine*. 1985;20(1):79-93.
36. Cancer Intervention and Surveillance Modeling Network (CISNET). Accessed June 12, 2020. https://cisnet.cancer.gov/
37. Berry DA, Cronin KA, Plevritis SK, et al. Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer. *The New England Journal of Medicine*. 2005;353(17):1784-1792.
38. Zelen M. A new design for randomized clinical trials. *The New England Journal of Medicine*. 1979;300(22):1242-1245.
39. Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: Introducing the "cohort multiple randomised controlled trial" design. *BMJ*. 2010;340(7753):c1066-undefined.
40. van der Velden JM, Verkooijen HM, Young-Afat DA, et al. The cohort multiple randomized controlled trial design: a valid and efficient alternative to pragmatic trials? *International Journal of Epidemiology*. Published online 2017:96-102.
41. KWF Kankerbestrijding. Wat zegt een mammogram over de agressiviteit van borstkanker? Published May 1, 2019. Accessed June 11, 2020. https://www.kwf.nl/onderzoek/onderzoeksdatabase/wat-zegt-een-mammogram-over-de-agressiviteit-van-borstkanker
42. 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. *Investigative Radiology*. 2020;55(7):438-444.
43. Blokhuis P. Adviesaanvraag Gezondheidsraad MRI in bevolkingsonderzoek borstkanker (Kamerbrief). Published November 27, 2019. Accessed June 13, 2020. https://www.rijksoverheid.nl/documenten/kamerstukken/2019/11/27/kamerbrief-over-adviesaanvraag-gezondheidsraad-mri-in-bevolkingsonderzoek-borstkanker

8

Summary

In 2011, the Dense tissue and Early breast Neoplasm ScrEening (DENSE) trial was initiated, the first randomized controlled trial to study the additional value of supplemental MRI screening in early detection of breast cancer in women with extremely dense breasts. The DENSE trial is a multicenter randomized controlled trial embedded in the Dutch population-based mammography screening program. The trial consists of three consecutive screening rounds, and the primary outcome of the trial is the reduction of interval cancers in the intervention arm compared to the control arm

Women with extremely dense breasts and a negative ('normal') result at mammography were randomized in a 1:4 ratio to either an invitation for supplemental MRI (intervention arm) or mammography screening only (control arm). A prerandomization single-consent design was used, in which randomization was performed before informed consent was obtained. Only women randomized to the intervention arm were asked to participate. This design was used to prevent anxiety in the control group and to reduce the probability that women in the control group would arrange for MRI examination on their own initiative. The primary outcome of the trial was the between-group difference in the incidence of interval cancers during a 2-year screening period.

This thesis reports the first results of the DENSE trial.

In **chapter 2** we studied the primary outcome of the first screening round of the DENSE trial: the effect of supplemental MRI on the incidence of interval cancers in women with extremely dense breast tissue in the first 2-year screening round of the trial. Data regarding the number of interval cancers and the tumor characteristics in the two groups was collected through linkage with the Netherlands Cancer Registry. Interval cancers included all breast cancers that were diagnosed after negative results on mammography or supplemental MRI, before the next scheduled mammography examination.

Between 2011 and 2015, in total 8,061 women were assigned to an invitation for supplemental MRI (intervention arm) and 32,312 were assigned to regular mammography screening only (control arm). Of the 8,061 women who were invited to undergo supplemental MRI, 4,783 (59.3%) actually underwent MRI screening.

In the MRI-invitation group, an interval cancer was diagnosed in 20 women (4 among the MRI screening participants and 16 among the nonparticipants who were invited but did not undergo screening) of 8,061. In the mammography-only group, an interval cancer was diagnosed in 161 of 32,312 women, which resulted in an interval-cancer rate of 2.5 per 1000 screenings (95% confidence interval [CI], 1.6 to 3.8) in the MRI-invitation group and 5.0 per 1000 screenings (95% CI, 4.3 to 5.8) in the mammography-only group. In the intention-to-screen analysis, the interval-cancer rate was lower by

2.5 per 1000 screenings (95% CI, 1.0 to 3.7) in the MRI-invitation group than in the mammography-only group (p<0.001).

Complier average causal effect (CACE) analysis was applied to estimate the effect of actually undergoing supplemental MRI screening in the subpopulation of women who said that they would have accepted MRI screening if it had been offered. Using CACE analysis, we estimated that supplemental MRI screening among women who would have accepted screening if it had been offered, was associated with an interval-cancer rate that decreased to 0.8 per 1000 screenings, compared to 5.0 per 1000 screenings for mammography-only participants who would have accepted MRI screening (p<0.001).

Among the 4,783 MRI participants, the recall rate was 94.9 per 1000 screenings (95% CI, 86.9 to 103.6), and the cancer-detection rate with MRI was 16.5 per 1000 screenings (95% CI, 13.3 to 20.5). The false positive rate was 79.8 per 1000 screenings (95% CI, 72.4 to 87.9).

In **chapter 3** we determined the willingness of women with extremely dense breasts to undergo breast screening with supplemental MRI in a research setting (the DENSE trial), and we examined reasons for women to participate or not.

Of the 8,061 women who were invited to undergo MRI, 5,267(65.5%) were interested in MRI screening, and 4,783 (59.3%) actually underwent MRI screening. The invitation was declined by 1,872 (23.3%) and 913 (11.3%) women were nonrespondents.

Women with higher socio-economic status (SES) were more often interested in participation than women with lower SES (68.3% versus 58.7%, p<0.001). Participants were on average 54-years old (interquartile range 51 to 59 years), comparable to women with extremely dense breasts in the population-based screening program (55 years).

All invitees were asked to state their reason(s) for participation or nonparticipation on the online registration form or on the reply card that was sent with the invitation.

The most frequently stated reasons for non-participation were “MRI-related inconveniences (e.g., claustrophobia, refusing contrast agent) and/or self-reported contraindications to MRI” (26.9%) and “anxiety regarding the result of supplemental screening” (20.8%). Women interested in participation stated “expected personal health benefit” (68.4%) and “contribution to science” (42.6%) most frequently as reason for participation.

Incidental findings are a disadvantage of nearly all cross-sectional imaging techniques, and might lead to (unnecessary) anxiety and treatment. The field of view of breast MRI

does not only consist of breast tissue, surrounding tissues and organs are also visible on breast MRI. In **chapter 4** we studied the prevalence, type and clinical relevance of extramammary incidental findings on the first screening round MRI examinations of the DENSE trial.

Radiologists reported incidental findings in the MRI reports, and in a free-text space in the online trial database. In case of incidental findings for which further evaluation was advised by the radiologist, the woman's primary care physician was informed by the radiologist. Incidental findings were considered clinically relevant in case the radiologist advised further evaluation or follow-up.

In the first screening round, in total 63 extramammary incidental findings were reported for 61 women (1.3% of 4,783 women). Incidental findings detected were most frequently pulmonary (25.4%), followed by hepatic (22.2%), mediastinal (20.6%) and musculoskeletal findings (20.6%). Further evaluation was recommended for 35 women with an incidental finding (57.4% of incidental findings;0.7% of 4,783 first screening round participants). Extramammary malignancy (hematologic, pulmonary or thymic) was histologically confirmed in 7 women (11.1% of incidental findings). These results show that the frequency of incidental findings that required further evaluation was low, in this large group of MRI-screened women from the general population.

Two years after the first (prevalent) screening round of the DENSE trial, women were invited for a subsequent mammography screening round. In case of a negative mammography result, these women were again invited to participate in supplemental MRI screening. This second screening round shows cancer detection rates after the prevalence peak of the first screening round, and also the influence on the false-positive rate of having previous MRI examinations available. In **chapter 5** we described the first results of the second MRI screening round (first incidence round) of the DENSE trial. We determined the recall rate, biopsy rate, cancer detection rate, false positive rate and positive predictive values of second round MRI screenings, as well as tumor characteristics of MRI-detected cancers.

Of 4,783 women in the MRI arm, 3,436 (71.8%) underwent a second MRI; 110 (3.2%) were recalled of whom 84 (2.5%) underwent a biopsy. The cancer detection rate was 5.8 per 1000 screenings (95% CI, 3.8 to 9.0) and the false positive rate was 26.3 per 1000 (95% CI, 21.5 to 32.3). Positive predictive value for recall was 18.2% (95% CI, 12.1 to 26.4) and positive predictive value for biopsy 23.8% (95% CI, 16.0 to 33.9). No node positive or late stage cancers were observed in the incidence round.

Not all women with extremely dense breasts have the same breast cancer risk. Supplemental screening could be further personalized by using individual breast cancer risks, based on risk factors besides breast density. In **chapter 6** we investigated the

effect of stratification by breast cancer risk on screening performance of supplemental MRI in the first and second screening round of the DENSE trial, in order to assess whether the harm benefits ratio of supplemental MRI screening could be improved. We stratified the participants by 5-year Tyrer-Cuzick and 5-year Breast Cancer Surveillance Consortium (BCSC) breast cancer risk. We also explored stratification based on age and percent volumetric density. MRI screening performance measures (rates for recall, MRI cancer-detection, false positive rate and positive predictive value of MRI were calculated for women in different risk quartiles.

In the first and second screening round, MRI cancer-detection rates increased for increasing quartiles of Tyrer-Cuzick 5-year risk; all with statistically significant linear trends (p-trend <0.001 and 0.033 respectively). False positive rates increased for increasing quartiles of Tyrer-Cuzick risks, but the trend over quartiles was not statistically significant. Positive predictive value of MRI increased for increasing quartiles of Tyrer-Cuzick risks, however the trend was only statistically significant in the first screening round and combined screening rounds (p-trend 0.002 and <0.001 respectively). MRI cancer-detection rates increased for increasing quartiles of BCSC risks, with significant linear trends for the second screening round and combined screening rounds (p-trend 0.044 and 0.014 respectively). In both screening rounds and combined screening rounds, false positive rates decreased for increasing quartiles of BCSC risks, all with a significant linear trend (p-trend 0.018, 0.011 and 0.002 respectively). Positive predictive value of MRI increased for increasing quartiles of BCSC risk, all with a significant linear trend (p-trend 0.017, 0.007 and <0.001 for the first, second and combined screening rounds respectively).

Interval cancer rates are as yet only available after the first, but not the second screening round. With only 4 interval cancers in total after supplemental MRI, this number was too low for meaningful statistical comparison.

Risk stratifications based on age and percent density within women with extremely dense breasts, did not distinguish subgroups with a more optimal harm-benefit balance as well as Tyrer-Cuzick and BCSC risks did.

In **chapter 7** we discuss the benefits and harms of supplemental screening and possibilities of a transition from the current one-size fits all mammography breast cancer screening program to tailored breast cancer screening.

In summary, the first results of the DENSE trial showed a significant reduction of interval cancers in women with extremely dense breasts who were offered supplemental MRI screening. These results will be used in microsimulation models to estimate the effect on mortality and the harm-benefit ratios on the longer term.

The first results of the trial described in this thesis were reason for the Secretary of State of the Ministry of Health, Welfare and Sport to request official advice from the Dutch Health Council on the desirability of supplemental MRI screening for women with extremely dense breasts. At the same time the National Institute for Public Health and the Environment was asked to investigate what would be the consequences in terms of organization, logistics and costs. A formal response of the Secretary of State is expected while this thesis is being printed.

A

Summary in Dutch

List of publications

Acknowledgments

Curriculum vitae

SAMENVATTING IN HET NEDERLANDS

In het bevolkingsonderzoek borstkanker in Nederland worden vrouwen van 50 tot 75 jaar iedere 2 jaar uitgenodigd voor mammografische screening. Dit zijn meer dan 1 miljoen vrouwen per jaar. Ongeveer 8% van deze vrouwen heeft zeer dicht borstweefsel. Vrouwen met zeer dicht borstweefsel hebben een 3-6 keer hoger risico op borstkanker dan vrouwen met heel weinig dicht borstweefsel. Bovendien is bij deze vrouwen de sensitiviteit van het mammogram lager, omdat borstkanker minder goed is op te sporen in dicht borstweefsel.

Aanvullend gebruik van een screeningstechniek die gevoeliger is, zoals bijvoorbeeld echografie of MRI, kan de effectiviteit van de borstkankerscreening voor deze groep vrouwen mogelijk verbeteren. Het is bekend dat aanvullende screening de opsporing van borstkanker verbetert, maar het is nog onduidelijk of gezondheidsuitkomsten hierdoor ook verbeteren.

De DENSE (Dense tissue and Early breast Neoplasm ScrEening) studie is de eerste gerandomiseerde studie naar het effect van aanvullende MRI-screening op het optreden van intervalkankers bij vrouwen met zeer dicht borstweefsel. De studiepopulatie bestaat uit deelneemsters van het bevolkingsonderzoek borstkanker met zeer dicht borstweefsel, en van wie het screeningsmammogram negatief ('normaal') was. Deze vrouwen werden in een verhouding van 1:4 gerandomiseerd naar een uitnodiging voor aanvullende MRI-screening (MRI-arm) of alleen mammografische screening (mammografie-arm).
Alleen de vrouwen die naar de MRI-arm waren gerandomiseerd werden geïnformeerd over de studie en gevraagd om deel te nemen. De vrouwen in de mammografie-arm ontvingen alleen de uitnodigingen van de reguliere screening.
De studie werd in 2011 goedgekeurd door de minister van Volksgezondheid, Welzijn en Sport (VWS), op advies van de Gezondheidsraad (2011/19 WBO, Den Haag, Nederland).

In dit proefschrift worden de eerste resultaten van de DENSE-studie gerapporteerd.

De primaire uitkomst van de eerste screeningsronde van de DENSE-studie is het effect van aanvullende MRI op het in het optreden van intervalkankers tussen de twee studiearmen. Deze resultaten worden besproken in **hoofdstuk 2**.
Informatie over intervalkankers werd verkregen door koppeling met de Nederlandse Kankerregistratie. Intervalkankers waren alle borstkankers die werden gediagnosticeerd na een negatieve ("normale") uitslag van het mammogram of de aanvullende MRI, en voordat het mammogram van de volgende ronde van het bevolkingsonderzoek werd gemaakt.

Tussen 2011 en 2015 werden in totaal 8061 vrouwen gerandomiseerd naar een uitnodiging voor aanvullende MRI (MRI-arm) en 32.312 vrouwen naar het reguliere bevolkingsonderzoek (mammografie-arm). Van de 8061 vrouwen die werden uitgenodigd voor een aanvullend MRI-onderzoek hebben 4783 (59,3%) vrouwen daadwerkelijk deelgenomen.

In de MRI-arm werden 20 vrouwen met een intervalkanker gediagnosticeerd (4 MRI-deelneemsters en 16 vrouwen die waren uitgenodigd maar niet hadden deelgenomen). In de mammografie-arm werden 161 vrouwen gediagnosticeerd met een intervalkanker. Per 1000 gescreende vrouwen zijn in de MRI-arm zijn 2,5 intervalkankers gediagnosticeerd (95%-betrouwbaarheidsinterval (BI): 1,6-3,8) en in de mammografie-arm 5,0 per 1000 gescreende vrouwen (95%-betrouwbaarheidsinterval: 4,3-5,8). Dit is een reductie in intervalkankers van 2,5 per 1000 gescreende vrouwen (95%-BI: 1,0-3,7) in de MRI-arm, in vergelijking met de mammografie-arm (p<0,001).
Omdat slechts een gedeelte van de uitgenodigde vrouwen deelnam aan de MRI-screening hebben we met behulp van een "complier-average causal effect"(CACE)-analyse het effect van aanvullende MRI-screening geschat in de subpopulatie vrouwen die, indien dit was aangeboden, de uitnodiging voor MRI-screening hadden geaccepteerd.
Uit de CACE-analyse blijkt dat het aantal intervalkankers kan afnemen naar 0,8 per 1000 screeningsonderzoeken in vergelijking met 5,0 per 1000 screeningsonderzoeken voor vrouwen uit de mammografie-arm die de uitnodiging voor MRI-screening hadden geaccepteerd (p<0,001).

Per 1000 met MRI-gescreende vrouwen werden er 94,9 (95%-BI: 86,9-20,5) verwezen voor aanvullend onderzoek; de aanvullende borstkankerdetectie was 16,5 per 1000 gescreende vrouwen (95%-BI: 13,3-20,5). Het aantal fout-positieve uitslagen was 78,9 per 1000 screeningsonderzoeken (95%-BI: 72,4-87,9).

A

In **hoofdstuk 3** bespreken we de bereidheid van vrouwen met zeer dicht borstweefsel om aanvullende MRI-screening te ondergaan in het kader van wetenschappelijk onderzoek (de DENSE-studie), en we onderzoeken redenen voor vrouwen om wel of niet deel te willen nemen.

Van de 8061 vrouwen die werden uitgenodigd voor aanvullende MRI, waren 5267 (65,5%) geïnteresseerd in deelname, en 4783 (59,3%) hebben daadwerkelijk deelgenomen aan de MRI-screening. De uitnodiging voor aanvullende MRI werd afgewezen door 1872 (23,3%) vrouwen en 913 (11,3%) vrouwen hebben niet gereageerd op de uitnodiging.

Vrouwen met een hogere sociaaleconomische status (SES) waren vaker geïnteresseerd in deelname dan vrouwen met een lagere SES (68,3% versus 58,7%, p<0,001). Deelneemsters waren gemiddeld 54 jaar oud (interkwartielafstand 51 tot 59

jaar), vergelijkbaar met de leeftijd van vrouwen met zeer dicht borstweefsel in het bevolkingsonderzoek (55 jaar).

Alle uitgenodigde vrouwen werden gevraagd om hun redenen voor deelname of voor niet-deelname op een online formulier of antwoordkaart te noteren.

De meest genoemde reden voor afwijzing van deelname waren gerelateerd aan de MRI (bijvoorbeeld claustrofobie, weigeren van contrastvloeistof) en/of zelf gerapporteerde contra-indicaties voor de MRI (26,9%) en angst voor de uitslag van de aanvullende screening (20,8%). Vrouwen met interesse in deelname noemden het meest frequent de verwachte persoonlijke gezondheidswinst (68,4%) en bijdrage aan de wetenschap (42,6%) als reden voor deelname.

Nevenbevindingen kunnen leiden tot (onnodige) angst en behandeling, en zijn een nadeel van bijna alle cross-sectionele beeldvormingstechnieken. Op een MRI van de borsten wordt naast borstweefsel ook aangrenzende weefsels en organen afgebeeld. Het voorkomen van nevenbevindingen in de eerste screeningronde van de DENSE-studie, de soort nevenbevindingen en de klinische relevantie ervan worden besproken in **hoofdstuk 4**.

Radiologen konden de nevenbevindingen rapporteren in het MRI-verslag en in een open tekstveld in de online database van de studie. Wanneer de radioloog aanvullend onderzoek adviseerde voor de nevenbevinding werd de huisarts van de deelneemster op de hoogte gebracht. Nevenbevindingen waarvoor aanvullend onderzoek of follow-up werd geadviseerd werden als klinisch relevant beschouwd.

In de eerste screeningronde werden in totaal 63 nevenbevindingen buiten de borsten gerapporteerd bij 61 vrouwen (1,3% van 4783 vrouwen). De nevenbevindingen bevonden zich meestal in de longen (25,4%), in de lever (22,2%), het mediastinum (20,6%) en musculoskeletaal (20,6%). Voor 35 vrouwen met een nevenbevinding werd aanvullend onderzoek geadviseerd (57,4% van de nevenbevindingen; 0,7% van 4783 eerste ronde deelneemsters). Bij 7 vrouwen werd een maligniteit buiten de borst (hematologisch, pulmonaal en in de thymus) vastgesteld (11,1% van de nevenbevindingen). De resultaten laten zien dat het voorkomen van nevenbevindingen in een grote groep met MRI-gescreende vrouwen uit de algehele populatie laag is.

Twee jaar na de eerste (prevalentie) screeningsronde van de DENSE-studie, werden vrouwen opnieuw uitgenodigd voor de volgende mammografische screeningsronde. De vrouwen met een negatieve ("normale") uitslag van deze mammografische screening werden uitgenodigd om opnieuw deel te nemen aan aanvullende MRI-screening.

Deze tweede screeningronde laat kankerdetectie zien na de prevalentiepiek van de eerste screeningronde. Daarnaast kan het effect van het hebben van een eerste ronde MRI op de foutpositieve uitslagen worden geëvalueerd. In **hoofdstuk 5** beschrijven we de eerste resultaten van de tweede screeningsronde (eerste incidentie ronde) van de DENSE-studie. We bepaalden de verwijscijfers, aantal biopten, kankerdetectie en foutpositieve uitslagen en de positief voorspellende waarde van de tweede screeningsronde MRI. Daarnaast werden tumorkenmerken van de met MRI-gedetecteerde kankers beschreven.

Van de 4783 deelneemsters in de MRI-arm hebben 3436 (71,8%) een tweede ronde MRI ondergaan; 110 (3,2%) vrouwen werden verwezen van wie 84 (2,5%) een biopt hebben ondergaan. De aanvullende borstkankerdetectie was 5,8 per 1000 gescreende vrouwen (95%-BI: 3,8-9,0) en het aantal fout-positieve uitslagen was 26,3 per 1000 screeningsonderzoeken (95%-BI: 21,5-32,3). De positief voorspellende waarde voor verwijzing was 18,2% (95%-BI: 12,1-26,4) en de positief voorspellende waarde voor biopt 23,8% (95%-BI: 16,0-33,9). In deze incidentie ronde werden geen borstkankers met positieve lymfeklieren of hoog stadium kankers gedetecteerd.

Niet alle vrouwen met zeer dicht borstweefsel hebben hetzelfde borstkankerrisico. Aanvullende screening zou beter gepersonaliseerd kunnen worden door het gebruik van individuele borstkankerrisico's, gebaseerd op andere risicofactoren naast de dichtheid van het borstweefsel. Om te onderzoeken of de verhouding tussen baten en lasten van aanvullende MRI-screening verbeterd kan worden, onderzochten we de effecten van stratificatie naar borstkankerrisico op de screeningsresultaten van de eerste en tweede ronde van de DENSE-studie. Deze resultaten worden gepresenteerd in **hoofdstuk 6**.

Deelneemsters werden gestratificeerd naar 5-jaar Tyrer-Cuzick en 5-jaar Breast Cancer Surveillance Consortium (BCSC) borstkankerrisico. Daarnaast onderzochten we stratificatie gebaseerd op leeftijd en percentage borstdensiteit. Het aantal verwijzingen, aantal biopten, MRI-gedetecteerde kankers en fout-positieve uitkomsten per 1000 gescreende vrouwen, de positief voorspellende waarde voor MRI werden berekend voor de verschillende risicokwartielen.
MRI-gedetecteerde kankers in de eerste en tweede screeningsronde waren het hoogste in de hoogste kwartielen het 5-jaar Tyrer-Cuzick borstkankerrisico; met statistisch significante lineaire trends (p-trend <0,001 en 0,033 respectievelijk). Fout-positieve uitkomsten namen eveneens toe bij hogere Tyrer-Cuzick risico's, maar de trend over de kwartielen was niet statistisch significant. Positief voorspellende waarde van MRI nam toe met over de kwartielen van Tyrer-Cuzick risico's, echter was de trend alleen statistisch significant in de eerste en de gecombineerde screeningsrondes (p-trend 0,002 en <0,001, respectievelijk).

MRI-gedetecteerde kankers namen toe met toenemende kwartielen van 5-jaar BCSC-borstkankerrisico, met een significante lineaire trend voor de tweede screeningsronde en de gecombineerde screeningsrondes (p-trend 0,044 en 0,014 respectievelijk).
De fout-positieve uitkomsten namen in beide screeningsronden en de gecombineerde screeningsronde af met toenemende kwartielen van BCSC-risico's, allen met een significante lineaire trend (p-trend 0,018, 0,011 en 0,002 respectievelijk).
De positief voorspellende waarde van MRI nam toe met toenemende kwartielen van BCSC-risico, allen met een significante lineaire trend (p-trend 0,017, 0,007 en <0,001 voor de eerste ronde, tweede ronde en de screeningsrondes gecombineerd).

Gegevens over het aantal intervalkankers zijn tot nu toe alleen beschikbaar van de eerste screeningronde, maar nog niet van de tweede screeningsronde. Het aantal intervalkankers van de eerste ronde, slechts 4 intervalkankers na aanvullende MRI-screening, is te laag voor statistische vergelijkingen.

Risicostratificaties gebaseerd op leeftijd en percentage borstdensiteit in vrouwen met zeer dicht borstweefsel, maakten geen beter onderscheid in subgroepen met een gunstigere baten-lasten verhouding dan de Tyrer-Cuzick en BCSC-risico's deden.

In **hoofdstuk 7** bediscussiëren we de baten en lasten van aanvullende screening en de mogelijkheden van een transitie van het huidige "one-size fits all" mammografische screeningsprogramma naar een geïndividualiseerd screeningsprogramma.

De eerste resultaten van de DENSE-studie laten een significante daling in intervalkankers zien in vrouwen met zeer dicht borstweefsel die aanvullende MRI-screening aangeboden hebben gekregen. Deze resultaten zullen gebruikt worden in microsimulatiemodellen om het effect op borstkankersterfte en lange termijn baten-lasten verhoudingen te schatten.

De resultaten van de DENSE-studie, beschreven in dit proefschrift, hebben geresulteerd in het verzoek van de staatssecretaris van het Ministerie van Volksgezondheid, Welzijn en Sport aan de Gezondheidsraad om officieel advies uit te brengen over de wenselijkheid van aanvullende MRI-screening voor vrouwen met zeer dicht borstweefsel. Tegelijkertijd heeft de staatssecretaris het Centrum voor Bevolkingsonderzoek van het Rijksinstituut voor Volksgezondheid en Milieu gevraagd om een uitvoeringstoets te verrichten, hierin wordt onderzocht wat nodig is voor de invoering van MRI-screening, en of dit haalbaar en betaalbaar is.
Terwijl dit proefschrift wordt gedrukt, wordt een formeel besluit van de staatssecretaris verwacht.

LIST OF PUBLICATIONS

Bakker MF*, de Lange SV*, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, Emaus MJ, Loo CE, Bisschops RHC, Lobbes MBI, de Jong MDF, Duvivier KM, Veltman J, Karssemeijer N, de Koning HJ, van Diest PJ, Mali WPThM, van den Bosch MAAJ, Veldhuis WB**, van Gils CH**, for the DENSE trial study group. Supplemental MRI screening for women with extremely dense breast tissue. The New England Journal of Medicine. 2019; 381:2091-2102.
*Shared first authorship, **Shared last authorship

Bakker MF, de Lange SV, van Gils CH. MRI screening in women with dense breasts. Reply. The New England Journal of Medicine. 2020; 382:1284

de Lange SV*, Bakker MF*, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, Emaus MJ, Loo CE, Bisschops RHC, Lobbes MBI, de Jong MDF, Duvivier KM, Veltman J, Karssemeijer N, de Koning HJ, van Diest PJ, Mali WPThM, van den Bosch MAAJ, Veldhuis WB**, van Gils CH**. Meerwaarde van MRI-screening bij zeer dicht borstweefsel in het bevolkingsonderzoek borstkanker. Nederlands Tijdschrift voor Geneeskunde. 2020; 164:D4673 (in Dutch)
**Shared first authorship, **Shared last authorship*

de Lange SV, Bakker MF, Monninkhof EM, Peeters PHM, de Koekkoek-Doll PK, Mann RM, Rutten MJCM , Bisschops RHC, Veltman J, Duvivier KM, Lobbes MBI, de Koning HJ, Karssemeijer N, Pijnappel RM, Veldhuis WB, van Gils CH. Reasons for (non)participation in the randomized DENSE trial on population- based MRI breast screening for women with extremely dense breasts. *Clinical Radiology. 2018; 73(8):759.e1-759.e9.*

Veenhuizen SGA, de Lange SV, Bakker MF, Pijnappel RM, Mann RM, Monninkhof EM, Emaus MJ, de Koekkoek-Doll PK, Bisschops RHC, Lobbes MBI, de Jong MDF, Duvivier KM, Veltman J, Karssemeijer N, de Koning HJ, van Diest PJ, Mali WPThM, van den Bosch MAAJ, Veldhuis WB*, van Gils CH*, for the DENSE trial study group. Supplemental MRI for women with extremely dense breasts – results of the second screening round of the DENSE trial. *Submitted.*
**Shared last authorship*

de Lange SV, Bakker MF, Pijnappel RM, Monninkhof EM, Emaus MJ, Mann RM, Loo CE, Bisschops RHC, Lobbes MBI, Rutten MJCM, Duvivier KM, Veltman J, Karssemeijer N, de Koning HJ, van Diest PJ, Mali WPThM, van den Bosch MAAJ, Veldhuis WB**, van Gils CH**, for the DENSE trial study group. Screening performance of supplemental MRI in women with extremely dense breasts stratified by breast cancer risk. *Submitted.*

de Lange SV, Bakker MF, Mann RM, Emaus MJ, de Koekkoek-Doll PK, Bisschops RHC, Lobbes MBI, de Jong MDF, Duvivier KM, Veltman J, Pijnappel RM, van Gils CH, Veldhuis WB, for the DENSE trial study group. Extramammary incidental findings in a population-based breast MRI screening trial for women with extremely dense breasts. *Submitted.*

DANKWOORD

Dit proefschrift was nooit tot stand gekomen zonder de bijdragen en hulp van velen. Een aantal mensen wil ik op deze plek in het bijzonder bedanken.

Allereerst gaat mijn dank uit naar alle deelneemsters van de DENSE-studie, die door deel te nemen aan de aanvullende MRI-screening hun medewerking hebben verleend aan dit onderzoek.

Geachte prof. dr. C.H. van Gils, beste Carla, wat vind ik het een eer dat ik je eerste promovenda mag zijn. Jij hebt niet alleen altijd interesse getoond in mijn werk, maar ook in mij als persoon. Naast dat jij mij heel veel hebt geleerd over epidemiologie, heb je mij een kans gegeven om nog meer van de wereld te zien en om vakgenoten vanuit de hele wereld te ontmoeten.

Geachte dr. M.F. Bakker, beste Marije, ik heb het erg gewaardeerd dat je altijd benaderbaar was en bereid was om mij te helpen als ik vragen had. Je georganiseerdheid en de wijze waarop je het onderzoek in goede banen leidt hebben veel indruk op mij gemaakt.

Geachte prof. dr. R.M. Pijnappel, beste Ruud, ondanks dat wij weinig persoonlijk overleg hebben gehad, was jij altijd in staat om mijn stukken, supersnel, van relevante feedback te voorzien.

Geachte dr. W.B. Veldhuis, beste Wouter, jij was altijd bereid en in staat om mij te helpen wanneer ik dit nodig had. Dank voor al je opmerkingen en waardevolle inzichten.

De leden van de beoordelingscommissie, prof. dr. S.C. Linn, prof. dr. H.M. Verkooijen, prof. dr. J. Wesseling, prof. dr. E. J. Th. Rutgers, en dr. L.E.M. Duijm wil ik hierbij graag bedanken voor hun bereidheid om mijn proefschrift te beoordelen.

Alle radiologen, laboranten, trialbureaumedewerkers en datamanagers van de DENSE-centra wil ik bedanken voor de fijne samenwerking en hun bijdragen aan de DENSE-studie. De regionale screeningsorganisaties, het LRCB, het RIVM en Volpara Health Technologies wil ik bedanken voor hun bijdragen aan de DENSE-studie.
De coauteurs wil ik graag bedanken voor de bijdragen aan de manuscripten in dit proefschrift.

Collega's van de oncologiegroep, dank voor de waardevolle tweewekelijkse overleggen.

Collega AIOS, fellows en stafleden van de radiologie in het UMC Utrecht, dank voor al jullie begrip en support in de afgelopen jaren.

Oud-kamergenoten van 5.122 en van Geunsgenoten, wat was het fijn om met jullie niet alleen inhoudelijke zaken en problemen te kunnen delen, maar ook de nodige koffiepauzes bij MíCaffè, de tostilunches in de Brink en de gezellige etentjes.

Lieve vrienden en vriendinnen, dank voor jullie interesse in mijn promotietraject en voor jullie steun.

Lieve paranimfen Nicoline en Renee. Nicoline, wat fijn dat wij in hetzelfde promotie-schuitje zaten en altijd alle successen en issues met elkaar kunnen delen. Daar waar ik beren op de weg zie, zie jij uitdagingen. Ik bewonder jouw veelzijdigheid en lef.
Renee, we begonnen als huisgenoten, maar zijn inmiddels alweer 10 jaar vriendinnen. Dank voor alle gezellige etentjes, een luisterend oor en de broodnodige relativering.

Lieve Rinus en Barry, wat prijs ik mij gelukkig met jullie als schoonouders.
Leonard & Eileen, Willeke & Awais, Christiaan & Laura, jullie hebben altijd interesse in mijn promotietraject getoond, dank daarvoor. Ik hoop dat er nog vele gezamenlijke etentjes (en wijntjes) mogen volgen!

Lieve Désirée en Rogier, dank voor het meedelen in al het lief en leed dat gepaard ging met het tot stand komen van dit proefschrift. Dees, bedankt voor het meelezen van het manuscript, zonder jou was dit dankwoord er niet geweest ;).

Lieve papa en mama, wanneer ik het moeilijk had, hielp jullie relativeringsvermogen en Noord-Nederlandse nuchterheid mij om niet op te geven. Ik wil jullie bedanken voor jullie onvoorwaardelijke steun en stimulans om altijd te gaan voor het beste resultaat.

Lieve Q, jij staat altijd voor mij klaar. Ik hou van jou.

CURRICULUM VITAE

Stéphanie werd geboren op 30 mei 1989 in Zuidlaren.

In 2007 behaalde zij haar vwo-diploma op het Maartens College in Haren, waarna zij in hetzelfde jaar startte met de studie geneeskunde aan de Universiteit Utrecht.

Na haar afstuderen in 2014 deed zij haar eerste werkervaring op als arts niet in opleiding op de afdeling klinische geriatrie in het Groene Hart Ziekenhuis in Gouda. In het voorjaar van 2015 startte zij met haar promotieonderzoek in het UMC Utrecht, onder supervisie van prof. dr. C.H. van Gils en dr. M.F. Bakker (Julius Centrum) en prof. dr. R.M. Pijnappel en dr. W.B. Veldhuis (afdeling Radiologie).

Tijdens haar promotieonderzoek heeft zij de postgraduate master epidemiologie behaald.

In september 2018 is zij begonnen aan de opleiding tot radioloog in het UMC Utrecht onder supervisie van dr. R.A.J. Nievelstein.