University of Groningen

MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc)

FaMRIsc Study Grp; Saadatmand, Sepideh; Geuzinge, H. Amarens; Rutgers, Emiel J. T.; Mann, Ritse M.; van Zuidewijn, Diderick B. W. de Roy; Zonderland, Harmien M.; Tollenaar, Rob A. E. M.; Lobbes, Marc B.; Ausems, Margreet G. E. M.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
FaMRIsc Study Grp, Saadatmand, S., Geuzinge, H. A., Rutgers, E. J. T., Mann, R. M., van Zuidewijn, D. B. W. D. R., Zonderland, H. M., Tollenaar, R. A. E. M., Lobbes, M. B., Ausems, M. G. E. M., van t'Riet, M., Hooning, M. J., Mares-Engelberts, I., Luiten, E. J. T., Heijnsdijk, E. A. M., Verhoef, C., Karssemeijer, N., Oosterwijk, J. C., Obdeijn, I-M., ... Tilanus-Linthorst, M. M. A. (2019). MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncology*, *20*(8), 1136-1147. https://doi.org/10.1016/S1470-2045(19)30275-X

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

Sepideh Saadatmand, MD[1*]; H Amarens Geuzinge[2*]; Professor Emiel JT Rutgers MD[3]; Ritse M Mann MD[4]; Diderick BW de Roy van Zuidewijn, MD[5], Harmien M Zonderland, MD[6]; Professor Rob AEM Tollenaar, MD[7]; Marc BI Lobbes, MD[8]; Margreet GEM Ausems, MD[9]; Martijne van 't Riet, MD[10]; Maartje J Hooning MD[11]; Ingeborg Mares-Engelberts, MD[12]; Ernest JT Luiten MD[13]; Eveline AM Heijnsdijk PhD[2]; Professor Cees Verhoef MD[1]; Nico Karssemeijer PhD[4]; Jan C Oosterwijk, MD[14]; Inge-Marie Obdeijn, MD[15]; Professor Harry J de Koning, MD[2]; Madeleine MA Tilanus-Linthorst, MD.[1] on behalf of the FaMRIsc study-group

* These authors contributed equally to this article.

1. Department of Surgery, Erasmus University Medical Centre, Rotterdam, Netherlands, sepideh1985@hotmail.com, c.verhoef@erasmusmc.nl, madeleinetilanus@hotmail.com

2. Department of Public Health, Erasmus University Medical Centre, Rotterdam, Netherlands, h.geuzinge@erasmusmc.nl, e.heijnsdijk@erasmusmc.nl, h.dekoning@erasmusmc.nl

3. Department of Surgery, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands, e.rutgers@nki.nl

4. Department of Radiology and Nuclear Medicine, Radboud University Hospital, Nijmegen, Netherlands, ritse.mann@radboudumc.nl; nico.karssemeijer@radboudumc.nl

5. Department of Surgery, Medical Centre Leeuwarden, Leeuwarden, Netherlands, d.b.w.de.roy.van.zuidewijn@znb.nl

6. Department of Radiology and Nuclear Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands, hmzonderland@outlook.com

7. Department of Surgery, Leiden University Medical Centre, Leiden, Netherlands, r.a.e.m.tollenaar@lumc.nl

8. Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, Maastricht, Netherlands, marc.lobbes@mumc.nl

9. Department of Genetics, University Medical Centre, Utrecht, Netherlands, m.g.e.m.ausems@umcutrecht.nl

10. Department of Surgery, Reinier de Graaf Gasthuis, Delft, Netherlands, rietm@rdgg.nl

11. Department of Medical Oncology, Erasmus University Medical Centre, Rotterdam, Netherlands, m.hooning@erasmusmc.nl

12. Department of Surgery, Vlietland ziekenhuis, Schiedam, Netherlands, i.mares@franciscus.nl

13. Department of Surgery, Amphia ziekenhuis, Breda, Netherlands, eluiten@amphia.nl

14. Department of Genetics, Groningen University, University Medical Centre Groningen, Groningen, Netherlands, j.c.oosterwijk@umcg.nl

15. Department of Radiology and Nuclear Medicine, Erasmus University Medical Centre, Rotterdam, Netherlands, a.obdeijn@erasmusmc.nl

Article type: Original article
Key words: breast cancer, MRI, mammography, screening, familial risk, breast density, randomised trial

Corresponding author:
Madeleine M. A. Tilanus-Linthorst
Department of Surgery
Erasmus Medical Center
P.O. Box 70780, 3070 WB Rotterdam
The Netherlands
Phone number: +31-641389580
Email: madeleinetilanus@hotmail.com

RESEARCH IN CONTEXT

Evidence before this study

We searched PubMed on November 30, 2018, with terms 'breast cancer screening', 'MRI screening', 'breast cancer', 'family history', 'familial risk' in several combinations, for prospective studies written in English with no restrictions in publication date. Several screening trials in women with a familial or genetic predisposition have been performed, all applying MRI and mammography at the same time, showing that by the addition of MRI to mammography screening, many breast cancers are detected in an earlier stage. A meta-analysis of these studies showed that MRI and mammography combined, resulted in a sensitivity of 98% while sensitivity of MRI and mammography alone were 89% and 55%, respectively. Specificity of MRI combined with mammography was 79%.

Unfortunately, all previous studies were performed in a non-randomised setting. As all participants were screened with MRI and mammography at the same time, it is unknown after what time a tumour, detected by MRI only, would have been detected by mammography, and whether this would cause a stage difference that is clinically relevant. With this limited evidence, screening guidelines for women with familial risk differ between countries.

Added value of this study

Our study is the first randomised controlled trial to our knowledge, to show the extent of the stage shift by adding MRI to mammography screening. We showed that the median size of invasive cancers detected under the MRI protocol was significantly (8 mm) smaller and cancers were far less often node positive than those detected under the mammography protocol. Importantly, in the incident rounds, without interval cancers with MRI, the absolute numbers of late-stage tumours (large or node positive) were also smaller. High density indicated poorer stage, and lower specificity for both MRI and mammography and was more informative than age.

The downside of screening of false-positive results occurs especially in the highest density category, also in the MRI-protocol. Some overdiagnosis (e.g. part of ductal carcinoma in situ, especially grade 1) occurs also with MRI-screening

Implications of all the available evidence

In addition to previously published evidence, our study indicates that MRI-screening in high risk women leads to significant and relevantly earlier detection of breast cancer, and fewer late-stage cancers, which may reduce the need for adjuvant chemotherapy and most likely reduce mortality. Density is relevant for the choice of a screenings strategy. Our findings can be used to inform policy discussions about the implementation of MRI in (high risk) screening.

ABSTRACT

Background: Screening guidelines for women at familial risk of breast cancer without a known causative gene mutation differ internationally. To our knowledge, no randomised controlled MRI-screening trial has been performed. The FaMRIsc-study aims to assess the efficacy of MRI versus mammography screening for familial risk and furthermore assesses the influence of breast density.

Methods: In 12 Dutch hospitals, 1355 women aged 30–55 years with a cumulative lifetime risk of ≥20% without a *BRCA1/2* mutation were randomised with a web-based computer generated hospital sequence, concealed for participants, physicians and researchers, in either the MRI-group with yearly MRI, clinical breast examination and mammography biennially, or the Mammography-group with yearly mammography and clinical breast examination. Breastfeeding, pregnancy, previous screening and previous ductal carcinoma in situ were permitted, but no previous invasive cancer. Primary outcomes were number, size and nodal-stage of breast cancers. Secondary outcomes were sensitivity, specificity and positive predictive value. Results were also stratified by mammographic density (BI-RADS A-D). Intention to screen analyses were performed. This trial was registered with the Netherlands Trial Register, NTR2789.

Findings: Between Jan 1 2011, and Dec 31 2017 in the MRI-group (674 women) compared to the Mammography-group (680 women) with a median follow-up of 5.2 years for both groups, more breast cancers were detected (40 versus 15, $p<0.001$), invasive cancers were smaller (median size 9 versus 17 mm, $p=0.010$) and less frequently node positive (4/24, 17% versus 5/8, 63%, $p=0.023$). In the MRI-group, specificity was significantly lower compared to the Mammography-group (83·8% versus 91·0%, $p<0.001$), while sensitivity hardly differed (97·5% versus 86.7%, $p=0.18$). Clinical breast examination contributed hardly to detection (1/55). In incident cancers, tumour stage was better in the MRI-group ($p=0.035$), with lower numbers of node positive and ≥T2 tumours, while specificity improved in both arms (MRI-group: 87·4%, Mammography-group: 92·6%, $p<0.001$). All tumours ≥T2 were in the two highest density categories. In BI-RADS density A-C MRI was most effective.

Interpretation: The earlier detection by MRI screening and especially the fewer late-stage cancers in incident rounds, may reduce adjuvant chemotherapy and mortality. However, especially for women with the highest breast density at the cost of more false positive results.

Funding: Dutch Government ZonMw, The Dutch Cancer Society, A Sisters Hope, Pink Ribbon, Stichting Coolsingel, J&T Rijke Stichting

INTRODUCTION

Approximately 15% of all female breast cancers occur in women with a family history of breast cancer, in whom no causative hereditary gene mutation has been found (familial risk).[1] These women are at greater risk for breast cancer, also at a relatively young age.[2]

Overall survival decreases considerably with increasing breast cancer size at detection and number of axillary lymph nodes involved, even with optimal adjuvant systemic therapy.[3,4] Screening aims to improve survival, by detecting breast cancer in an early stage. However, it also causes false positive results.

A decade ago several screening trials, comparing Magnetic Resonance Imaging (MRI) and mammography screening in high-risk women concluded that adding MRI to mammography screening improves early breast cancer detection in women with a familial or genetic predisposition.[5-7] As a result, guidelines for breast cancer screening were modified globally.[8-10]

Unfortunately, these studies were all in a non-randomised paired design in which MRI and mammography were performed simultaneously.[5-7,11] Therefore, it is unknown when an MRI-only detected tumour would have been detected by mammography, and whether a possible stage difference is clinically relevant. With this limited evidence, screening guidelines for women with familial risk differ between countries. American guidelines advise annual mammography, clinical breast examination, and MRI for women ≥30 years with a cumulative lifetime risk ≥ 20%.[8] Dutch and British guidelines omit MRI for women at familial risk without a BRCA1/2 mutation.[9,10]

Furthermore, breast density has not been considered in these studies.[5,6] Higher breast density, caused by more glandular and connective breast tissue in relation to fat, indicates a higher cancer risk also in women

with familial risk.[12] It impairs sensitivity of mammography,[12] but less of MRI,[13] and may cause more false positive results. Breast density is high in about 74% of women between 40 to 49 years, and in 45% of women in their 60s.[14] MRI might not be necessary for all women with familial breast cancer risk,[15] but breast density might be a parameter to identify subgroups for whom MRI screening could be useful.

The Familial MRI Screening study (FaMRIsc) was performed to address these issues. In this multicentre randomised controlled trial, women with a familial breast cancer risk were randomised between screening with 1) annual MRI and clinical breast examination plus biennial mammography; and 2) annual mammography and clinical breast examination.

METHODS

Study design and participants

The FaMRIsc study was a Dutch multicentre randomized controlled trial . The study follows the Helsinki declaration and was approved by the Institutional Review Board of the Erasmus University Medical Centre, Rotterdam, the Netherlands (reference-number: MEC-2010-292). The study protocol was published previously.[16]

Women aged 30–55 years with a cumulative lifetime risk ≥ 20% because of a familial predisposition according to the modified tables of Claus[5,17] or as assessed at a Clinical Genetics Centre were eligible. Exclusion criteria were previous invasive cancer, *BRCA1, BRCA2* or *TP53* mutation (proven or 50% risk of) since MRI screening is already advised for this group,[8-10] and a contraindication for contrast-enhanced MRI. Previous screening, a ductal carcinoma in situ diagnosis, pregnancy and breast feeding were permitted.

Participants were recruited from outpatient breast clinics or family cancer clinics at seven academic medical centres in the Netherlands and five of the larger hospitals (see Appendix). The physician of the outpatient clinic or family cancer clinic enrolled participants after written informed consent,

Randomisation

Randomization was performed via a web-based system and stratified per centre. Allocation was based on a general number between 1-100 that was randomly generated by the computer, half of them for either group. An algorithm decreased the possibility of the computer generating a number that led to allocation in an overrepresented study-group by a factor 5 minus 1. However, it remained impossible to predict what allocation would follow for the randomising physician, participant or researcher.

Procedures

The Mammography-group received annual mammography according to Dutch guidelines[9] plus clinical breast examination, which Dutch guidelines recommend in women with a lifetime risk of ≥30%. The MRI-group was screened with annual MRI and clinical breast examination, and mammography biennially. Leaving out mammography every other year was considered safe in the MRI-group and might prevent overdiagnosis of low-grade ductal carcinoma in situ .[18] Women who did not provide consent for randomisation, were asked consent for participation in the registration group arm (reg-group). Women in the reg-Mammography-group were screened according to the Mammography-group, usually because they did not want an extra investigation but screening according to our national guidelines ; and women in the reg-MRI-group according to the MRI-group, a choice of both the woman and her physician. A participant was removed from the study if she met one of the exclusion criteria (e.g. she developed invasive cancer, not being a screen-detected or interval breast cancer, or she appeared to be BRCA1/2 mutation carrier), or no longer met the inclusion criteria (e.g. lifetime risk fell below 20% because the family history was now explained by a BRCA1/2 mutation of which she was no carrier)

Mammographic examination was performed using full field digital mammography (FFDM). All mammography examinations were assessed according to the Breast Imaging Reporting and Data System (BI-RADS, 4th edition) and all MRI examinations were assessed to the Breast Imaging Reporting and Data System (BI-RADS, 1th

edition) of the American College of Radiology.[19] MRI and mammography were preferably scheduled on the same day.

A positive screening test was defined as a mammographic or MRI examination with a BI-RADS score of 3, for which additional investigation or a repeat examination at 6 months per radiologic judgement followed, or a score of 4-5, indicating histology; or a clinical breast examination with an abnormality, for which additional diagnostic testing was recommended. In the MRI-group, MRI and mammography were not independently read. To determine mammographic density, an automated breast density measurement (Volpara, version 1.3.0)[20] was done on raw data of the first FFDM of all participants, and estimated by radiologists at the mammograms according to the ACR BI-RADS breast composition categories: A = fatty; B = scattered fibroglandular; C = heterogeneously dense; D = extremely dense.[19] Dynamic contrast-enhanced breast MRI exams were performed according to the published protocol.[16] The study was planned to take 4 years.

Outcomes

Primary outcomes of this study were number, size and nodal status of detected breast cancers, both ductal carcinoma in situ and invasive. Secondary outcomes were false-positive results, sensitivity, specificity, and Positive predictive value (BI-RADS ≥3) and Positive predictive value of biopsies. Cost-effectiveness will be assessed in future analyses. Outcomes were also stratified by mammographic density (BI-RADS A-D).

Statistical analysis

The sample size needed was calculated based on the incidence rate of 7/1000 women years among women at familial risk screened in the Dutch MRI Screening Study (MRISC).[5] We expected, based on previous studies a sensitivity of 70% for MRI and 40% for mammography and after 4000 women year at risk in the study-groups, the detection of 32 tumours in the MRI-group and 18 in the Mx-group. With these 50 cancers, a difference in tumour size of 8 mm (SD tumour size: 9 mm) was expected to be statistically significant (two sided alpha=0·05) with a power of 80%. A difference of 8 mm was considered to be clinically relevant. With fewer women included, the number of 50 breast cancers was not reached in 4 years, so the study was continued for 3 more years.

All women who provided consent for randomisation and were accordingly screened at least once, were assessed. Randomised women, who requested the screening protocol of the other group during follow-up remained in the analyses in the group they were randomised to (intention to screen analyses).

Women were excluded after randomisation when they ultimately proved to have a cumulative lifetime risk below 20%, or because of diagnosis with a *BRCA1/2* or *TP53* mutation. However, data up until exclusion were used in the analyses.

Tumour type (invasive or ductal carcinoma in situ), tumour stage stage (pT), lymph node status (pN) Bloom-Richardson grade (BR grade), estrogen receptor (ER) status, progesterone (PR) status, HER2 status and ductal carcinoma in situ grade were compared between the randomisation groups, using two-sided Fisher's exact tests; and age at detection, tumour size and ductal carcinoma in situ size using Mann-Whitney *U* test.

Numbers of cancers were calculated per 1000 screening rounds or woman-years at risk and compared using Exact Rate Ratio Test assuming Poisson counts. Corresponding 95% confidence intervals (CIs) were calculated using a Poisson distribution. Woman-years at risk were calculated from the first screening examination to the date of discontinuation, bilateral prophylactic mastectomy, detection of invasive cancer, reaching the age of 60, death, or one year after the last screening visit to be able to account for interval cancers. In case a woman was lost to follow up after a screening visit, one year at risk was added after the last screening moment, with the same aim.

Interval cancers were defined as cancers diagnosed between two screening rounds due to symptoms, while the result of the previous screening round was negative. We performed linkage with the Dutch national Pathology Registry PALGA in 2017 and Jan 1 2019, in order not to miss interval cancers. To compare biopsy rates and false positive rates between the randomisation group, we used the Exact Rate Ratio Test. Sensitivity was calculated by dividing the number of screen-detected breast cancers by the number of all cancers.

Specificity was defined as the proportion of negative screens of all screens in women without breast cancer. Positive predictive value was calculated by dividing the number of screen-detected cancers by the number of positive screening tests (BI-RADS ≥3). Positive predictive value for biopsy was calculated by dividing the number of breast cancers by the number of biopsies. To compare sensitivity, specificity and Positive predictive value between the randomization groups, we used Fisher's exact test, and confidence intervals were calculated using the Clopper-Pearson interval. We also investigated these results of incident screens only (all screens after the first screening round).

The results of the registration groups were only used to examine the influence of density on the screening in both groups. Therefore all analyses were also performed after combining the reg-MRI-group with the MRI-group and reg-Mammography-group with the Mammography-group.

To test for linear trends in the number of breast cancers, interval cancers, pT stage, sensitivity, specificity and the number of false positive results when stratified by both BI-RADS breast density and automated Volpara density, we used Linear-by-Linear association tests.

To determine the level of agreement between the automated density measures and BI-RADS density estimates by the radiologists, Cohen's Kappa was calculated. A post-hoc analysis of pT stage, pN status and specificity stratified by age (<50 years, ≥50 years) was performed per group.

Analysis data consisted of all data of women during the study up until withdrawal or the end of the study follow-up. In case a woman was withdrawn from the study, data until withdrawal was included in the analyses. Statistical analyses were performed using IBM SPSS Statistics (version 24) and RStudio (version 1.0.44). A two-sided p-value ≤0.05 was considered statistically significant. No independent data monitoring committee oversaw the study.

This trial was registered with the Netherlands Trial Register, number NTR2789.

Role of funding source

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

RESULTS

Inclusion and follow up took place from Jan 1, 2011 until Dec 30, 2017. In total, 1355 women provided consent for randomisation and 231 for registration (Figure 1). Figure 2 shows the number of participants per screening round. In total, 675 women were randomised to the MRI-group, and 680 women to the Mx-group. Data of all randomised women were included in the analyses, except for one woman in the MRI-group due to a breast cancer diagnosis after randomisation but before the first MRI screening. Therefore, our modified intention to treat analyses contained 674 women in de MRI-arm. The average number of screening rounds per woman was 4.3 (SD:1.76), and the median follow-up after inclusion was 5.2 years for both groups (IQR MRI-group 3.4-6.2; IQR Mammography-group 3.6-6.3). Of the randomised women, 57 requested the screening protocol of the other group during follow-up (MRI-group 45/675, 7%; Mammography-group 13/680, 2%) but in the analyses they remained in the group they were randomised to (intention to screen analyses). Before the end of follow-up, 234 women withdrew (MRI-group 107/675, 16%; Mx-group 127/680, 19%), of whom 166 (71%) did not provide a reason. Thirteen women were excluded after randomisation, because they ultimately proved to have a cumulative lifetime risk below 20%, another twelve because of diagnosis with a *BRCA1/2* mutation and one with a *TP53* mutation. Table 1 shows the characteristics of the participants per group (Table S1 including registered participants). MRI and mammography were mostly performed on the same day, median 1 day (IQR: 0-14) , on average 12.8 days (SD 26.6), between the MRI and mammogram in the incident rounds.

In the randomised participants, 55 cancers (32 invasive cancers, 23 ductal carcinoma in situ) were detected: 24 invasive cancers and 16 ductal carcinoma in situ in the MRI-group, and 8 invasive cancers and 7 ductal carcinoma in situ in the Mammography-group (*p*0·002). Table 2 lists details of the cancers per group. No bilateral breast cancers were detected and none was metastasized. Two triple-negative cancers were detected

in the Mx-group, one in the MRI-group. Invasive cancers in the MRI-group were smaller than those in the Mammography-group (median size 9 versus 17 mm; p=0·010). Fourteen (58%) cancers were ≤10 mm in the MRI-group, versus one (13%) in the Mammography-group. The difference in pT-stage was significant in the incident cancers (p=0·035).

Invasive breast cancers in the MRI-group were less often node positive compared with the Mammography-group (4/24; 17% versus 5/8; 63%, p=0·023). BR grade, ER/PR/HER2 status, ductal carcinoma in situ grade and ductal carcinoma in situ size were not statistically significantly different.

Results after combining the results of the registration groups to the randomisation groups are shown in table S2.

Table 3 shows the performance of the two screening strategies within the randomisation groups. The number of breast cancers per 1000 screening rounds was significantly higher in the MRI-group than in the Mammography-group (14·2 versus 4·9 per 1000 screening rounds; p<0·001). The difference decreased and was no longer significant after the first screening round (MRI-group: 10·0 versus Mammography-group: 5·9; p=0·722). Figure 2 shows the incidence per group per screening round. One of the forty cancers in the MRI-group was an interval cancers and two of the fifteen cancers in the Mammography-group. The interval cancer in the MRI-group occurred ten months after screening (T2, node positive, BI-RADS density D). One interval cancer in the Mammography-group occurred nine months after screening (T2, node positive, BI-RADS density C), the second in the year after closure of the study (T1c, node negative, BI-RADS density B). Sensitivity was non-significantly higher and specificity was significantly lower in the MRI-group, compared with the Mammography-group.

The fourteen (61%) out of 23 invasive screen-detected cancers in the MRI-group detected by MRI only were eight T1a/T1b, five T1c, one T2; two (T1c) of them node positive; and the three (13%) out of 23 detected by mammography only were one T1a, two T1b. One ductal carcinoma in situ grade 1 was detected by MRI only, one by mammography only and three by both MRI and mammography.

Specificity was significantly lower in the MRI-group due to more false positive results. Of the false positive results in the MRI-group, 22% (99/450) resulted from a positive mammogram while MRI was negative. In incident rounds, specificity improved in both groups.

Positive predictive value (for BI-RADS ≥3) was non-significantly higher in the MRI-group than in the Mammography-group, whereas Positive predictive value for biopsy was similar.

With increasing breast density, numbers of detected cancers (including reg-group data) increased significantly for the mammography protocol (p=0·018). All ≥T2 tumours (N=7) and three of five interval cancers when including the registered women, were in the two highest density categories (Table 4). Estimated by radiologists, MRI detected only more early stage cancers, and with ≤ 25% positive nodes, in the three lower density categories A-C, in which it performed best. However, with automated (Volpara density) measurements this was also true for the highest density category (D).

Sensitivity did not differ significantly with increasing density in either protocol. Specificity rates decreased with increasing density for both screening protocols (p<0·001), as false positives were increasing (p<0·001). Automated breast density measures were available for 80% of the participants and in slight agreement with the density assessments by radiologists, with a kappa of 0·205.[21] However, results stratified by automated density grades (Appendix, table S4) were in accordance with those of BI-RADS breast density stratification.[19] When stratifying our results by age (women <50 years versus women ≥50 years), we observed no difference in tumour and nodal stage in either group (Table 2), but higher specificity in both groups in women aged ≥50 years compared to women <50 years(Table 3).

Median follow-up of patients after a breast cancer diagnosis was 4.3 years. None of the breast cancer patients of the randomisation groups died during follow-up. One patient in the reg-Mammography-group died from breast cancer. The trial was ended the year that the number of cancers as determined by the power calculation was reached.

DISCUSSION

In this randomized controlled trial, the first to our knowledge comparing MRI screening with mammography in high risk women, median tumour size of the invasive breast cancers was 8 mm smaller ($p<0·010$) in the MRI-group and they were far more often node negative20 (83%) of 24 versus 3 (38%) of 8 $p=0·023$) than in the Mammography-group. Even more important for the effectiveness in the long run, are the results of the incident rounds: MRI screening resulted in lower numbers of late stage cancers (≥T2 1/18 versus 2/8; and node positive cancers 2/18 versus 5/8 $p=0.014$), both tumour stage and nodal status were significantly more favourable. MRI screening may lead to a substantial mortality reduction, as in a study of 93,569 patients diagnosed with primary breast cancer in the Netherlands in 2006-2012, five-year relative survival for T1c tumours was 98%, decreasing in T2 tumours to 92%; for all N0 tumours 98%, but 86% for N2 tumours, despite up to date adjuvant therapy.[4] Furthermore, with substantially fewer node positive patients, less adjuvant chemotherapy will be indicated, sparing many women the early and late side effects and cost.We certainly intend to publish in the future, as prespecified, 10 year mortality results after linkage with our national database, as mortality reduction is the aim of screening. The current follow-up, in which one patient from the Reg-Mammography-group died from breast cancer, is too short, but tumour stage is a reliable proxy.[3,4]

Our MRI protocol caused an impressive favourable stage-shift compared to the mammography protocol, while the tumour stages in our Mx-group were very comparable to the stages detected by MRI screening in older multicentre studies in the familial group.[5-7] This demonstrates how much both mammography and MRI have improved over the last decade. Therefore, we cannot use the results of those older studies anymore[6] to estimate the pros and cons of MRI screening as already shown by Obdeijn et al.[22]

Not unexpectedly, MRI had the clear disadvantage of more false positive results and thus lower specificity for the MRI protocol. Despite improvement in the incident rounds of both randomisation groups, the difference between the groups remained significant and substantial. This is in accordance with other high-risk MRI screening studies.[6] The false positive rate is maybe explained by the average young age of our participants as it clearly decreased above the age of 50 years and increased with increasing density, and it may furthermore be the consequence of the very early stage at detection. The ACR expects a Positive predictive value of 24% for performed biopsies, our Positive predictive value is just above this[19]. Our Positive predictive value for BI-RADS ≥3 and for biopsy are comparable to the Positive predictive value 's in two recent cohort studies of MRI-screening.[11,23]

Another drawback of screening is overdiagnosis. The incidence of all cancers was higher in the MRI-group than in the Mammography-group. The difference declined after the first round and was not significant anymore, although incidence remained higher up till the fourth round (figure 2). However, with a mean age at detection of 49 years, at which the average Dutch life expectancy is 35 additional years, hardly any of the invasive cancers are expected to be overdiagnosis, as even early stage estrogen positive breast cancers may have metastasized after 20 years.[3] Nevertheless, the substantial increase in ductal carcinoma in situ is expected to be partly overdiagnosis: especially ductal carcinoma in situ grade 1 for which trials are ongoing to investigate whether active surveillance is safe.[24,25]

A possible unwanted side-effect of MRI-screening is retention of minute amounts of Gadolinium in brain and other tissues after MRI investigations, although less with the macrocyclic gadolinium products used in our 12 hospitals, of which so far no harmful effect has been identified. A letter we sent to all participants in the MRI-group in 2016 on this new evidence did not lead to substantial withdrawal of participants.

Both when breast density was estimated by radiologists as well as measured fully automated (Volpara density), all ≥T2 tumours, as well as most of the interval cancers were only seen in the two highest density categories. Estimated by radiologists, MRI detected only more early stage cancers, and with ≤ 25% positive nodes, in the three lower density categories A-C, in which it performed best. However, with automate Volpara density measurements this was also true for the highest density category (D).With increasing breast density, specificity decreased in both the MRI protocol and Mammography protocol, consistent with the results of Kerlikowske et al.[12] Density seems to be more important than age when choosing a screening strategy.

Also previous studies concluded that mammography is of limited additional value to MRI screening in women with familial risk.[26] Of the false positive results in the MRI-group 22% was caused by mammography only. On the other hand three (12%) minimal cancers (≤1 cm) with a considerable area of ductal carcinoma in situ with microcalcifications were only detected by our low frequent mammography. We do not know in which stage they would have been detected by MRI, but either an even lower frequency of mammography should be considered or omitting mammography. Clinical breast examination generated also substantial false positive results in both groups, and detected only one of all cancers, making the additional value of clinical breast examination negligible.

Studies have demonstrated that digital breast tomosynthesis (DBT) has the potential to increase screening sensitivity and specificity in comparison with digital mammography. Therefore, if DBT would have replaced mammography, we would expect a gain in diagnostic accuracy in the Mammography-group. However the average additional cancer yield published for DBT is 1·2 per 1000 cases [27,28], for additional ultrasound 3·5-4·4 per 1000 cases (with considerable false positive rate increase: the reason why it is not recommended in Dutch screening guidelines)[9,29], but for MRI 15·5 per 1000 cases.[30] We therefore do not expect as large a stage shift as demonstrated in our MRI-group, could have been attained with either DBT or ultrasound.

A limitation of our study is, that being powered to demonstrate a difference in tumour size between the two screening groups, the numbers of the detected cancers were small when stratified according to density or age categories. Maybe therefore we were unable to demonstrate a significant sensitivity decrease with increasing density which has been shown previously.[12] Importantly, in the MRI-group the number of later stage cancers decreased clearly in incident rounds, but we also have to evaluate long-term survival, and cost-effectiveness. Another limitation is, that previous screening may have influenced our incidence. However, a nearly equal amount of previous MRI-screening was performed in both groups: ≤ 2 years 9·2% in the MRI-group and 11·9% in the Mammography-group. It may have reduced the incidence possibly more in the Mammography-group. Fortunately the study continued for 7 years, with an average of 4·3 screening rounds per person. We see in figure 1 the highest cancer incidence in the Mammography-groupat the second year and a nearly equally steep decline in both groups thereafter. This suggests it is a quite limited influence for the complete study and it will not have influenced our primary endpoint: tumour size and nodal status.

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

We conclude that in real-life practice the MRI-screening causes an important favourable stage shift compared to mammography and can by reducing late-stage cancers reduce chemotherapy use and mortality. Certainly at density categories A-C,.but especially in density D, at the cost of lower specificity. Clinical breast examination may be omitted and the frequency of mammography beside MRI-screening may be further reduced.

Declaration of interests

The authors declared no conflicts of interest"

Data Sharing

De-identified participant data will be made available 6 months after the trial primary and secondary endpoints have been published. Any requests for trial data and supporting material (data dictionary, protocol, and statistical analysis plan) will be reviewed by the trial PI group in the first instance. Requests that have a methodologically sound proposal and whose proposed use of the data has been approved by the trial steering committee and do not interfere with planned analyses, e.g. cost-effectiveness and 10 yr. mortality results, will be considered in the first instance. To gain access, data requestors will need to sign a data access agreement.

REFERENCES

1.   Margolin S, Johansson H, Rutqvist LE, Lindblom A, Fornander T. Family history, and impact on clinical presentation and prognosis, in a population-based breast cancer cohort from the Stockholm County. *Fam Cancer* 2006; **5**(4): 309-21.

2.   Brandt A, Bermejo JL, Sundquist J, Hemminki K. Age of onset in familial breast cancer as background data for medical surveillance. *BR J Cancer* 2010; **102**(1): 42-7.

3.   Pan H, Gray R, Braybrooke J, et al. 20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years. *N Engl J Med* 2017; **377**(19): 1836-46.

4.   Saadatmand S, Bretveld R, Siesling S, Tilanus-Linthorst MMA. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients. *BMJ* 2015; **351**: h4901.

5.   Kriege M, Brekelmans CTM, Boetes C, et al. Efficacy of MRI and mammography for breast-cancer screening in women with familial or genetic predisposition. *New England Journal of Medicine* 2004; **351**: 427-37.

6.   Phi XA, Houssami N, Hooning MJ, et al. Accuracy of screening in women at familial risk of breast cancer without a known gene mutation: individual patient data meta-analysis. *European Journal of Cancer* 2017; **85**: 31-8.

7.   Leach MO, Boggis CR, Dixon AK, et al. Screening with magnetic resonance imaging and mammography of a UK population at high risk of breast cancer: a prospective multicenter cohort study (MARIBS). *Lancet* 2005; **365**(9473): 1769-78.

8.   Smith RA, Andrews KS, Brooks D, et al. Cancer screening in the United States, 2017: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin* 2017; **67**(2): 100-21.

9.   Richtlijn mammacarcinoom (Breast Cancer National Guideline). 2017. (accessed 2019-03-05 at https://www.oncoline.nl/borstkanker).

10.   NICE Clinical Guideline 164. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. 2017.(accessed 2018-04-15 https://www.nice.org.uk/guidance/cg164/chapter/Recommendations#surveillance-and-strategies-for-early-detection-of-breast-cancer

11.   Bick U, Engel C, Krug B, et al. High-risk breast cancer surveillance with MRI: 10-year experience from the German consortium for hereditary breast and ovarian cancer. *Breast Cancer Res Treat* 2019.

12. Kerlikowske K, Scott CG, Mahmoudzadeh AP, et al. Automated and clinical breast imaging reporting and data system density measures predict risk for screen-detected and interval cancers: a case-control study. *Ann Intern Med* 2018; **168**(11): 757-65.

13. Lehman CD, Gatsonis C, Kuhl CK, et al. MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer. *N Engl J Med* 2007; **356**(13): 1295-303.

14. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. *American Journal of Roentgenology* 2012; **198**(3): W292-W5.

15. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, et al. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst* 2013; **105**(17): 1314-21.

16. Saadatmand S, Rutgers EJ, Tollenaar RA, et al. Breast density as indicator for the use of mammography or MRI to screen women with familial risk for breast cancer (FaMRIsc): a multicenter randomized controlled trial. *BMC Cancer* 2012; **12**: 440.

17. Claus EB, Risch NR, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer* 1994; **73**(3): 643-51.

18. Kuhl CK, Schrading S, Bieling HB, et al. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *Lancet* 2007; **370**(9586): 485-92.

19. American College of Radiology. American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas). Reston, Va: American College of Radiology; 2003.

20. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. *IEEE Trans Med Imaging* 2006; **25**(3): 273-82.

21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**(1): 159-74.

22. Obdeijn IM, Winter-Warnars GA, Mann RM, Hooning MJ, Hunink MG, Tilanus-Linthorst MMA. Should we screen BRCA1 mutation carriers only with MRI? A multicenter study. *Breast Cancer Res Treat* 2014; **144**(3): 577-82.

23. Lee JM, Ichikawa L, Valencia E et al. Performance Benchmarks for Screening Breast MR Imaging in Community Practice. Radiology. 2017 Oct;285(1):44-52.

24. Elshof LE, Tryfonidis K, Slaets L, et al. Feasibility of a prospective, randomized, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ—the LORD study. *Eur J Cancer* 2015; **51**(12): 1497-510.

25. Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen-detected DCIS; the LORIS trial. *Eur J Cancer* 2015; **51**(16): 2296-303.

26. Vreemann S, van Zelst JCM, Schlooz-Vries M, et al. The added value of mammography in different age-groups of women with and without BRCA mutation screened with breast MRI. *Breast Cancer Res* 2018; **20**: 84.

27. Gilbert FJ, Tucker L, Young KC. Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool. *Clin Radiol* 2016; **71**(2): 141-50.

28. Zackrisson S, Lang K, Rosso A, et al. One-view breast tomosynthesis versus two-view mammography in the Malmo Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. *Lancet Oncol* 2018; **19**(11): 1493-503.

29. Berg WA, Blume JD, Cormack JB, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* 2008; **299**(18): 2151-63.

30.   Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women with Average Risk of Breast Cancer. *Radiology* 2017; **283**(2): 361-70.

Table 1. Characteristics of randomised women at baseline, according to study group

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

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

13

Figure 1. Trial profile

1768 women were invited

413 were not randomised:
- 53 did not meet inclusion criteria or met exclusion criteria
- 14 because of a contraindication
- 102 because of psychological reasons or claustrophobia
- 73 did not want additional tests or had no time for it
- 31 did not want MRI screening
- 11 wanted MRI screening
- 107 provided no reason
- 22 because of oth

1355 randomised

Of the declining women:
218 participated in the Reg-Mammography-group

675 were allocated to the MRI-group

680 were allocated to the Mammography-group

107 withdrew themselves
54 were excluded
- 26[a] because of breast cancer diagnosis
- 5 because of other cancer diagnosis
- 6 ultimately proved to have a cumulative lifetime risk below 20%
- 2 underwent prophylactic mastectomy
- 8 proved to be BRCA1/2 mutation carriers
- 6 because of a contraindication
- 1 was included at a too old age

127 withdrew themselves
27 were excluded
- 7 because of breast cancer diagnosis
- 3 because of other cancer diagnosis
- 7 ultimately proved to have a cumulative lifetime risk below 20%
- 5 underwent prophylactic mastectomy
- BRCA1/2 mutation carriers
- Li-Fraumeni syndrome

674[a] were included in the intention to screen analy

680 were included in intention to screen the analyses

MRI screening was performed. Due to our exclusion criteria, we excluded this woman from our analysis data.
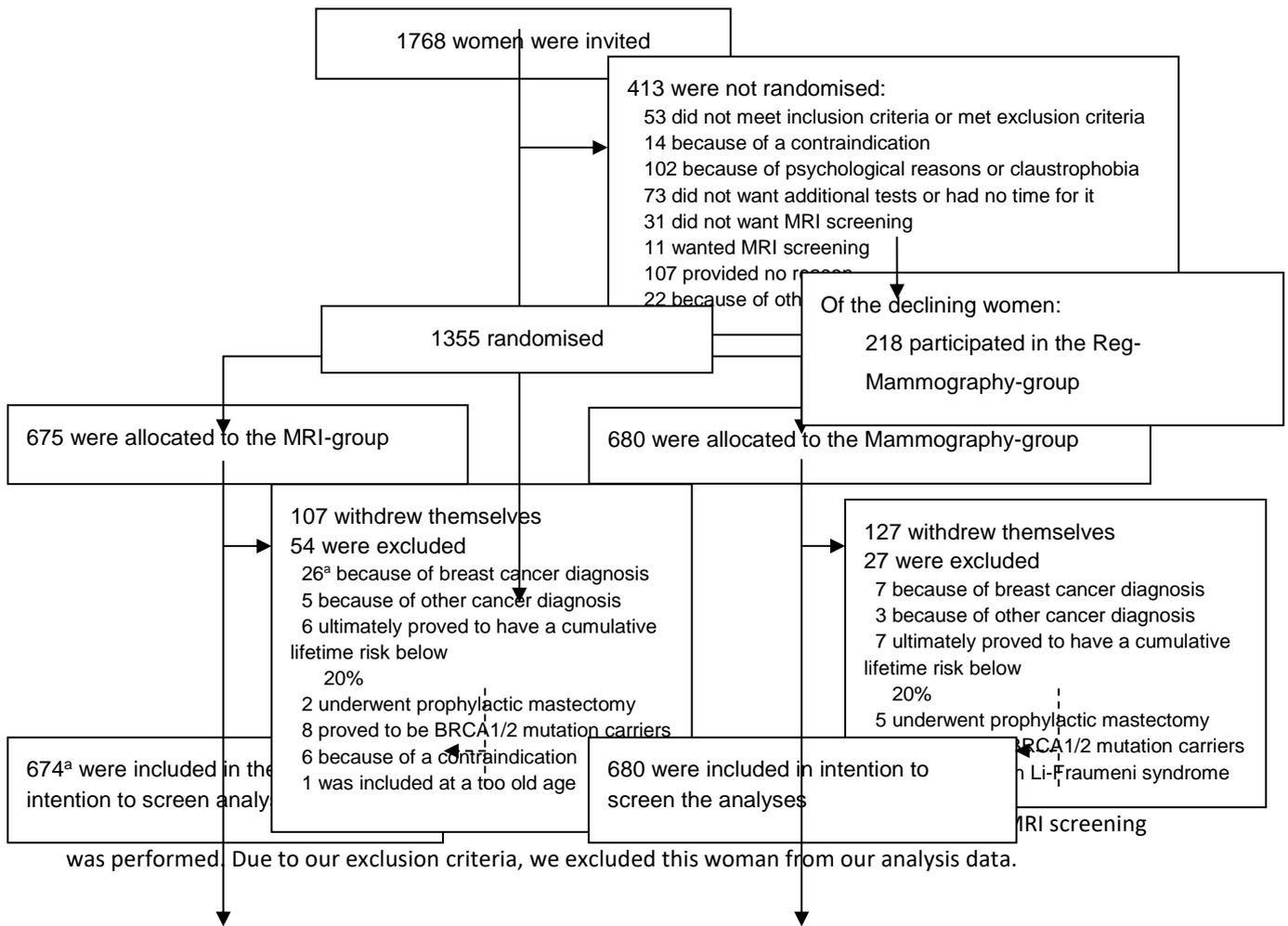
14

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

	MRI-group (N=674)	Mammography-group (N=680)	Comparison MRI-group **vs** Mammography-group, *p*-value
Mean age at detection (years ± SD)	49·4 ± 7·1	50·0 ± 4·6	0·880
No cancer	634 (94%)	665 (98%)	
Invasive breast cancers	24 (4%)	8 (1%)	
ductal carcinoma in situ	16 (2%)	7[a] (1%)	0·002
Size of invasive cancers (mm ± SD)	Mean: 11·9 ± 12·3 Median: 9	Mean: 18·0 ± 8·1 Median: 17	0·010
T stage			
T1a	7/24 (292%) ⎤	0 ⎤	
T1b	7/24 (29%) ⎬ 14 (58%)	1/8 (13%) ⎬ 1 (12·5%)	
T1c	7/24 (29%) ⎦	5/8 (63%) ⎦	0·065[b]
T2	2/24 (8%) ⎤	2/8 (25%) ⎤	
T3	1/24 (4%) ⎬ 3 (13%)	0 ⎬ 2 (25%)	
T4	0 ⎦	0 ⎦	
Node status			
Positive	4/24 (17%)	5/8 (63%)	
Negative	20/24 (83%)	3/8 (38%)	0·023
BR grade			
1	10/24 (41·7%)	2/8 (25%)	
2	9/24 (37·5%)	3/8 (38%)	0·504
3	4/24 (16·7%)	3/8 (38%)	
Missing	1/24 (4·2%)	0	
ER positive	22/24 (91·7%)	6/8 (75%)	0·254
PR positive	18/24 (75%)	5/8 (63%)	0·654
HER2 positive	2/24 (8%)	0	1.000
Ductal carcinoma in situ grade			
1			
2	5/16 (31%)	2/7 (29%)	
3	8/16 (50%)	4/7 (57%)	1.000
	3/16 (19%)	1/7 (14%)	
ductal carcinoma in situS size (mm ± SD)	Mean: 34·18 ± 43·8 Median: 14[c]	Mean: 30·29 ± 26·9 Median: 20	1.000
T stage incident rounds			
Tis	7/ 25 (28%)	7/15 (47%)	
T1a+T1b	12/25 (48%)	1/15 (7%)	0·035
T1c	5/25 (20%)	5/15 (33%)	
≥ T2	1/25 (4%)	2/15 (13%)	
Node status incident rounds			
Positive	2/18 (11%)	5/8 (63%)	
Negative	16/18 (89%)	3/8 (38%)	0·014
T stage <50 years			
Tis	7/18 (39%)	5/8 (63%)	
T1a+T1b	6/18 (33%)	0	

T1c	4/18 (22%)	1/8 (13%)	
≥T2	1/18 (6%)	2/8 (25%)	0·125[d]
T stage ≥50 years			
Tis	9/22 (409%)	2/7 (29%)	
T1a+b	8/22 (36%)	1/7 (14%)	
T1c	3/22 (14%)	4/7 (57%)	
≥T2	2/22 (9%)	0	0·180 [d]
Node status, <50 years			
Positive	1/11 (9%)	3/3 (100%)	
Negative	10/11 (91%)	0	0·011 [d]
Node status ≥50 years			
Positive	3/13 (23%)	2/5 (40%)	
Negative	10/13 (77%)	3/5 (60%)	0·583 [d]

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

b. Based on categories 'T1a+T1b'; 'T1c' and 'T2+'

c. Contains missing values

d. *p*-value for T stage MRI-group < 50 versus ≥ 50 years: 0·928; *p*-value for T stage Mammography-group < 50 versus ≥ 50 yrs: 0·092; *p*-value for node status MRI-group < 50 versus ≥ 50 years: 0·596; *p*-value for node status Mammography-group < 50 versus ≥ 50 years: 0·196

Table 3. Screen-detected cancers, interval cancers, detection rates, detection technique, biopsies, false positives, sensitivity, specificity, and PPV, according to study group and rounds

	MRI-group (N=674)	Mammography-group (N=680)	Comparison MRI-group vs Mammography-group, *p*-value
Screening rounds	2812	3075	
Woman-years at risk	3220	3326	
Screen-detected cancers	39[a]/40 (98%)	13[b]/15 (87%)	
Interval breast cancers	1/40 (3%)	2/15 (13%)	0·177
No. of breast cancers per 1000 screening rounds (95% CI)			
All breast cancers	14·2 (10·0-18·8)	4·9 (2·6-7·5)	<0·001
Screen-detected cancers	13·9[a] (9·6-18·5)	4·2[b] (2·0-6·8)	<0·001
Invasive screen-detected cancers	8·2 (5.0-11·7)	2·0 (0·7-3·6)	0·001
ductal carcinoma in situ	5·7[a] (3·2-8·5)	2·3[b] (0·7-4·3)	0·058
Interval cancers per 1000 woman-years at risk (95% CI)	0·3 (0·0-0·9)	0·6 (0·0-1·5)	1·000
Detection technique of invasive cancers			
Mammography[c]	3/23 (13%)	6/6 (100%)	
MRI[c]	14/23 (61%)	n.a.	
Both mammography and MRI	5/23 (22%)	n.a.	
clinical breast examination only	1/23 (4%)	0	
Biopsies (rate[d])	149 (53·0)	54 (17·6)	<0·001
False positives (rate[d]) ≥BI-RADS 3	449 (159·7)	276 (89·8)	<0·001
By mammography[c]	98/449 (22%)	157/276 (57%)	
By MRI[c]	275/449 (61%)	9/276 (3%)[e]	
By both mammography and MRI[c]	19/449 (4%)	0	
By clinical breast examination only	57/449 (13%)	110/276 (40%)	
Sensitivity (95% CI)	97·5% (86·8-99·9)	86·7% (59·5-98·3)	0.177
Specificity (95% CI)	83·8% (82·4-85·2)	91·0% (89·9-92·0)	<0·001
Positive predictive value BI-RADS ≥ 3 (95% CI)	8·0% (5·7-10·7)	4·5% (2·4-7·6)	0·074
Positive predictive value for biopsy (95% CI)	26·8% (20·0-34·7)	27·8% (16·5-41·6)	1·000
Incident screening rounds	2141	2407	
No. of breast cancers in incident round per 1000 screening rounds (95% CI)	10·0 (6·4-14·0)	5·9 (3·2-9·1)	0·722
Screen-detected in incident rounds	25[a]/25 (100%)	13/15 (87%)	
Interval cancers in incident rounds	0	2/15 (13%)	0·135
Biopsies in incident rounds (rate[d])	82 (38·3)	38 (15·8)	<0·001
False positives in incident rounds (rate[d])	266 (124·2)	176 (73·1)	<0·001
Sensitivity in incident rounds (95% CI)	100·0% (86·3-100·0)	86·7% (59·51-98·3)	0·135
Specificity in incident rounds (95% CI)	87·4% (85·9-88·8)	92·6% (91·5-93·7)	<0·001
Positive predictive value BI-RADS ≥ 3 (95% CI) in incident rounds	8·6% (5·6-12·4)	6·9% (3·7-11·5)	0·605
Positive predictive value for biopsy in incident rounds (95% CI)	30.5% (20·8-41·6)	39·5% (24·0-56·6)	0·538
Specificity < 50 yrs	81·9% (80·1-83·6)	89·6% (88·2-90·9)	<0·001[f]

| Specificity ≥ 50 yrs | 87·7% (85·4-89·8) | 93·5% (91.9-94·9) | <0·001 [f] |

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

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

c. Possibly in combination with a positive clinical breast examination

d. Rate per 1000 screening rounds

e. These false positives occurred in women who requested the MRI protocol while being randomised to the Mammography-group (see paragraph 'Study population' of the Results)

f. p-value for results of specificity difference < 50 versus ≥ 50 yrs for MRI-group: <0·001; for Mammography-group <0·001; and Positive predictive value difference <50 versus ≥ 50 yrs for MRI-group <0·001; and for Mammography-group <0·001

Table 4. All breast cancers, T staging, and false positives by BI-RADS density categories

	BI-RADS breast density[a]				
	A	**B**	**C**	**D**	*p* **trend**
All participants	N=206	N=549	N=562	N=239	
Screening rounds	993	2412	2413	1022	
All breast cancers (rate[b])	5 (5.0)	22 (9.1)	27 (11.2)	11 (10.8)	0.132
Interval cancers (rate[b])	0	2 (0.8)	2[c] (0.8)	1 (1.0)	0.474
T stage					
Tis	1 (20%)	8 (36%)	11 (41%)	5 (50%)	0.108
T1a + T1b	2 (40%)	5 (23%)	8 (30%)	1 (9%)	0.978
T1c	2 (40%)	9 (41%)	4 (15%)	2 (18%)	0.489
T2+	0	0	4 (15%)	3 (27%)	0.008
Node status					
Positive	1 (25%)	3 (21%)	6 (38%)	3 (50%)	
Negative	3 (75%)	11 (79%)	10 (63%)	3 (50%)	
MRI-group + **RegMRI**-group	N=86	N=249	N=238	N=105	
Screening rounds	403	1033	973	440	
All breast cancers (rate[b])	5 (12.4)	15 (14.5)	17 (17.5)	5 (11.4)	0.916
Interval cancers (rate[b])	0	0	0	1 (2.3)	0.104
Sensitivity (95% CI)	100.0% (47.8-10.0)	100.0% (78.2-100.0)	100.0% (80.5-100.0)	80.0% (28.4-99.5)	0.079
Specificity (95% CI)	90.5% (87.1-93.2)	85.3% (82.9-87.4)	82.8% (80.3-85.2)	77.0% (72.8-80.9)	<0.001
False positives (rate[b])	38 (94.3)	150 (145.2)	164 (168.6)	100 (227.3)	<0.001
T stage					
Tis	1 (20%)	7 (46.7%)	7 (41.2%)	1 (20%)	0.969
T1a + T1b	2 (40%)	5 (33.3%)	7 (41.2%)	0	0.535
T1c	2 (40%)	3 (20.0%)	2 (11.8%)	1 (20%)	0.424
T2+	0	0	1 (5.9%)	3 (60%)	0.007
Node status					
Positive	1 (25%)	0	2 (20.0%)	2 (50%)	
Negative	3 (75%)	8 (100.0%)	8 (80.0%)	2 (50%)	
Mammography-group + **RegMammography**-group	N=120	N=300	N=324	N=134	
Screening rounds	590	1379	1440	582	
All breast cancers (rate[b])	0	7 (5.1)	10[c] (6.9)	6 (10.3)	0.018
Interval cancers (rate[b])	0	2 (1.5)	2 (1.4)	0	0.992
Sensitivity (95% CI)	n.a.	71.4% (29.0-96.3)	80.0% (44.4-97.5)	100.0% (54.1-100.0)	0.181
Specificity (95% CI)	93.7% (91.5-95.5)	93.0% (92.3-93.6)	89.0% (87.3-90.6)	86.3% (83.2-90.0)	<0.001
False positives (rate[b])	37 (62.7)	96 (69.6)	157 (109.0)	79 (135.7)	<0.001
T stage					
Tis	0	1 (14%)	4 (40%)	4 (67%)	0.007
T1a + T1b	0	0	1 (10%)	1 (17%)	0.124
T1c	0	6 (86%)	2 (20%)	1 (17%)	0.843
T2+	0	0	3 (30%)	0	0.348
Node status					

Positive	0	3 (50%)	4 (67%)	1 (50%)
Negative	0	3 (50%)	2 (33%)	1 (50%)

a. BI-RADS breast density estimated at baseline

b. Rate per 1000 screening rounds

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

Figure 2. Incidence of all cancers (invasive + ductal carcinoma in situ) per screening round, according to randomisation group

No. of women per screening round:
Mx-group   668   640   587   481   372
MRI-group   672   622   542   440   317