Contrast-enhanced mammography versus conventional imaging in women recalled from breast cancer screening (RACER trial): a multicentre, open-label, randomised controlled clinical trial

Lidewij M. F. H. Neeter,[a,b] Patricia J. Nelemans,[c] H. P. J. Raat (Frank),[d] Caroline Frotscher,[e] Katya M. Duvivier,[f] Brigitte A. B. Essers,[g,h] Marjolein L. Smidt,[b,i] Joachim E. Wildberger,[a] and Marc B. I. Lobbes[a,b,e,*]

[a]Maastricht University Medical Center, Department of Radiology and Nuclear Medicine, P.O. Box 5800, 6202AZ, Maastricht, the Netherlands
[b]GROW School for Oncology and Reproduction, Maastricht University, P.O. Box 616, 6200MD, Maastricht, the Netherlands
[c]Maastricht University, Department of Epidemiology, P.O. Box 616, 6200MD, Maastricht, the Netherlands
[d]Laurentius Hospital, Department of Radiology, P.O. Box 920, 6040 AX, Roermond, the Netherlands
[e]Zuyderland Medical Center, Department of Medical Imaging, P.O. Box 5500, 6130 MB, Sittard-Geleen, the Netherlands
[f]Amsterdam University Medical Center, Department of Radiology, P.O. Box 7057, 1007MB, Amsterdam, the Netherlands
[g]Maastricht University Medical Center, Department of Clinical Epidemiology and Medical Technology Assessment, P.O. Box 5800, 6202AZ, Maastricht, the Netherlands
[h]CAPHRI Care and Public Health Research Institute, Maastricht University, P.O. Box 616, 6200MD, Maastricht, the Netherlands
[i]Maastricht University Medical Center, Department of Surgery, P.O. Box 5800, 6202AZ, Maastricht, the Netherlands

The Lancet Regional Health - Europe 2024;44: 100987
Published Online xxx
https://doi.org/10.1016/j.lanepe.2024.100987

Summary

Background Women recalled from breast cancer screening receive post-screening work-up in the hospital with conventional breast imaging. The RACER trial aimed to study whether contrast-enhanced mammography (CEM) as primary imaging instead of conventional imaging resulted in more accurate and efficient diagnostic work-up in recalled women.

Methods In this randomised, controlled trial (registered under NL6413/NTR6589) participants were allocated using deterministic minimisation to CEM or conventional imaging as a primary work-up tool in two general and two academic hospitals. Predefined patients' factors were reason for recall, BI-RADS score, and study centre. Primary outcomes were sensitivity and specificity. Secondary outcomes were the proportion of women needing supplemental examinations, and number of days until diagnosis.

Findings Between April, 2018, and September, 2021, 529 patients recalled from the Dutch screening program were randomised, 265 to conventional imaging and 264 to CEM. Three patients in the control arm had to be excluded from analysis due to a protocol breach. After the entire work-up, sensitivity was 98.0% (95% CI; 92.2–99.7%) in the intervention arm and 97.7% (91.8–99.6%) in the control arm (p = 1.0), and specificity was 75.6% (72.5–76.6%) and 75.4% (72.5–76.4%, p = 1.0), respectively. Based on only primary full-field digital mammography/digital breast tomosynthesis or CEM, final diagnosis was reached in 27.7% (73/264) in the intervention arm and 1.1% (3/262) in the control arm. The frequency of supplemental imaging was significantly higher in the control arm (p < 0.0001). Median time needed to reach final diagnosis was comparable: 1 day (control arm: IQR 0–4; intervention arm: IQR 0–3). Thirteen malignant occult lesions were detected using CEM, versus three using conventional imaging. No serious adverse events occurred.

Interpretation Diagnostic accuracy of CEM in the work-up of recalled women is comparable with conventional imaging. However, work-up with CEM as primary imaging is a more efficient pathway.

Funding ZonMw (grant number 843001801) and GE Healthcare.

*Corresponding author. Zuyderland Medical Center, Department of Medical Imaging, P.O. Box 5500, 6130 MB, Sittard-Geleen, the Netherlands.
E-mail address: m.lobbes@zuyderland.nl (M.B.I. Lobbes).

Keywords: Breast neoplasm; Mammography; Screening

Research in context

Evidence before this study

We searched PubMed on July 1, 2017, for studies written in English with no restrictions on publication date, with the terms 'contrast-enhanced (spectral) mammography (CEM), breast cancer screening, and breast cancer. We found several clinical trials studying the use of CEM in women recalled from breast cancer screening, showing that the sensitivity and specificity of CEM are superior to full-field digital mammography (FFDM) in recalled women. Unfortunately, these were all single-centre, retrospective studies with relatively small sample sizes. This promising, but limited evidence, formed the basis of the current randomised controlled clinical trial, performed in different centres using equipment of different vendors.

Added value of this study

To our knowledge, this is the first randomised controlled clinical trial on the clinical use of CEM in women recalled from screening. It shows that although the sensitivity and specificity of CEM and conventional imaging are comparable when they are used as primary imaging tools for these women, the work-up route with CEM results in the use of fewer resources. There are fewer additional imaging or tissue sampling examinations required and more occult breast lesions are detected when using CEM instead of conventional imaging.

Implications of all the available evidence

The collective evidence shows that CEM as a work-up tool in women recalled from breast cancer screening is more efficient in terms of resources needed, and it detects more occult lesions than the control group. However, we will also perform a future cost-effectiveness study to evaluate whether this new approach will also result in a more cost-effective strategy for these women.

Introduction

The Dutch breast cancer screening program invites women between 50 and 75 years biennially to participate, using full-field digital mammography (FFDM) as an imaging modality.[1] About 2.4% of all participants will be recalled for further work-up via one of two routes: women with a Breast Imaging and Reporting and Data System (BI-RADS) 0 recall (*i.e.*, low suspicion of breast cancer) are recalled directly to a Radiology department, whereas women with a BI-RADS 4 or 5 recall (*i.e.*, high suspicion of breast cancer) will also visit an outpatient breast clinic, where a baseline history and physical examination is performed combined with imaging assessment during the same appointment.[2] CEM can help to limit the number of additional examinations and thereby reduce psychological distress and anxiety that is often experienced by women recalled from breast cancer screening for a suspicious finding detected on FFDM.

In most hospitals, the primary imaging modality in the work-up of these recalled women is FFDM. When deemed necessary, supplemental imaging in the form of digital breast tomosynthesis (DBT), ultrasound (US), or breast magnetic resonance imaging (MRI), with or without tissue sampling, can be added to reach the final diagnosis. During work-up, breast cancer is ruled out in 70% of recalled women.[3] These (false positive) recalls cause unnecessary anxiety and downstream testing, with the latter also increasing healthcare costs.[4,5] Also, re-attendance rates after a false positive recall are known to be lower than in women who were not recalled.[6]

Repeating FFDM as the primary imaging modality after a recall is debatable, as there are other complementary mammographic modalities available, such as contrast-enhanced mammography (CEM). Studies have shown that the diagnostic performance of CEM is consistently superior to FFDM, even matching the diagnostic performance of breast MRI.[7–11] Previous studies have specifically shown the clinical feasibility, and the promising accuracy and reproducibility of CEM as a work-up tool in screening recalls.[12,13] However, these single-centre studies were primarily retrospective in design.[12–14]

The RACER trial is the first multicentre, randomised controlled clinical trial using CEM compared to conventional imaging (such as FFDM or DBT) as primary imaging modality to evaluate the post-screening work-up in women recalled from breast cancer screening. Primary study outcomes were diagnostic accuracy expressed as sensitivity and specificity, and secondary outcomes were the proportion of women needing additional imaging and/or tissue sampling, and the number of days until a final diagnosis was reached.

Methods

Study design

The detailed study protocol of this multicentre, open-label, randomised controlled clinical trial has been published earlier.[15] All women (natal sex) recalled from the Dutch breast cancer screening program who attended one of four participating centres in The Netherlands, two general and two academic hospitals, were eligible for participation. Patients were excluded from participation in case of known hypersensitivity

reactions to iodinated contrast or known acute or chronic severe renal insufficiency (*i.e.,* estimated glomerular filtration rate below 30 mL/min/1.73^2).

The study was approved by our institutional review board (reference (METC171082/NL62788.068.17). Local ethical approval was obtained in each study centre. The study was registered in the Netherlands Trial Registry (NL6413/NTR6589). Since this study is marked as a 'low-risk study' by the Clinical Trial Center Maastricht, a data monitoring committee was not commissioned.[15]

Investigators of the participating centres provided the study information to patients by phone and -after oral consent-also by e-mail. All participants provided written informed consent.

Randomisation

Group allocation was performed using deterministic minimisation to ensure the balancing of study arms concerning predefined patient factors: predominant reasons for recall (mass, calcification, asymmetry, or architectural distortion), recall BI-RADS score (BI-RADS 0 versus BI-RADS 4/5), and study centre.[15] For this purpose, the computer-generated randomisation screening and enrolment application software ALEA (version 3.0.2083.212r, ALEA Clinical, Abcoude, the Netherlands) was used.

Imaging study protocol

Patients underwent CEM as the primary imaging tool when allocated to the intervention study arm. The CEM image acquisition protocol has previously been described in detail.[16] CEM was followed by supplemental imaging and/or tissue sampling when indicated by the radiologist. Additional breast imaging could consist of US of the breast and/or axilla, spot compression views, extended craniocaudal lateral views or other special views, DBT, and breast MRI. Standard image acquisition protocols for these modalities were used.

Conventional breast imaging modalities (*e.g.,* FFDM, DBT) were used as primary imaging tools when participants were allocated to the control study arm. A minimum of one view had to be repeated to allow comparison with the screening FFDM. Primary imaging could be further supplemented with any other breast imaging modality, except CEM, when indicated by the radiologist.

Cysts were aspirated to confirm its aqueous content in the control arm, while the presence of an 'eclipse sign' on CEM was applied for this diagnosis in the intervention arm.[12,17] Tissue sampling was performed of suspicious solid lesions or calcifications in both study arms. A follow-up appointment after three, six, or twelve months could be used in case of inconclusive findings, irrespective of the study arm.

Primary and secondary outcomes

The primary outcomes and the most important secondary outcomes are presented in this article. Primary outcomes were diagnostic accuracy parameters sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), the receiver operating characteristic (ROC) curve, and area under the ROC curve (AUC). The final BI-RADS score during the first visit was based on the primary imaging and supplemental imaging combined. BI-RADS scores of 1–3 were considered 'benign', and BI-RADS scores of 4–5 were considered 'malignant'. Histopathological results were used as reference standard to confirm true positive or true negative disease status. For patients without tissue sampling, true negative disease status was confirmed after a follow-up of 24 months. When no interval cancer was mentioned in the patient's hospital file within this period, the final disease status was defined as 'true negative'.

Information on supplemental imaging, tissue sampling, and outpatient breast clinic visits was collected from the hospital files per patient. The need for additional examinations during the initial visit and follow-up visits was recorded.

The number of days until final diagnosis was calculated from the day of enrolment, which was the date of the first recall visit, until the day of a final diagnosis. The date on which pathology results became available was used as date of final diagnosis when tissue sampling was performed. The date of the last follow-up visit was used as date of final diagnosis for women for whom follow-up examination between 3 and 12 months was recommended. When a follow-up appointment was indicated, but the patient was a 'no show', we decided to appoint 90 days for a 3-month follow-up, 180 days for a 6-month follow-up, and 365 days for a 12-month follow-up as the date of the final diagnosis.

The remaining secondary outcome parameters that were mentioned in the published trial design[15] (*i.e.,* quality-of-life, cost-effectiveness, and experienced patient anxiety) will be published in a second article.

Sample size calculation

Prior research showed a specificity of 40% for FFDM in this population. To enable the detection of a clinically relevant increase of specificity of 15%, and assuming a power of 80% (alpha 5%) and a prevalence of malignant disease status in this population of approximately 30% (*i.e.,* 70% have true negative disease status), 251 patients needed to be included per group. To account for a 5% loss to follow-up, the minimum number of patients that needed to be included was 528 (502/0.95).[15]

Statistical analysis

The modified intention-to-treat analysis and the per-protocol population included all patients who were randomised and received either conventional imaging or CEM according to the protocol for whom the true disease status was known. In the intention-to-treat analysis, patients were analysed in the arm to which they were assigned by randomisation. In the

per-protocol analysis, patients were analysed according to the imaging modality that they received.

Sensitivity, specificity, PPV, and NPV, were calculated and compared between study arms using the Fisher's exact test, reporting the p-value as double the exact one-tailed probability. The area under the receiver operating curve (AUC) was compared using an algorithm suggested by DeLong et al. for paired samples using the command "roccomp" in STATA.[18] Proportions of women diagnosed at the initial visit were compared between arms using the Chi-square test. The number of supplemental imaging examinations as well as the number of days to reach a final diagnosis was expressed as median with interquartile range (IQR) and the difference between arms was tested for significance using the Mann–Whitney test.

p-values <0.05 were considered statistically significant. Statistical analyses were performed with SPSS (version 28, IBM Corporation, Armonk, New York, USA) and STATA (version 17, StataCorp LCC, College Station, Texas, USA).

Role of the funding source

This trial received funding from the Netherlands Organisation for Health Research and Development (ZonMw Efficiency Studies Grant Number 843001801) and GE Healthcare. ZonMw was involved in the trial design. GE Healthcare was not involved in data collection, analyses, or manuscript writing. No authors were precluded from accessing data in the study. All authors accept responsibility to submit for publication.

Results

Baseline characteristics and adverse events

A total of 1267 patients were eligible to participate between April 3rd, 2018, and September 14th, 2021. Excluded were 738 potential study candidates, leaving 529 women (42%) for the randomisation process, in which 265 patients were allocated to the control arm and 264 patients to the intervention arm. Women declining to participate in the trial was the major reason for non-participation. The condition that at least one FFDM and/or DBT view had to be repeated was not met in three patients in the control arm. These patients were excluded from further analyses. A flowchart of the patient randomisation process is presented in Fig. 1.

Table 1 shows that the distribution of baseline characteristics in both randomised groups was comparable. Most women were recalled for a mass (54.8%; 288/526), followed by calcification (23.6%; 124/526), asymmetry (12.5%; 66/526), and architectural distortion (9.1%; 48/526), which was comparable between the two groups. Thirty-five patients were recalled for two lesions and one patient for three. Malignant disease was present in 37.9% (100/264) in the intervention arm and 33.2% (87/262) in the control arm.

An adverse event occurred in 1.5% of patients (4/264) who underwent intravenous contrast administration during CEM: two mild hypersensitivity reactions to the contrast agent and two contrast extravasations occurred, none of which required medical intervention. One patient with contrast extravasation did not undergo CEM but received FFDM instead. Since no other crossovers occurred, we decided that per-protocol analyses were unnecessary.

Two patients deceased during their 24 months of follow-up. Both patients died of metastatic renal cell carcinoma. However, histopathology determined true negative disease status in one patient and true positive disease status for breast cancer in the other. All other patients had a minimum follow-up of 24 months.

Diagnostic accuracy

Diagnostic parameters (Table 2) were comparable for both the control arm (conventional imaging) and the intervention arm (CEM). After entire post-screening work-up, sensitivity was 97.7% (85/87) for the control arm and 98.0% (98/100) for the intervention arm (p = 1.0). Specificity was 75.4% (132/175) and 75.6% (124/164), respectively (p = 1.0). PPV was with 71.0% (98/138) higher in the intervention arm than the 66.4% (86/128) in the control arm (p = 0.50), while NPV was comparable (98.4% (124/126) versus 98.5% (132/134), respectively; p = 1.0). The AUC-value was comparable for both study arms: 0.866 (95% CI; 0.821–0.910) in the control arm and 0.868 (95% CI; 0.824–0.912) in the intervention arm (p = 0.93).

Supplemental examinations to reach a diagnosis

The process of the diagnostic work-up after recall to reach a final diagnosis is complex and is described in Fig. 2. The flowchart represents a schematical model of the work-up that (in some cases) consisted of multiple recursive loops in the process towards a diagnosis.

Table 3 shows the extent of the diagnostic work-up that was required to reach the final diagnosis for both study arms. A diagnosis was reached with only a single FFDM and/or DBT in 1.1% (3/262) of the patients, versus 27.7% (73/264) using solely CEM in the intervention group (p < 0.0001). The three patients with a final diagnosis based on only FFDM and/or DBT in the control arm were referred for a BIRADS 0 lesion. Of the 73 patients in the intervention arm in whom CEM without supplemental imaging resulted in a final diagnosis, 66 were recalled for a BIRADS 0 lesion and seven for a BIRADS 4 lesion.

In one patient from the control group core needle biopsy (CNB) was performed, but no supplemental imaging. Supplemental imaging was required in 258 patients in the control arm and 191 patients in the intervention arm. In most patients, US of at least one breast was part of supplemental imaging and was performed in 98.8% (255/258) and 96.9% (185/191) of

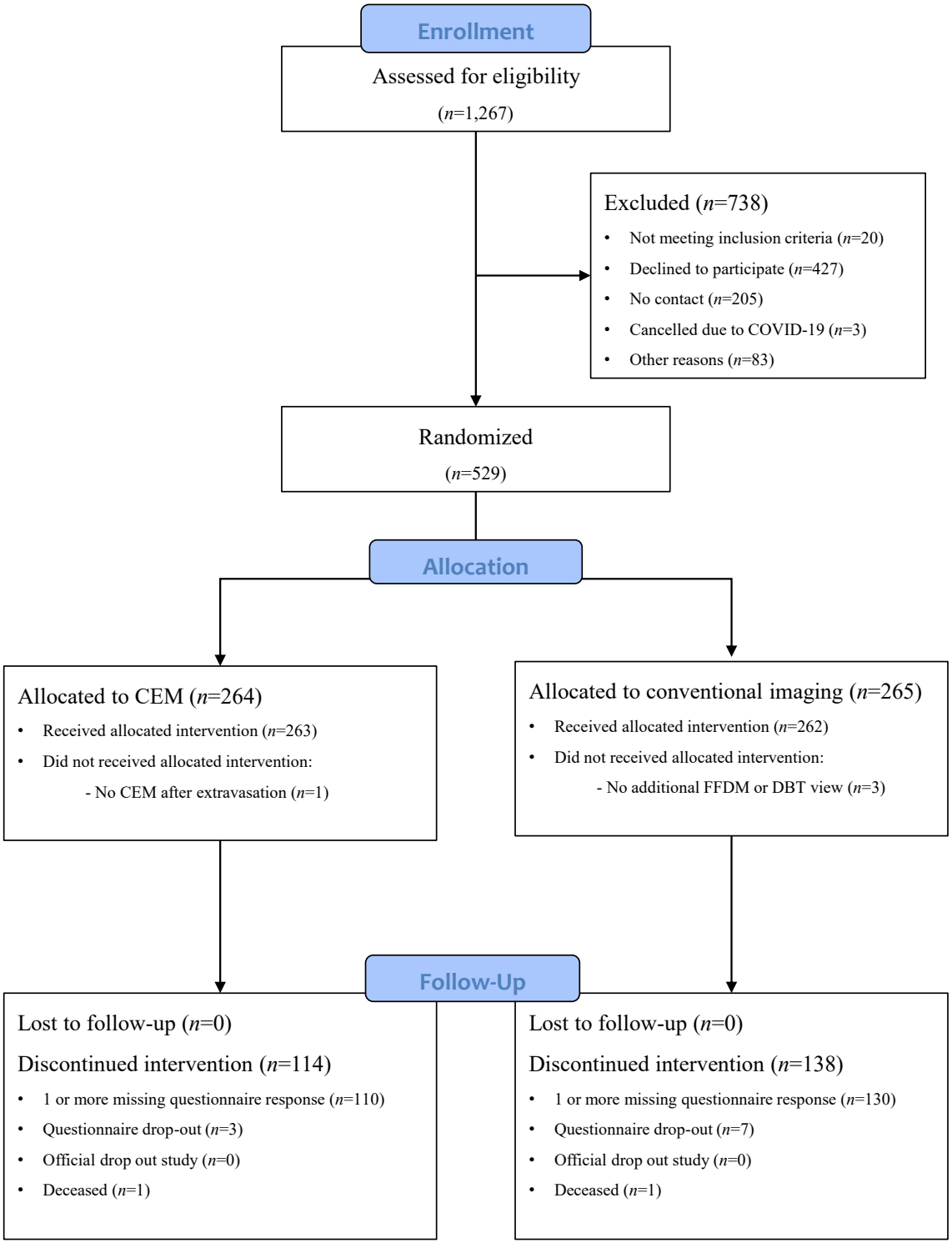

Fig. 1: CONSORT flow chart of patient inclusion.

Patient and lesion characteristics	Control arm n = 262	Intervention arm n = 264
Mean age in years (SD)	59 (8)	60 (8)
BI-RADS score recall		
0	140/262 (53.4%)	141/264 (53.4%)
4	104/262 (39.7%)	109/264 (41.3%)
5	18/262 (6.9%)	14/264 (5.3%)
Reason for recall		
Mass	143/262 (54.6%)	145/264 (54.9%)
Calcifications	62/262 (23.7%)	62/264 (23.5%)
Asymmetry	34/262 (13.0%)	32/264 (12.1%)
Architectural distortion	23/262 (8.8%)	25/264 (9.5%)
Number of lesions		
1	244/262 (93.1%)	246/264 (93.2%)
2	18/262 (6.9%)	17/264 (6.4%)
3	0	1/264 (0.4%)
Prevalence of malignancy	87/262 (33.2%)	100/264 (37.9%)
Breast cancer subtype	N = 87	N = 100
Invasive carcinoma NST	47/87 (54.0%)	69/100 (69.0%)
Invasive lobular carcinoma	6/87 (6.9%)	7/100 (7.0%)
Ductal carcinoma *in situ*	26/87 (29.9%)	20/100 (20.0%)
Other invasive breast cancer	8/87 (9.2%)	4/100 (4.0%)

Abbreviations - SD: standard deviation; BI-RADS: breast imaging reporting and data system; NST: no special type.

Table 1: **Distribution of patient and lesion characteristics in both randomised arms.**

patients in the control and intervention arm, respectively. The number of times patients needed additional imaging was significantly higher in the control group ($p < 0.0001$). In the control group, 38.5% (101/262) of the women had three or more rounds of additional imaging versus 22.0% (58/264) in the intervention arm.

Tissue sampling was required in 74.4% (195/262) in the control arm and 58.3% (154/264) in the CEM arm ($p < 0.0001$). Cyst aspiration and fine needle aspiration cytology (FNAC) were performed significantly more often in the control group. The frequency of breast or axillary CNB was 53.1% (139/262) in the control group and 60.6% (150/264) in the CEM group. The extent of the work-up in the cases that required follow-up examination is given in Supplementary Table SI.

	Control arm n = 262 (%) (95% CI) [absolute numbers]	Intervention arm n = 264 (%) (95% CI) [absolute numbers]	p-value
Sensitivity	97.7 (91.8–99.6) [85/87]	98.0 (92.9–99.7) [98/100]	1.0
Specificity	75.4 (72.5–76.4) [132/175]	75.6 (72.5–76.6) [124/164]	1.0
PPV	66.4 (62.4–67.7) [85/128]	71.0 (67.3–72.2) [98/138]	0.50
NPV	98.5 (94.7–99.7) [132/134]	98.4 (94.4–99.7) [124/126]	1.0
AUC	0.866 (0.821–0.910)	0.868 (0.824–0.912)	0.93

Abbreviations - CI: confidence interval; PPV: positive predictive value; NPV: negative predictive value; AUC: area under the ROC curve.

Table 2: **Diagnostic parameters in both study arms.**

Time needed to reach a final diagnosis

Definitive diagnosis, with or without tissue sampling and without the need for follow-up, was reached in 92.8% (245/264) in the intervention arm and 90.1% (236/262) in the control arm. One patient in the intervention arm and eight patients in the control arm did not undergo follow-up evaluation ('no shows'). At the follow-up visit, CNB was performed in two patients in the control arm and one patient in the intervention arm, confirming malignancy in the latter case in the intervention arm.

The distribution of the days until diagnosis for both study arms is presented in Table 4. There was no significant difference in days until diagnoses between study arms ($p = 0.17$). The median of days until final diagnosis was 1 day in both arms (control arm: IQR 0–4; intervention arm: IQR 0–3). It was possible to reach a definitive diagnosis on the day of enrolment in 37.8% (99/262) of participants in the control arm and 40.5% (107/264) of participants in the intervention arm. Over 80% of women had a definitive diagnosis within 5 days in both study arms. In 5.3% (14/262) of participants in the control arm and 4.5% (12/264) of participants in the intervention arm, it took over six months to reach a final diagnosis.

Occult lesions

Additional (pre)malignant lesions, *i.e.*, not the lesions for which women were initially recalled, were detected more often in patients in the intervention arm (13 versus 3, respectively). Two of these additional lesions were ductal carcinoma *in situ*. All other occult lesions in both study arms were invasive cancers. Moreover, in the intervention arm, detection of the occult lesion led to the diagnosis of unilateral breast cancer in two patients (while the initial lesion for which the women were recalled was benign), bilateral breast cancer in four patients, and multifocal breast cancer in the remaining seven patients. In the control arm, detection of the occult lesion led to the diagnosis of unilateral breast cancer in one patient (while the lesion for which the patient was recalled was caused by superimposition of breast tissue), bilateral breast cancer in a second patient, and multifocal breast cancer in a third patient.

It should be noted that diagnostic accuracy in this study was based on per-patient level, not on per-lesion level. Consequently, the detection of these additional lesions hardly affected overall diagnostic accuracy. An example of bilateral cancer detection by CEM is given in Fig. 3. An overview of the lesions characteristics of all occult lesions is given in Supplementary Table SII.

Discussion

CEM is an emerging breast imaging technique with a diagnostic accuracy superior to FFDM. Retrospective studies have demonstrated the excellent problem-solving

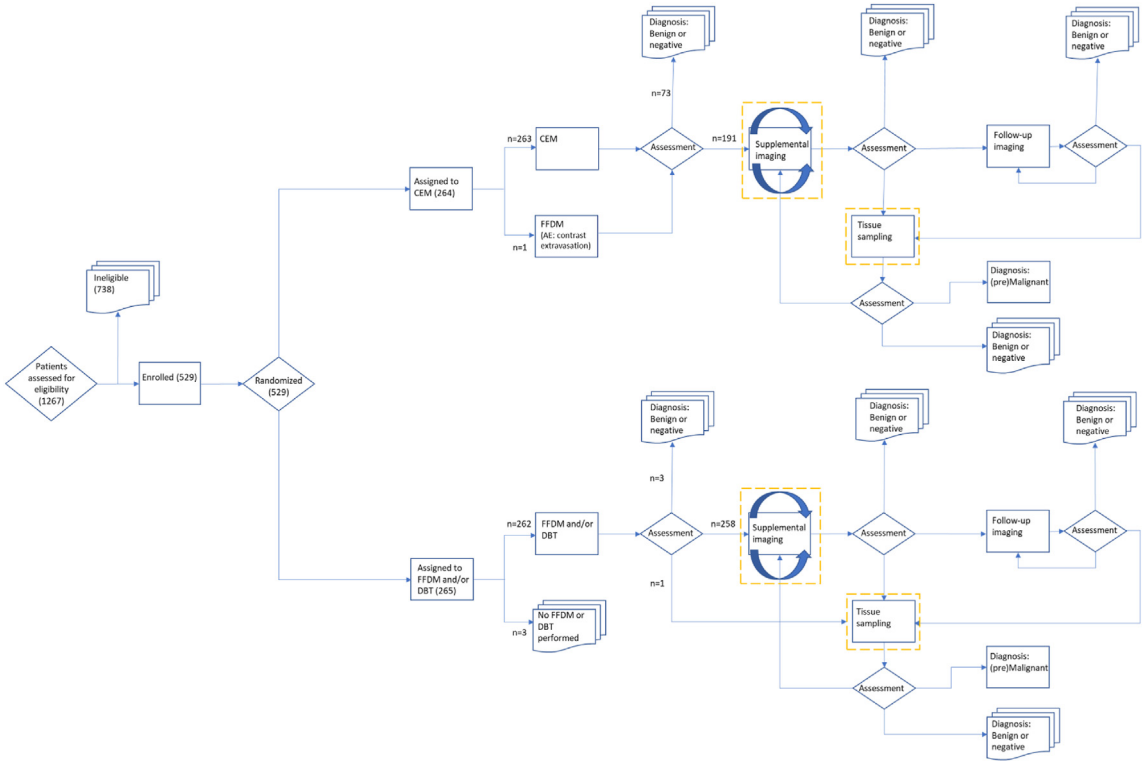

Fig. 2: Trial profile with a schematic overview of the work-up performed in both study arms. Blue arrows emphasize the recursive process. Details of the orange boxes are in Table 3. Abbreviations - CEM: contrast-enhanced mammography; FFDM: full-field digital mammography; DBT: digital breast tomosynthesis.

capabilities of CEM in women recalled from the Dutch screening program.[12,13] The RACER trial is the first multicentre, randomised, controlled clinical trial that compares the diagnostic accuracy and efficiency of CEM and conventional imaging as primary work-up tools in recalled women. Our results showed that the diagnostic accuracy of both study arms was comparable. Sensitivity and specificity were 98.0% (95% CI; 92.9–99.7) and

	Control arm n = 262 (absolute numbers)	Intervention arm n = 264 (absolute numbers)	p-value
Primary imaging only	1.1% (3/262)	27.7% (73/264)	<0.0001[c]
Supplemental imaging			
1x	16.4% (43/262)	8.7% (23/264)	
2x	43.5% (114/262)	41.7% (110/264)	
3x	13.0% (34/262)	5.7% (15/264)	
4x	19.5% (51/262)	14.4% (38/264)	
5x	5.7% (15/262)	1.5% (4/264)	
6x	0.4% (1/262)	0.0% (0/264)	
7x	0.0% (0/262)	0.4% (1/264)	
Tissue sampling	74.4% (195/262)[a]	58.3% (154/264)[b]	<0.0001
Cyst aspiration	18.3% (48/262)	2.7% (7/264)	<0.0001
FNAC	6.9% (18/262)	1.5% (4/264)	0.0044
Core needle biopsy	53.1% (139/262)	60.6% (150/264)	0.44

NB. In one patient in the control arm, tissue sampling was performed without supplemental imaging (0.4%). Abbreviations - FNAC: fine needle aspiration cytology. [a]30 patients had two and six patients three different tissue samples taken in the control arm. [b]41 patients had two, nine patients three, and one patient five different tissue samples taken. [c]p-value from Mann–Whitney U test.

Table 3: **Extent of diagnostic work-up that was required to reach a final diagnosis in both study arms.**

Diagnosis	Control arm	Intervention arm	p-value
Diagnosis at initial visit	n = 262	n = 264	0.39
Benign/negative	149/262 (60.7%)	146/264 (55.3%)	
Malignant	87/262 (33.2%)	99/264 (37.5%)	
Inconclusive and follow-up required	26/262 (6.1%)	19/264 (7.2%)	
Days until diagnosis	n = 262	n = 264	
Median	1	1	0.17[a]
IQR	0–4	0–3	
Range	0–748	0–653	
Category			
0 days	37.8% (99/262)	40.5% (107/264)	
1–5 days	45.4% (119/262)	45.8% (121/264)	
6–14 days	6.1% (16/262)	5.7% (15/264)	
15–180 days	5.3% (14/262)	3.4% (9/264)	
181–365 days	3.4% (9/262	3.4% (9/264)	
>365 days	1.9% (5/262)	1.1% (3/264)	

Abbreviations – IQR: interquartile range. [a]p-value from Mann–Whitney U test.

Table 4: **Results on visits and days until final diagnosis in both study arms.**

75.6% (95% CI; 72.5–76.6) for the intervention arm and 97.7% (95% CI; 91.8–99.6) and 75.4% (95% CI; 72.5–76.4) for the control arm. However, the use of CEM in the studied population was more efficient: CEM use required fewer additional examinations in the form of imaging or tissue sampling and resulted in the detection of more occult lesions when compared with the control group. The time needed to reach a final diagnosis was comparable between the study arms with a median of one day.

Systematic reviews and meta-analyses comparing the diagnostic accuracy of FFDM with CEM have consistently shown that CEM is superior.[9–11] In clinical practice, however, FFDM/DBT or CEM are often used in combination with US. Within the RACER trial, the radiologist was allowed to use additional tools, such as DBT, US, tissue sampling, MRI, or even follow-up to establish the final diagnosis, like in everyday clinical practice. Within the intervention (CEM) arm, fewer supplemental examinations were required to achieve this goal. For example, breast US was less frequently used because the sensitivity of CEM itself is already very high. Sorin et al. even demonstrated that low-threshold use of US in addition to CEM should be avoided as it can decrease overall specificity.[19] Even the use of breast MRI as a problem-solving modality in equivocal findings on conventional imaging decreased in women undergoing CEM.[19] After all, recent reviews have repeatedly shown no large differences in sensitivity and specificity between CEM and breast MRI.[8,20,21]

If a cyst was suspected on imaging in the control arm, it needed to be visualised by ultrasound and aspirated, after which at least one control FFDM/DBT view was required to confirm the disappearance of the lesion. In the CEM arm, a cyst diagnosis was often made based on CEM, where cysts appear as an 'eclipse sign' on the recombined image.[12,17] Consequently, the number of cyst aspirations dropped from 18.3% (48/262) to 2.7% (7/264) in the CEM arm without loss of sensitivity, proving that the 'eclipse sign' is indeed a reliable imaging feature on CEM to diagnose cysts.

It was possible to reach a final diagnosis without further diagnostic work-up in 27.7% (73/264) of participants in the intervention arm, versus only 1.1% (3/262) of the participants in the control arm. CEM can help to limit the number of additional examinations and thereby reduce psychological distress and anxiety that is often experienced by recalled women with suspicious findings on screening mammograms. A study by Van der Steeg et al. on 385 recalled women showed that women with a false positive finding suffer from feelings of anxiety, which could last for at least one year.[22]

Regardless of the study arm, a final diagnosis was achieved in most women on the day of or the day after enrolment, and otherwise often within 14 days (89.3% (234/262) in the control arm and 92.0% (243/264) in the intervention arm). This outcome was unexpected because we hypothesised that with conventional imaging more time might be needed to reach the diagnosis since more examinations were required. This might be explained by the well-organised outpatient breast clinics and (breast) Radiology departments, with access to different imaging modalities often within a few days.

Importantly, CEM detected more (occult) malignant lesions compared to conventional imaging. Houben et al. retrospectively found that CEM led to the detection of 70 breast lesions that were occult on FFDM in 839 women.[14] Of these, 38 (4.5%) were diagnosed as breast cancer. In line with these observations, we detected additional malignant lesions in 4.9% (13/264) of women undergoing CEM, versus only in 1.1% (3/262) of women appointed to conventional imaging. Whereas the three occult lesions in the control arm were grade I (maximum size 6.5 cm), four of the thirteen occult lesions in the intervention arm were grade I (maximum size 1.0 cm), whilst the three grade III occult lesions varied in size between 0.8 and 2.7 cm, most of them being stage cT1. Especially when an occult lesion leads to the diagnosis of multifocal or bilateral breast cancer, the patient's treatment plan might change significantly, and failing to detect these lesions increases the risk of recurrent breast cancer. Occult lesions that we observed in this trial also demonstrated that invasive lobular carcinoma (ILC) and its multifocality are more difficult to detect on FFDM than on CEM or breast MRI.[23] The screening FFDM had missed multifocality of ILC, while two patients (grade II and grade III) on CEM and one patient (grade I) on MRI showed multiple additional lesions only a couple of days after the screening FFDM was performed.

Although the disadvantages of CEM include the use of iodinated contrast media and increased radiation exposure, the more efficient workflow that we observed

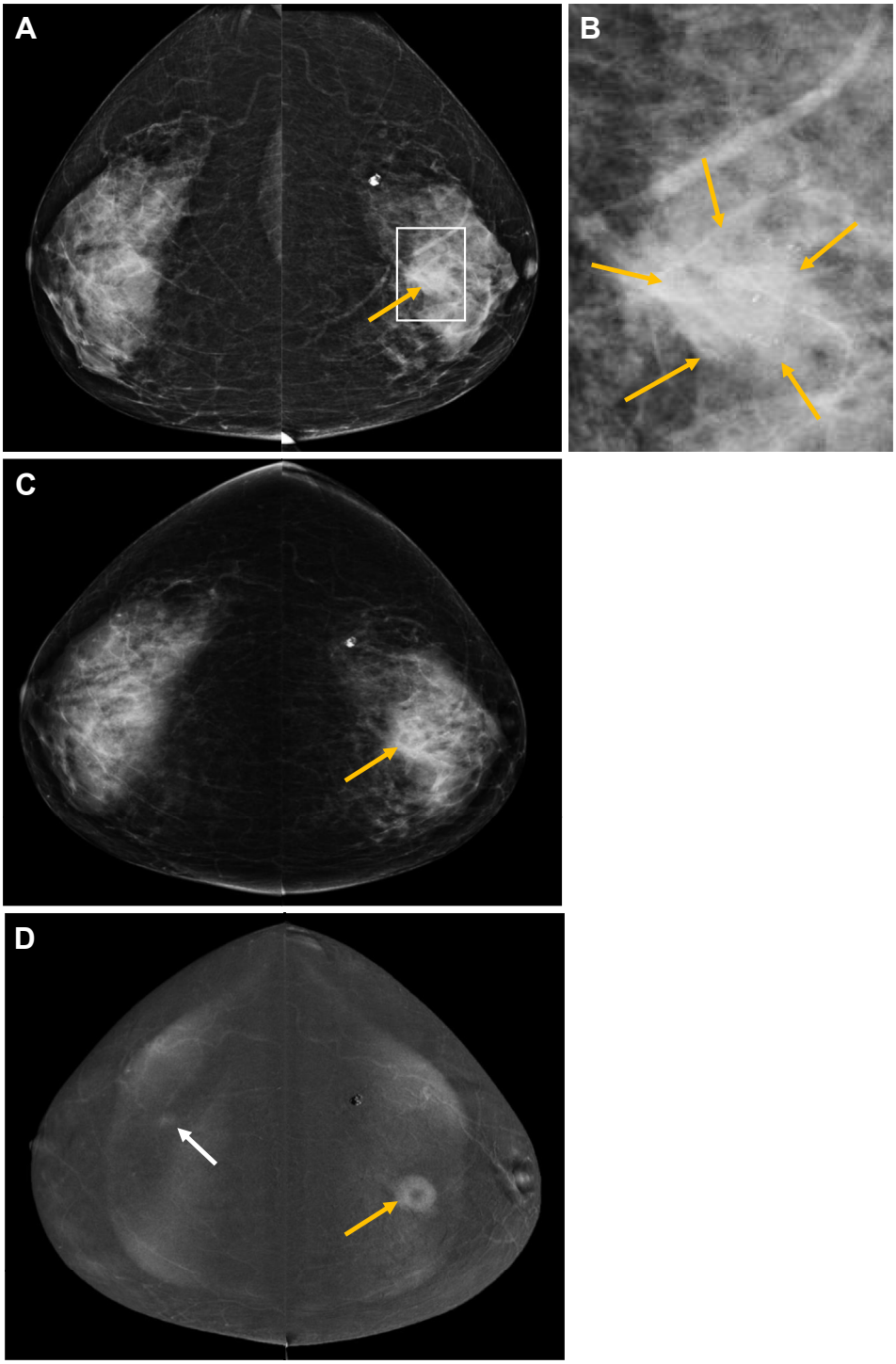

Fig. 3: Example of bilateral breast cancer detected by CEM in a 71-year-old woman. Only CC views are shown. (A) original screening mammogram showed an obscured irregular mass in the left breast with progressive fine pleiomorphic calcifications (orange arrows). (B) shows an outtake of the recalled lesion (C) shows the low-energy images of the CEM exam, similar to (A). However, the recombined images show not only an ill-defined mass in the left breast (orange arrow) but also a minimally enhancing mass in the left breast (white arrow) which was also biopsied after targeted ultrasound. Final pathology showed 12 mm papillary carcinoma *in situ* (right) and 30 mm grade II invasive carcinoma of no special type (left).

using CEM does not come at the expense of unacceptable risks. A limited number of extravasations can be expected in any intravenous access placement. The risk of hypersensitivity reactions to iodinated contrast agents was low (0.8%), with (self-limiting) mild symptoms. These observations were in line with previously published results by Houben et al.[14] Studies also showed that CEM increased radiation exposure by 81% from 1.6 mGy to 2.8 mGy.[24] However, in a population of women >50 years the risk of developing or dying from radiation-induced breast cancer is negligible. Yaffe et al. showed that the danger of radiation exposure is limited or perhaps even hypothetical. They suggested that in a cohort of 100,000 women (if screened with FFDM from 40 to 74 years) eleven deaths due to radiation-induced breast cancer could be expected.[25] Using CEM would increase this number, but only to a small extent. In addition, a Position Statement of the American Association of Physicists in Medicine (AAPM, policy number PS 4-A) stated that 'epidemiological evidence supporting increased cancer incidence or mortality from radiation doses below 100 mSv is inconclusive'.[26] Therefore, the increased radiation exposure of CEM should never serve as a deterrent when any clear benefit of CEM over FFDM is demonstrated.

This study has limitations. First, regarding the number of examinations, we believe that the number of US examinations performed might be higher in both study arms than medically required to establish a diagnosis.[27] All the participating radiologists of all centres, who had been reviewing examinations (>75) before the initiation of the RACER trial and had reached the level of experienced reader,[28] were instructed to rely as much as possible on CEM findings in this trial. Although the steep learning curve in reading a CEM,[13,28] as creatures of habit the threshold was probably low to also perform (bilateral) ultrasound to confirm a diagnosis which was already reliably established on CEM.[27] With growing confidence and experience in CEM, the requirement for supplemental US will perhaps diminish further. Second, the outcomes of this trial are specific to the breast cancer screening and care environment in the Netherlands. The Dutch screening recall rate was 2.4% in 2019 and therefore low compared to European and American programs, in which a recall rate of 5% and 5–12% is observed, respectively.[29,30] Therefore, differences between nations need to be considered when extrapolating the current trial results to other regions of the world. Third, we did not perform a cost-effectiveness analysis to demonstrate that the use of fewer resources in the CEM arm is also a more efficient workflow from a healthcare cost perspective. We intend to perform extensive cost-effectiveness analyses in a follow-up study.

Finally, to achieve comparability of randomised groups deterministic minimisation was used without a probabilistic element. However, this may not be a serious deficiency since in a multi-centre trial it is difficult for investigators to keep track of past assignments.

In conclusion, we showed that the risks associated with the use of CEM as a work-up tool for women recalled from screening were negligible. The diagnostic performance of the work-up of women recalled from breast cancer screening when using CEM as primary imaging tool is similar to when conventional imaging is used. However, using CEM as primary imaging modality in the recalled women is more efficient in terms of the number of examinations needed to establish a final diagnosis and it detects more (occult) breast cancers that were not screen-detected with FFDM. Therefore, CEM should be strongly considered as the preferred primary imaging modality in women recalled from breast cancer screening.

Contributors

L. Neeter: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, software, validation, visualisation, writing original draft, review and editing; P. Nelemans: conceptualisation, formal analysis, investigation, methodology, validation, visualisation, funding acquisition, writing original draft, review and editing; H. Raat: data curation, investigation, validation, visualisation, review and editing; C. Frotscher: data curation, investigation, validation, visualisation, review and editing; K. Duvivier: data curation, investigation, validation, visualisation, review and editing; B. Essers: conceptualisation, formal analysis, investigation, methodology, validation, funding acquisition, visualisation, writing original draft, review and editing; M. Smidt: investigation, validation, visualisation, review and editing; J. Wildberger: conceptualisation, investigation, methodology, validation, visualisation, funding acquisition, writing original draft, review and editing; M. Lobbes: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, software, supervision, validation, visualisation, funding acquisition, writing original draft, review and editing; M. Lobbes and L. Neeter directly assessed and verified the underlying data reported.

Data sharing statement

The study protocol of this trial has already been published and is available via Open Access. The pseudo-anonymised data set is available upon request for research purposes only. The authors will review the request based on novelty and feasibility of the research question before sharing pseudo-anonymised individual participant data. A data transfer agreement between institutes will be required. The final decision to share the study data resides with the principal investigator and sponsor.

Declaration of interests

M.L. received research funding from GE Healthcare, Hologic Inc., and Sirius Medical B.V. M.L. received honorary fees for presentations and participation of medical advisory boards from GE Healthcare, Hologic Inc., Sirius Medical B.V., Bayer, Guerbet, and Tromp Medical B.V. J.W. received institutional grants from Adora/Oldelft, Anaconda, Bard, Bayer, Bentley, Boston, Brainlab, GE Healthcare, Inari Medical, Johnson & Johnson, Merit Medical, Nico-Lab, Philips, Sectra, Siemens, and Stryker. J.W. received speaker's fees from Bayer and Siemens. L.N., K.D., P.N., H.R., C.F., M.S., and B.E. had no conflicts of interest to report.

Acknowledgements

We wish to acknowledge the Netherlands Organisation for Health Research and Development, better known as ZonMw (ZonMw Efficiency Studies Grant Number 843001801) and GE Healthcare for funding this study. During the preparation of this work the authors used Grammarly in order to check for residual English language and grammar errors. After using this tool/service, the authors reviewed and

edited the content as needed and take full responsibility for the content of the publication.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanepe.2024.100987.

References

1 Paap E, Verbeek ALM, Botterweck AAM, et al. Breast cancer screening halves the risk of breast cancer death: a case-referent study. *Breast*. 2014;23:439–444.
2 Rijksinstituut voor Volksgezondheid en Milieu (RIVM). *Uitvoeringskader bevolkingsonderzoek borstkanker*; 2023:1–83. Webpage https://www.rivm.nl/documenten/uitvoeringskader-bevolkingsonderzoek-borstkanker. Accessed February 13, 2024.
3 Rijksinstituut voor Volksgezondheid en Milieu (RIVM). *Bevolkingsonderzoek borstkanker - factsheet*; 2021. Webpage https://www.rivm.nl/documenten/factsheet-borstkanker. Accessed February 13, 2024.
4 Chubak J, Boudreau DM, Fishman PA, Elmore JG. Cost of breast-related care in the year following false positive screening mammograms. *Med Care*. 2010;48:815–820.
5 Bond M, Pavey T, Welch K, et al. Psychological consequences of false-positive screening mammograms in the UK. *BMJ Evid Based Med*. 2013;18:54–61.
6 Klompenhouwer EG, Duijm LEM, Voogd AC, et al. Re-attendance at biennial screening mammography following a repeated false positive recall. *Breast Cancer Res Treat*. 2014;145:429–437.
7 Zanardo M, Cozzi A, Trimboli RM, et al. Technique, protocols and adverse reactions for contrast-enhanced spectral mammography (CESM): a systematic review. *Insights Imaging*. 2019;10:76.
8 Xiang W, Rao H, Zhou L. A meta-analysis of contrast-enhanced spectral mammography versus MRI in the diagnosis of breast cancer. Thorac. *Cancer*. 2020;11:1423–1432.
9 Zhu X, Huang JM, Zhang K, et al. Diagnostic value of contrast-enhanced spectral mammography for screening breast cancer: systematic review and meta-analysis. *Clin Breast Cancer*. 2018;18:e985–e995.
10 Suter MB, Pesapane F, Agazzi GM, et al. Diagnostic accuracy of contrast-enhanced spectral mammography for breast lesions: a systematic review and meta-analysis. *Breast*. 2020;53:8–17.
11 Tagliafico AS, Bignotti B, Rossi F, et al. Diagnostic performance of contrast-enhanced spectral mammography: systematic review and meta-analysis. *Breast*. 2016;28:13–19.
12 Lobbes MBI, Lalji U, Houwers J, et al. Contrast-enhanced spectral mammography in patients referred from the breast cancer screening programme. *Eur Radiol*. 2014;24:1668–1676.
13 Lalji UC, Houben IPL, Prevos R, et al. Contrast-enhanced spectral mammography in recalls from the Dutch breast cancer screening program: validation of results in a large multireader, multicase study. *Eur Radiol*. 2016;26:4371–4379.
14 Houben IPL, Van de Voorde P, Jeukens CRLPN, et al. Contrast-enhanced spectral mammography as work-up tool in patients recalled from breast cancer screening has low risks and might hold clinical benefits. *Eur J Radiol*. 2017;94:31–37.
15 Neeter LMFH, Houben IPL, Nelemans PJ, et al. Rapid access to contrast-enhanced spectral mammography in women recalled from breast cancer screening: the RACER trial study design. *Trials*. 2019;20:759.
16 Jochelson MS, Lobbes MBI. Contrast-enhanced mammography: state of the art. *Radiology*. 2021;299:36–48.
17 Lee CH, Phillips J, Sung JS, Lewin JM, Newell MS. *Contrast enhanced mammography (CEM): a supplement to ACR BI-RADS® Mammography*; 2013:1–64. Webpage https://www.acr.org/-/media/ACR/Files/RADS/BI-RADS/BIRADS_CEM_2022.pdf. Accessed February 13, 2024.
18 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–845.
19 Sorin V, Yagil Y, Yosepovich A, et al. Contrast-enhanced spectral mammography in women with intermediate breast cancer risk and dense breasts. *AJR Am J Roentgenol*. 2018;211:W267–W274.
20 Pötsch N, Vatteroni G, Clauser P, Helbich TH, Baltzer PAT. Contrast-enhanced mammography versus contrast- enhanced breast MRI: a systematic review and meta-analysis. *Radiology*. 2022;305:94–103.
21 Lobbes MBI, Jochelson MS, Neeter LMFH, Nelemans PJ. Contrast-enhanced mammography and breast MRI: friends or foes? *Radiology*. 2022;307:221558.
22 Van der Steeg AFW, Keyzer-Dekker CMG, De Vries J, Roukema JA. Effect of abnormal screening mammogram on quality of life. *Br J Surg*. 2011;98:537–542.
23 Thomas M, Kelly ED, Abraham J, Kruse M. Invasive lobular breast cancer: a review of pathogenesis, diagnosis, management, and future directions of early stage disease. *Semin Oncol*. 2019;46:121–132.
24 Jeukens CRLPN, Lalji UC, Meijer E, et al. Radiation exposure of contrast-enhanced spectral mammography compared with full-field digital mammography. *Invest Radiol*. 2014;49:659–665.
25 Yaffe MJ, Mainprize JG. Risk of radiation-induced breast cancer from mammographic screening. *Radiology*. 2011;258:98–105.
26 American Association of Physicists in Medicine. *Position statement on radiation risks from medical imaging procedures*. American Association of Physicists in Medicine; 2018:330–331.
27 Klang E, Krosser A, Amitai MM, et al. Utility of routine use of breast ultrasound following contrast-enhanced spectral mammography. *Clin Radiol*. 2018;73:908.e11–908.e16.
28 Cheung YC, Lin YC, Wan YL, et al. Diagnostic performance of dual-energy contrast-enhanced subtracted mammography in dense breasts compared to mammography alone: interobserver blind-reading analysis. *Eur Radiol*. 2014;24:2394–2403.
29 Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. *European guidelines for quality assurance in breast cancer screening and diagnosis*. European Communities; 2006:1–30.
30 D'Orsi CJ, Sickles E, Mendelson E, Morris E. *AtlasBI-RADS BBreast imaging reporting and data system*. 2013.