Recall and Outcome of Screen-detected Microcalcifications during 2 Decades of Mammography Screening in the Netherlands National Breast Screening Program

Radiology

Recall and Outcome of Screen-detected Microcalcifications during 2 Decades of Mammography Screening in the Netherlands National Breast Screening Program

Jacky D. Luiten, MD • Adri C. Voogd, PhD • Ernest J. T. Luiten, MD, PhD • Mireille J. M. Broeders, MSc • Kit C. B. Roes, PhD • Vivianne C. G. Tjan-Heijnen, MD, PhD • Lucien E. M. Duijm, MD, PhD

From the Department of Surgery, Elisabeth-Tweesteden Hospital, Hilvarenbeekseweg 60, 5022 GC Tilburg, the Netherlands (J.D.L.); School for Oncology and Developmental Biology, GROW Research Institute (J.D.L., V.C.G.T.), and Department of Epidemiology (A.C.V.), Maastricht University, Maastricht, the Netherlands; Department of Research and Development, Netherlands Comprehensive Cancer Organization, Utrecht, the Netherlands (A.C.V.); Department of Breast Surgery, Amphia Hospital, Breda, the Netherlands (E.J.T.L.); Department of Health Evidence, Biostatistics Section, Radboud University Medical Centre, Nijmegen, the Netherlands (M.J.M.B., K.C.B.R.); Dutch Expert Centre for Screening, Nijmegen, the Netherlands (M.J.M.B., L.E.M.D.); Department of Internal Medicine, Division of Medical Oncology, GROW, Maastricht University Medical Centre, Maastricht, the Netherlands (V.C.G.T.); and Department of Radiology, Canisius Wilhelmina Hospital, Nijmegen, the Netherlands (L.E.M.D.). Received June 4, 2019; revision requested July 16; revision received October 28; accepted November 14. **Address correspondence to** J.D.L. (e-mail: *jd.luiten@etz.nl*).

Conflicts of interest are listed at the end of this article.

Radiology 2020; 294:528–537 • https://doi.org/10.1148/radiol.2020191266 • Content code: BR

Background: Trends in the detection of suspicious microcalcifications at mammography screening and the yield of these lesions after recall are unknown.

Purpose: To determine trends in recall and outcome of screen-detected microcalcifications during 20 years of mammography screening.

Materials and Methods: The authors performed a retrospective analysis of a consecutive series of 817 656 screening examinations (January 1997 to January 2017) in a national breast screening program. In 2009–2010 (transition period), screen-film mammography (SFM) was gradually replaced by full-field digital mammography (FFDM). The recalls of suspicious microcalcifications from all radiology reports and pathologic outcome of recalled women with 2-year follow-up were analyzed. Screening outcome in the era of SFM (1997–2008), the transition period (2009–2010), and the era of FFDM (2011–2016) were compared. Trends over time and variations between the SFM and FFDM periods were expressed by using proportions with 95% confidence intervals (CIs). In cases where the analysis based on the CI confirmed clear periods (eg, before and after introduction of FFDM), pre- and postchange outcomes were compared by using χ^2 tests.

Results: A total of 18 592 women (median age, 59 years; interquartile range, 14 years) were recalled at mammography screening, 3556 of whom had suspicious microcalcifications. The recall rate for microcalcifications increased from 0.1% in 1997–1998 to 0.5% in 2015–2016 ($P < .001$). This was temporally associated with the change from SFM to FFDM. The recalls yielding ductal carcinoma in situ (DCIS) increased from 0.3 per 1000 screening examinations with SFM to 1.1 per 1000 screening examinations with FFDM ($P < .001$), resulting in a decrease in the positive predictive value for recall for suspicious microcalcifications from 51% to 33% ($P < .001$). More than half of all DCIS lesions were high grade (52.6%; 393 of 747). The distribution of DCIS grades was stable during the 20-year screening period ($P = .36$).

Conclusion: The recall rate for suspicious microcalcifications at mammographic screening increased during the past 2 decades, whereas the ductal carcinoma in situ detection rate increased less rapidly, resulting in a lower positive predictive value for recall.

Fine calcifications of the breast, so-called microcalcifications, were first described by the German surgeon Salomon in 1913 (1). They are defined as tiny grouped calcareous deposits in terminal ductal lobular units of the breast and originate from intraluminal secretions or necrosis of epithelial cells. Microcalcifications visible at mammography may be an early sign—and sometimes even the only sign—of underlying early breast cancer or ductal carcinoma in situ (DCIS) (2–4). Up to 90% of DCIS is not palpable and is diagnosed at the work-up of suspicious microcalcifications seen at mammography (5,6).

Following the introduction of a nationwide biennial mammography screening program in the Netherlands in 1989, a considerable increase in the incidence of DCIS has been reported (7,8).

Moreover, the replacement of screen-film mammography (SFM) by full-field digital mammography (FFDM) in the Dutch mammography screening program between 2009 and 2010 (transition period) further contributed to this increase (7–9).

Since 1995, data on all screened women in the southern region of the Netherlands have been recorded in a database. This database is used for quality assurance and improvement of the screening program. We have recently reported on the overall trends in incidence and tumor grade of screening-detected DCIS and invasive carcinoma and the use of surgical excision biopsies during 2 decades of mammography screening (9,10).

Most DCIS lesions are detected at mammography screening by the presence of microcalcifications. However,

Abbreviations

CDR = cancer detection rate, DCIS = ductal carcinoma in situ, FFDM = full-field digital mammography, PPV = positive predictive value, SCNB = stereotactic core needle biopsy, SFM = screen-film mammography

Summary

Over 2 decades in the Netherlands National Breast Screening Program, the recall rate for suspicious microcalcifications increased fivefold while the positive predictive value of screening mammography decreased by 40% (from 42% to 25%). These changes were temporally associated with the introduction of digital mammography.

Key Results

- Screening mammograms from the past 2 decades in the southern region of the Netherlands National Breast Screening Program showed a fivefold increase in the recall rate for suspicious microcalcifications, from 0.1% (1997–1998) to 0.5% (2015–2016).
- Recalls yielding ductal carcinoma in situ (DCIS) increased from 0.3 (17 of 48, $n = 721$) to 1.1 (147 of 131, $n = 757$) per 1000 screening examinations ($P < .001$), resulting in a decrease in the positive predictive value for recall for suspicious microcalcifications from 51% to 33% ($P < .001$).
- More than half of the DCIS lesions were high grade (53%; 393 of 747), and the distribution in DCIS grade was stable during the 20-year screening period ($P = .36$).

not all microcalcifications found at screening are related to underlying DCIS. Little is known about the trends in the detection of microcalcifications at mammography screening and the yield of these findings after recall. Digital mammography has a higher sensitivity than SFM for the detection of microcalcifications, resulting in increased recall rates for calcifications during the transition from SFM to FFDM screening (11). Furthermore, once microcalcification-associated DCIS is diagnosed, the potential for overdiagnosis and overtreatment, especially with low-grade DCIS, is a serious concern.

Therefore, the purpose of this study was to determine the trends in recall of suspicious microcalcifications and their outcomes during 2 decades of mammography screening. We analyzed the yield of microcalcifications, expressed as the cancer detection rate (CDR; the number of [pre]malignant abnormalities per 1000 screening examinations) and the histologic grade at current digital mammography screening compared with that in the era of SFM. More specifically, we determined to which degree the changes in recall of microcalcifications and the histologic grade of DCIS, observed during the transition from SFM screening to digital screening, persisted several years after this transition.

Materials and Methods

Mammography Screening Program and Study Cohort

A biennial nationwide mammography screening program was introduced in the Netherlands between 1989 and 1996 for all women aged 50–70 years. Between 1998 and 1999, the upper age limit was extended to 75 years. In the southern part of the Netherlands, SFM was gradually replaced by FFDM between May 2009 and April 2010. The Dutch screening program has been described in more detail previously (7,12–14). In summary, all screening mammograms are obtained by certified radiographers, after which the images are routinely double-read by certified screening radiologists. The radiologists classify the mammographic abnormality in case of recall (suspicious mass, suspicious microcalcifications, suspicious mass in combination with microcalcifications, asymmetry, architectural distortion, or other suspicious abnormality). A Breast Imaging Reporting and Data System category is routinely provided in clinical breast imaging reports after 2001 (15,16). Recalled women are then referred by their general practitioner to a hospital breast unit for further analysis. In case of a false-positive recall, a woman returns to the screening program and is screened with the same frequency (biennial) as all other women who attend the program.

We retrospectively analyzed all women who attended the breast cancer screening program at four specialized breast cancer screening units in a southern part of the Netherlands between January 1997 and January 2017. Our study population consisted of a consecutive series of 817 656 screening examinations. Women participating in the screening program were offered the option to opt out of the use of their data for quality assessment and scientific purposes. Three recalled women refrained from giving permission and were excluded from analysis. According to the Dutch Central Committee on Research involving Human Subjects, ethical approval was not required for our study.

The authors declare no conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Follow-up of Recalled Women

During a 2-year follow-up period (until the next biennial screening examination), we collected the radiology reports, type of biopsy methods with their outcome, and breast surgery reports of all recalled women. The screening organization routinely received the follow-up data from the hospitals at which the women were analyzed after recall. To complete 2-year follow-up, one of the radiologists (L.E.M.D., with >25 years of experience in breast imaging) and several radiology residents collected additional reports, which were not received by the screening organization, through visits at these departments. All data were then entered into a database, which was created for quality control of the screening program and scientific purposes by the radiologist. The quality of data entry was not reviewed.

If a woman was recalled for more than one ipsilateral lesion or for one lesion in the right breast and one in the left breast (bilateral) during the same screening round, the mammographic lesion with the highest suspicion at mammography screening was considered as the index lesion for recall. For the purpose of this study, we scored one screening abnormality per recalled woman. If a woman was recalled again in a subsequent screening round, this counted as a new recall. A total of 60 women were recalled twice, and two were recalled three times. Only four women, one of whom had microcalcifications, experienced a repeated recall within the same 2-year period.

Screening-detected cancers were divided into DCIS and invasive cancers. Lobular carcinoma in situ is considered a benign lesion. Details on the methods for the detection of interval cancers in our screened cohort have been published previously (17,18).

Statistical Analysis

The main outcome measure of this study was the number of microcalcification recalls per 1000 screening examinations and positive predictive value (PPV) of microcalcification recalls during 2 decades of mammography screening. These trends are shown as graphs and reported as absolute numbers, proportions, PPVs with 95% confidence intervals (CIs), and rates per 1000 screening examinations for women screened from 1997 until 2017. Evaluation of the CIs over time, treating nonoverlapping CIs as evidence of difference, then provides conservative assessment of trends (19) that fit the graphical presentation and does not go beyond the limitations in the data. To allow clear interpretation of the data, results are presented separately for first (initial) screening examinations of participating women and subsequent screening examinations. The limitations of the data did not allow further refinement, and thus correlations due to the same women being screened more than once could not be modeled. In case the analysis based on CIs enabled clear differentiation of separate periods, such as before (through 2007 and 2008) and after (from 2011 to 2012) introduction of FFDM, pre- and post-change outcomes were compared by using χ^2 tests in addition to summary statistics. Interpretation of the tests is exploratory, rather than confirmatory. Preoperative DCIS confirmation was evaluated descriptively only.

The median differences and interquartile ranges were calculated for continuous variables. Statistical analyses were performed by using commercially available software (SPSS, version 22.0; SPSS, Chicago, Ill).

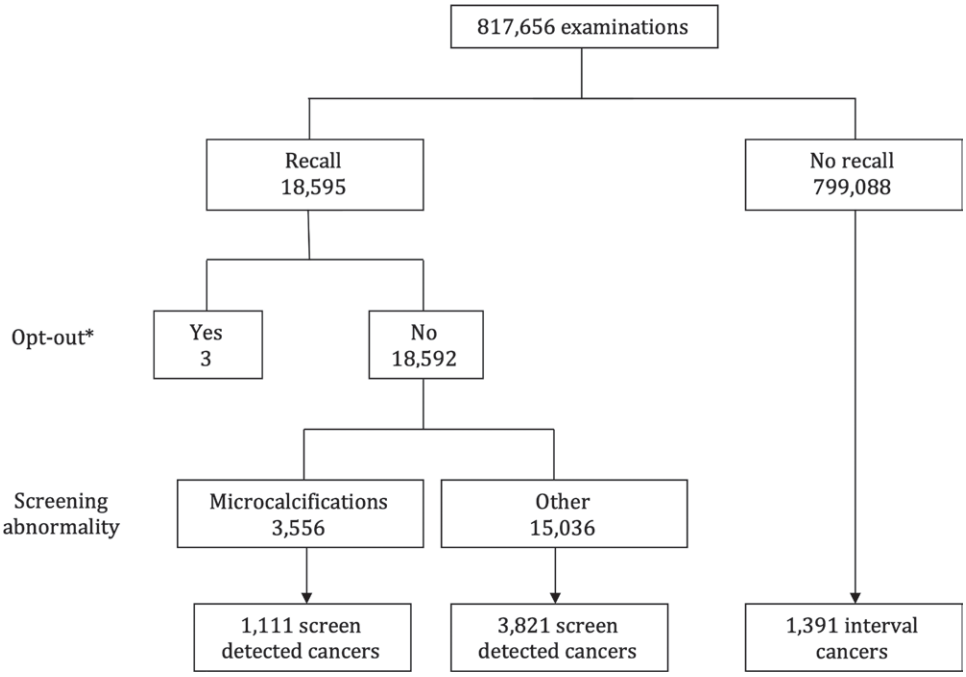

Figure 1: Flowchart of study population. * = The screening program requires women to opt out of mammography screening and/or use of their data.

Results

Overall Screening Results

A total of 817 656 mammography screening examinations were performed between January 1997 and January 2017. Of those 817 656 examinations, 97 541 (11.9%) were initial screening examinations and 720 115 (88.1%) were subsequent screening examinations. The median age of recalled women was 59 years (interquartile range, 14). Figure 1 shows a flowchart of the study participants.

The number of screening examinations increased from 48 721 in 1997–1998 to 131 757 in 2015–2016 (Table 1). Of the 817 656 women, 18 592 (2.3%) were recalled for further analysis of a suspicious lesion at mammography screening; the overall recall proportion was 4.7% (95% CI: 4.6%, 4.9%; 4612 of 97 541) at initial screening examination and 1.9% (95% CI: 1.9%, 1.9%; 13 980 of 720 115) at subsequent screening examinations.

From the trend analysis (Table 1), the recall rate for initial screening examinations increased from 2.3% (1039 of 46 155) during the SFM period (1997–1998) to 5.5% (558 of 10 182) during the transition period ($P < .001$) and to 7.3% (3015 of 41 204) during the FFDM period (2011–2016) ($P < .001$). The recall rate for subsequent screening examinations increased in the same period, from 1.1% (3413 of 304 854) in the SFM period to 2.3% (1801 of 79 946; $P < .001$) in the transition period and to 2.6% (8766 of 335 315) in the FFDM period ($P < .001$) (Table 1). Combined, these led to an increase of the overall recall rate from 1.3% (4452 of 351 009) during the SFM period to 3.1% (11 781 of 376 519) during the FFDM period ($P < .001$).

Trends in Recall of Suspicious Microcalcifications

Of the 817 656 women who underwent screening examinations, 3556 (0.4%) were recalled because of microcalcifications. The absolute number of women recalled with microcalcification abnormalities increased from 51 in 1997–1998 (0.1% of 48 721 screening examinations) to 680 in 2015–2016 (0.5% of 131 757 screening examinations) (Table 2). The trend in proportion of recalls for microcalcifications was very similar to the trend for recalls overall, with an increase from 0.2% (95% CI: 0.1%, 0.3%; 17 of 9602) for initial screening examinations in 1997–1998 to 1.3% (95% CI: 1.1%, 1.4%; 179 of 14 220) in 2015–2016 (Table 2) (SFM period vs FFDM period, $P < .001$). Subsequent screening examinations showed a similar pattern but with substantially smaller proportions.

Table 1: Overall Screening Results according to Year

Parameter	1997–1998	1999–2000	2001–2002	2003–2004	2005–2006	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016	Total
No. of screening examinations	48 721	53 718	53 489	61 251	66 300	67 530	90 128	113 335	131 427	131 757	817 656
Initial examinations	9602	8968	6902	7038	6588	7057	10 182	12 969	14 015	14 220	97 541
Subsequent examinations	39 119	44 750	46 587	54 213	59 712	60 473	79 946	100 366	117 412	117 537	720 115
No. of recalls*	536 (1.1)	499 (0.9)	555 (1.0)	985 (1.6)	874 (1.3)	1003 (1.5)	2359 (2.6)	3507 (3.1)	4653 (3.5)	3621 (2.7)	18 592 (2.3)
Initial examination†	163 (1.7) [1.4, 2.0]	128 (1.4) [1.2, 1.7]	113 (1.6) [1.3, 1.9]	211 (3.0) [2.6, 3.4]	198 (3.0) [2.6, 3.4]	226 (3.2) [2.8, 3.6]	558 (5.5) [5.0, 5.9]	865 (6.7) [6.2, 7.1]	1133 (8.1) [7.6, 8.5]	1017 (7.2) [6.7, 7.6]	4612 (4.7) [4.6, 4.9]
Subsequent examinations‡	373 (0.95) [0.9, 1.0]	371 (0.8) [0.7, 0.9]	442 (0.95) [0.9, 1.0]	774 (1.4) [1.3, 1.5]	676 (1.1) [1.0, 1.2]	777 (1.3) [1.2, 1.4]	1801 (2.3) [2.1, 2.4]	2642 (2.6 [2.5, 2.7]	3520 (3.0) [2.9, 3.1]	2604 (2.2) [2.1, 2.3]	13 980 (1.9) [1.9, 1.9]
No. of cancers	224	275	254	345	322	354	571	775	925	887	4932
CDR§	4.6	5.1	4.7	5.6	4.9	5.2	6.3	6.8	7.0	6.7	6.0
Initial examinations	5.7	7.5	5.1	5.5	8.3	6.5	7.1	6.9	8.1	8.4	7.1
Subsequent examinations	4.3	4.6	4.7	5.6	4.5	5.1	6.2	6.8	6.9	6.5	5.9
PPV for recall (%)‖	41.8	55.1	45.8	35.0	36.8	35.3	24.2	22.1	19.9	24.5	26.5
Initial examinations	33.7 (26.5, 41.0)	52.3 (43.7, 61.0)	31.0 (22.4, 39.5)	18.5 (13.2, 23.7)	27.8 (21.5, 34.0)	20.4 (15.1, 25.6)	12.9 (10.1, 15.7)	10.4 (8.4, 12.4)	10.1 (8.3, 11.8)	11.7 (9.7, 13.7)	15.0 (14.0, 16.0)
Subsequent examinations	45.3 (40.3, 50.4)	56.1 (51.0, 61.0)	49.5 (44,9, 54.2)	39.5 (36.1, 43.0)	39.5 (35.8, 43.2)	39.6 (36.2, 43.1)	27.7 (25.6, 29.8)	25.9 (24.3, 27.6)	23.0 (21.6, 24.4)	29.5 (27.7, 31.2)	30.3 (29.6, 31.1)

Note.—CDR = cancer detection rate, PPV = positive predictive value.

* Numbers in parentheses are percentages, and numbers in brackets are the 95% confidence interval.

† Because these initial screenings involve all different women, the intervals are correct (95% coverage).

‡ Because of dependence between repeat screenings, these intervals are not entirely correct but are likely conservative (ie, actual variance is smaller). Because data could not be retrieved per woman per repeat screening examination (anonymous database), full correlations cannot be modeled and estimated·

§ Per 1000 screening examinations.

‖ Numbers in parentheses are 95% confidence intervals.

Outcome after Recall

Breast cancer was diagnosed in 4932 recalled women (screening-detected cancers) and 1391 nonrecalled women (interval cancers), yielding an overall CDR of 6.0 per 1000 screening examinations (4932 of 817 656) (Table 1) and a program sensitivity of 78.0% (4932 of 6323). If the proportions reported earlier are translated into CDR, the rates for initial screening examinations increased from 6.4 per 1000 screening examinations (297 of 46 155) in the SFM period to 7.8 per 1000 screening examinations (323 of 41 204) in the FFDM period ($P = .01$) and the CDR for subsequent screening examinations increased from 4.8 per 1000 screening examinations (1477 of 304 854) to 6.8 per 1000 screening examinations (2264 of 335 315), respectively ($P < .001$).

Similarly, although the recall rate of suspicious microcalcifications increased from 2.1 per 1000 screening examinations (741 of 351 009) in the SFM period to 5.7 per 1000 screening examinations (2162 of 376 519) in the FFDM period ($P < .001$), the DCIS detection rate increased from 0.9 per 1000 screening examinations (321 of 351 009) to 1.7 per 1000 screening examinations (637 of 376 519; $P < .001$), respectively (Fig 2). In 2009–2010, an increased recall rate for suspicious microcalcifications to 7.2 per 1000 screening examinations (653 of 90 128; $P < .001$ vs the SFM period) was observed, as well as an increase in the DCIS detection rate to 1.7 per 1000 screening examinations (153 of 90 128; $P < .001$ vs the SFM period). These changes were temporally associated with the introduction of digital mammography. To more closely assess the change in the DCIS detection rate given the apparent shift after the introduction of FFDM, a χ^2 test comparing detection rates during the SFM period and FFDM period showed that detection rates significantly increased after

Table 2: Trends in Recall and Work-up of Suspicious Calcifications at Screening Mammography according to Year

Parameter	1997–1998	1999–2000	2001–2002	2003–2004	2005–2006	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016	Total
No. of screening examinations	48 721	53 718	53 489	61 251	66 300	67 530	90 128	113 335	131 427	131 757	817 656
No. of calcification-related recalls	51 (0.1)	68 (0.1)	77 (0.1)	213 (0.3)	165 (0.2)	167 (0.2)	653 (0.7)	761 (0.7)	721 (0.5)	680 (0.5)	3556 (0.4)
Initial examinations*	17 (0.2) [0.1, 0.3]	22 (0.2) [0.1, 0.3]	15 (0.2) [0.1, 0.3]	43 ((0.6) [0.4, 0.8]	32 (0.5) [0.3, 0.7]	37 (0.5) [0.4, 0.7]	148 (1.5) [1.2, 1.7]	209 (1.6) [1.4, 1.8]	224 (1.6) [1.4, 1.8]	179 (1.3) [1.1, 1.4]	926 (0.9) [0.9, 1.0]
Subsequent examinations†	34 (0.1) [0.1, 0.1]	46 (0.1) [0.1, 0.1]	62 (0.1) [0.1, 0.2]	170 (0.3) [0.3, 0.4]	133 (0.2) [0.2, 0.3]	130 (0.2) [0.2, 0.3]	505 (0.6) [0.6, 0.7]	552 (0.5) [0.5, 0.6]	497 (0.4) [0.4, 0.5]	501 (0.4) [0.4, 0.5]	2630 (0.4) [0.4, 0.4]
Preoperative assessment of recalled calcifications											
None	0	0	1 (1.3)	0	1 (0.6)	1 (0.6)	2 (0.3)	2 (0.3)	4 (0.6)	0	11 (0.3)
Additional imaging‡	32 (62.7)	48 (70.6)	21 (27.3)	39 (18.3)	29 (17.6)	26 (15.6)	117 (17.9)	151 (19.8)	85 (11.8)	69 (10.1)	617 (17.4)
Additional imaging and biopsy											
FNAC	2 (3.9)	0	4 (5.2)	8 (3.8)	4 (2.4)	2 (1.2)	5 (0.8)	2 (0.3)	1 (0.1)	0	28 (0.8)
CNB	0	4 (5.9)	3 (3.9)	16 (7.5)	9 (5.5)	16 (9.6)	47 (7.2)	45 (5.9)	56 (7.8)	57 (8.4)	253 (7.1)
SCNB	0	5 (7.4)	26 (33.8)	126 (59.2)	117 (70.9)	113 (67.7)	474 (72.6)	542 (71.2)	553 (76.7)	528 (77.6)	2484 (69.8)
Surgical biopsy	17 (33.3)	11 (16.2)	22 (28.6)	24 (11.3)	5 (3.0)	9 (5.4)	8 (1.2)	19 (2.5)	22 (3.1)	26 (3.8)	163 (4.6)
BI-RADS§	NA	NA									
1	…	…	0	1 (0.5)	1 (0.6)	3 (1.8)	5 (0.8)	8 (1.1)	0	2 (0.3)	…
2	…	…	7 (9.1)	18 (8.5)	16 (9.7)	15 (9.0)	74 (11.3)	92 (12.1)	54 (7.5)	38 (5.6)	…
3	…	…	9 (11.7)	46 (21.6)	40 (24.2)	42 (25.1)	141 (21.6)	144 (18.9)	111 (15.4)	53 (7.8)	…
4	…	…	52 (67.5)	143 (67.1)	97 (58.8)	94 (56.3)	413 (63.2)	506 (66.5)	538 (74.6)	562 (82.6)	…
5	…	…	9 (11.7)	5 (2.3)	9 (5.5)	11 (6.6)	15 (2.3)	8 (1.1)	13 (1.8)	20 (2.9)	…
Unknown	…	…	0	0	2 (1.2)	2 (1.2)	5 (0.8)	3 (0.4)	5 (0.7)	5 (0.7)	…

Note.—Numbers in parentheses are percentages, and numbers in brackets are 95% confidence intervals. BI-RADS = Breast Imaging Reporting and Data System, CNB =core needle biopsy, FNAC = fine needle aspiration cytology, NA = not available, SCNB = stereotactic core needle biopsy.

* Because these initial screenings involved all different women, the intervals are correct (95% coverage).

† Because of dependence between repeat screenings, these intervals are not entirely correct but are likely conservative (ie, actual variance is smaller). As data could not be retrieved per woman per repeat screening examination (anonymous database), full correlations cannot be modeled and estimated.

‡ Additional imaging included additional mammographic views, breast tomosynthesis, three-dimensional breast US, breast MRI, or a combination of these modalities.

§ BI-RADS classification was routinely available since 2001.

the introduction of FFDM ($P < .001$ for both initial and subsequent screening examinations).

The PPV for overall recall was 26.5% (4932 of 18 592). Table 1 shows the PPVs and 95% CIs for all recalls and for initial and subsequent screening examinations. The PPV for all recalls decreased from 41.8% (224 of 536) in 1997–1998 to 24.5% (887 of 3621) in 2015–2016. In the comparison of the SFM and the FFDM periods, the PPV for all recalls decreased from 39.8% (1774 of 4452) to 22.0% (2587 of 11 781) ($P < .001$).

As shown in Figure 3, the PPV for microcalcification recalls showed a similar decreasing trend over the years, from 51.0% (26 of 51) in 1997–1998 to 32.6% (222 of 680) in 2015–2016 (43.3% [321 of 741] during the SFM period vs 29.5% [637 of 2162] during the FFDM period; $P < .001$) (Table 3).

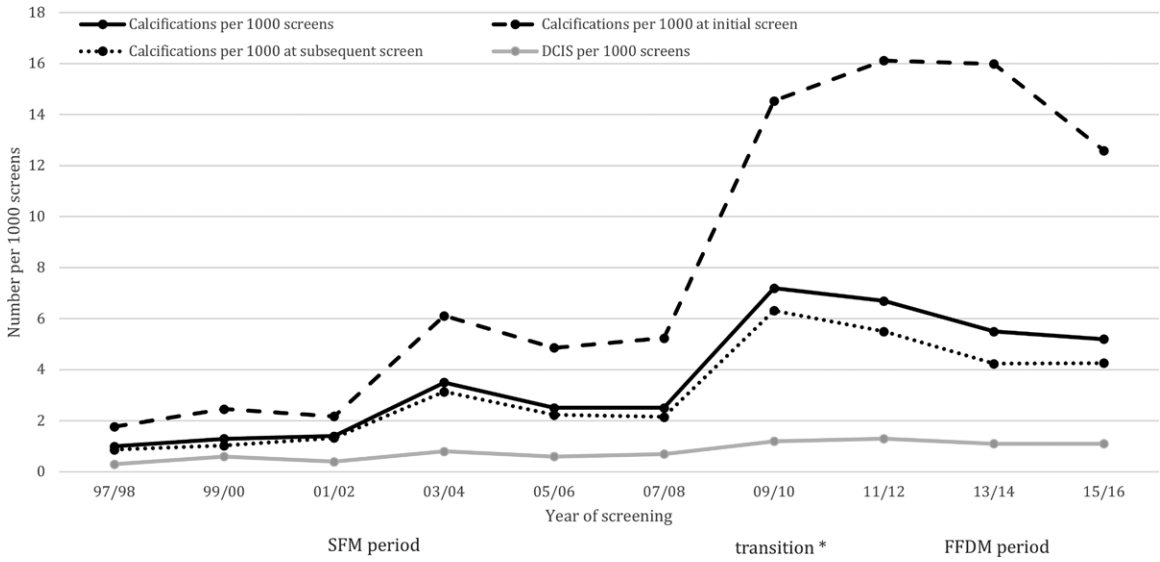

Figure 2: Graph shows trends in recall of calcifications and ductal carcinoma in situ (DCIS) yield per 1000 screening examinations. * = Transition from screen-film mammography (SFM) to full-field digital mammography (FFDM) took place between 2009 and 2010.

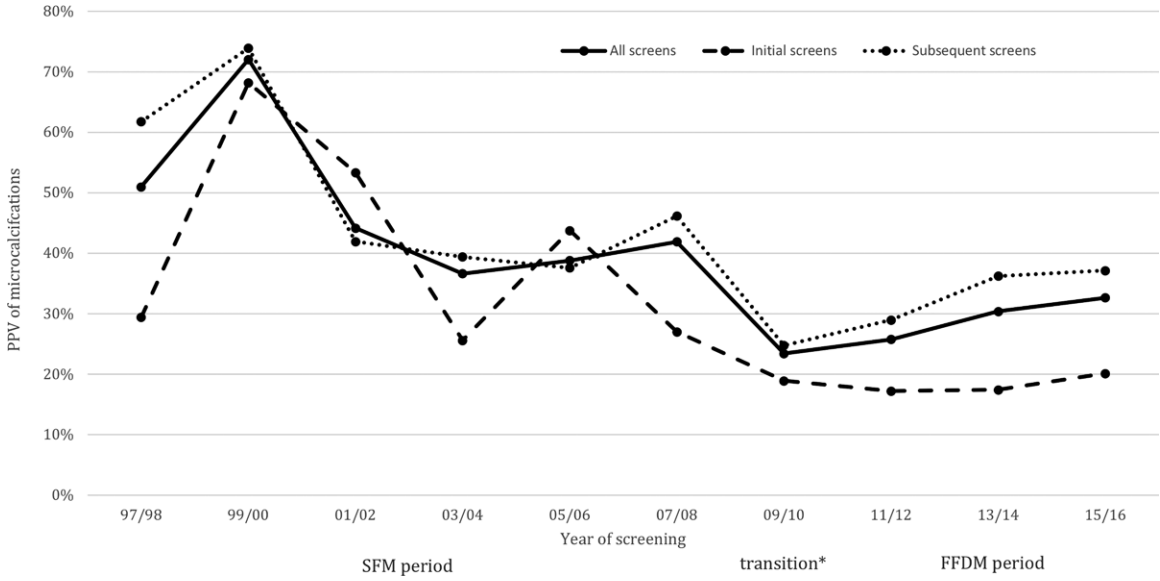

Figure 3: Graph shows trends in positive predictive value (PPV) of microcalcification recalls. * = Transition from screen-film mammography (SFM) to full-field digital mammography (FFDM) took place between 2009 and 2010.

Trends in Diagnostic Work-up of Microcalcifications

Of the 3556 women recalled for suspicious microcalcifications, 617 (17.4%) underwent additional imaging (additional mammographic views, tomosynthesis, breast US, breast MRI, or any combination of these modalities) as the only diagnostic procedure (Table 2).

Stereotactic core needle biopsy (SCNB) was the most frequently used type of biopsy, and a total of 2484 SCNB procedures were performed (Table 2). The SCNB rate per 1000 screening examinations increased from 0.4 (150 of 351 009) during the SFM period to 1.2 (109 of 90 128) during the transition period ($P < .001$), remaining stable in the FFDM period (1.2; 451 of 176 519). The proportion of women who underwent SCNB for microcalcifications increased from 0% (0 of 51) in 1997–1998 to 77.6% (528 of 680) in 2015–2016 ($P < .001$).

A total of 1111 suspicious microcalcifications proved to be (pre)malignant (31.2%; 1111 of 3556) and consisted of 747 DCIS (67.2%) and 364 invasive cancers (32.8%) (Table 3). The preoperative confirmation of DCIS with percutaneous biopsy increased over the years, from 5.9% (one of 17) in 1997–1998 to 91.8% (135 of 147) in 2015–2016 ($P < .001$) (Table 3).

Trends in Surgical Treatment of DCIS

Of all 747 women with DCIS lesions that presented with suspicious calcifications, 561 (75.1%) were treated with breast-conserving surgery and 173 (23.2%) with mastectomy. The remaining 13 women (1.7%) did not undergo surgical treatment (Table 3). The lowest proportion of breast-conserving surgery for screening-detected DCIS was observed in women in the first cohort of 1997–1998 (59%; 10 of 17) and subse-

Table 3: Suspicious Calcifications at Screening Mammography: Trends in PPV of Recall, DCIS Detection Rate, Grading, and Surgical Treatment

Parameter	1997–1998	1999–2000	2001–2002	2003–2004	2005–2006	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016	Total
No. of cancers	26 (51.0)	49 (72.1)	34 (44.2)	78 (36.6)	64 (38.8)	70 (41.9)	153 (23.4)	196 (25.8)	219 (30.4)	222 (32.6)	1111 (31.2)
Type of cancer											
DCIS	17 (65.4)	33 (67.3)	22 (64.7)	47 (60.3)	40 (62.5)	44 (62.9)	104 (68.0)	145 (74.0)	148 (67.6)	147 (66.2)	747 (67.2)
Invasive cancer	9 (34.6)	16 (32.7)	12 (35.3)	31 (39.7)	24 (37.5)	26 (37.1)	49 (32.0)	51 (26.0)	71 (32.4)	75 (33.8)	364 (32.8)
Mean DCIS (mm)*	18 (4–60)	16 (2–35)	21 (4–70)	20 (3–40)	25 (2–80)	21 (4–85)	15 (1–90)	15 (1–130)	22 (2–114)	21 (2–100)	19 (1–130)
PPV for recall (%)†	51.0	72.1	44.2	36.6	38.8	41.9	23.4	25.8	30.4	32.6	31.2
Initial examinations	29.4 (7.8, 51.1)	68.2 (48.7, 87.6)	53.3 (28.1, 78.6)	25.6 (12.5, 38.6)	43.8 (26.6, 60.9)	27.0 (12.7, 41.3)	18.9 (12.6, 25.2)	17.2 (12.1, 22.3)	17.4 (12.4, 22.4)	20.1 (14.2, 26.0)	21.8 (19.2, 24.5)
Subsequent examinations	61.8 (45.4, 78.1)	73.9 (61.2, 86.6)	41.9 (29.7, 54.2)	39.4 (32.1, 46.8)	37.6 (29.4, 45.8)	46.2 (37.6, 54.7)	24.8 (21.0, 28.5)	29.0 (25.2, 32.8)	36.2 (32.0, 40.4)	37.1 (32.9, 41.4)	34.6 (32.7, 36.4)
DCIS detection rate‡	0.3	0.6	0.4	0.8	0.6	0.7	1.2	1.3	1.1	1.1	0.9
Initial examinations	0.4	1.1	0.9	1.1	1.2	1.0	2.4	2.3	1.8	1.9	1.5
Subsequent examinations	0.3	0.5	0.3	0.7	0.5	0.6	1.0	1.1	1.0	1.0	0.8
DCIS grades											
Low	3 (17.6)	2 (6.1)	3 (13.6)	7 (14.9)	3 (7.5)	5 (11.4)	24 (23.1)	32 (22.1)	23 (15.5)	14 (9.5)	116 (15.5)
Intermediate	1 (5.9)	8 (24.2)	2 (9.1)	10 (21.3)	8 (20.0)	10 (22.7)	37 (35.6)	48 (33.1)	52 (35.1)	57 (38.8)	233 (31.2)
High	10 (58.8)	21 (63.6)	17 (77.3)	30 (63.8)	29 (72.5)	29 (65.9)	43 (41.3)	65 (44.8)	73 (49.3)	76 (51.7)	393 (52.6)
Unknown	3 (17.6)	2 (6.1)	0	0	0	0	0	0	0	0	5 (0.7)
Surgical treatment for DCIS											
None	0	0	0	0	1 (2.5)	1 (2.2)	2 (1.9)	4 (2.8)	2 (1.4)	3 (2.0)	13 (1.7)
BCS	10 (58.8)	24 (72.7)	16 (72.7)	38 (80.9)	28 (70.0)	31 (70.5)	80 (76.9)	110 (75.9)	114 (70.0)	110 (74.8)	561 (75.1)
Mastectomy	7 (41.2)	9 (27.3)	6 (27.3)	9 (19.1)	11 (27.5)	12 (27.3)	22 (21.2)	31 (21.4)	32 (21.6)	34 (23.1)	173 (23.2)
Preoperative DCIS confirmation	1 (5.9)	3 (9.1)	11 (50.0)	40 (85.1)	36 (90.0)	43 (97.7)	102 (98.1)	135 (93.1)	139 (93.9)	135 (91.8)	645 (86.3)

Note.—Except where indicated, numbers in parentheses are percentages. BCS = breast-conserving surgery, DCIS = ductal carcinoma in situ, PPV = positive predictive value.

* Numbers in parentheses are the range.

† Numbers in parentheses are the 95% confidence interval.

‡ Per 1000 screening examinations.

quently varied between 70% (28 of 40) and 81% (38 of 47) afterward.

DCIS Characteristics

The median DCIS size was 19 mm (range, 1–130 mm). Of the 747 DCIS lesions, 393 (52.6%) were high grade, 233 (31.2%) were intermediate grade, and 116 (15.5%) were low grade (Table 3). This distribution of low-, intermediate-, and high-grade DCIS did not change during the whole study period (SFM vs FFDM period, $P = .36$). The distributions of DCIS lesions according to histologic grade per 1000 screening examinations in the SFM and FFDM periods were 0.1 (41 of 351 009) and 0.3 (102 of 376 519), respectively, for low-grade DCIS ($P < .001$); 0.2 (58 of 351 009) and 0.5 (192 of 376 519) for intermediate-grade DCIS ($P < .001$); and 0.5 (175 of 351 009) and 0.6 (241 of 376 519; $P = .01$) for high-grade DCIS (Fig 4). Following SCNB, resection specimens showed no residual DCIS in 36 of the 734 women who underwent surgical treatment (4.9%). Of these 36 women, 10 (28%) had low-grade DCIS at SCNB (Fig 5), 14 (39%) had intermediate-grade DCIS, and 12 (33%) had high-grade DCIS.

Discussion

This retrospective 20-year analysis of screening-detected microcalcifications in the southern part of the Netherlands showed

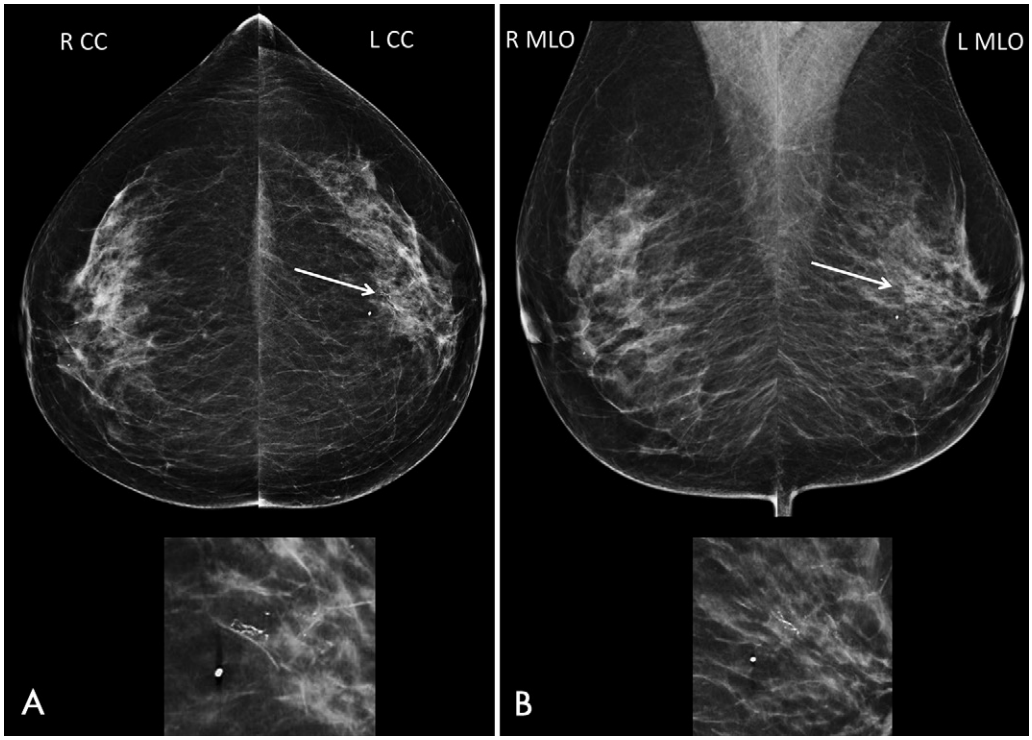

Figure 4: Two-view screening mammograms in, *A*, craniocaudal (CC) and, *B*, mediolateral oblique (MLO) views show fine linear calcifications located centrally in left breast (arrow). Calcifications are more clearly visible on spot magnification views (insets). Stereotactic vacuum-assisted core needle biopsy (9 gauge) revealed high-grade ductal carcinoma in situ. The surgical specimen was 25 mm.

a fivefold increase in the recall rate for suspicious microcalcifications from 0.1% (1997–1998, the screen-film mammography [SFM] period) to 0.5% (2015–2016, the full-field digital mammography [FFDM] period) ($P < .001$). Recalls yielding ductal carcinoma in situ (DCIS) increased from 0.3 per 1000 screening examinations (17 of 48 721) in the SFM period to 1.1 per 1000 screening examinations (147 of 131 757) in the FFDM period ($P < .001$), resulting in a decrease in the positive predictive value of recall for suspicious microcalcifications from 51% to 33% ($P < .001$). Overall, one-third of all recalled microcalcifications proved to be (pre)malignant. Almost 70% of those recalled microcalcifications were DCIS. The distribution in DCIS grade was stable during the 20-year screening period ($P = .36$), with more than half (53%) being high grade.

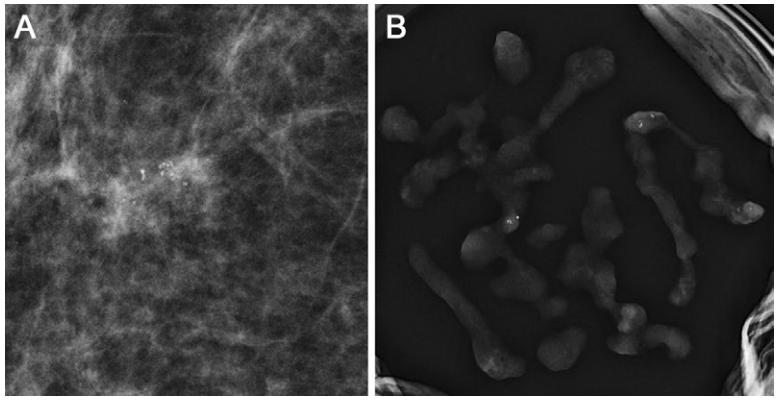

Figure 5: *A*, Spot magnification view shows grouped amorphic calcifications detected at mammography screening. Stereotactic vacuum-assisted core needle biopsy (9 gauge) yielded multiple specimens with calcifications. *B*, Photomicrograph shows usual ductal hyperplasia and adenosis.

When comparing the SFM period with the FFDM period, the increase in the number of DCIS per 1000 screening examinations was most pronounced for low-grade DCIS, which rose threefold in comparison with the lesser increase of both intermediate- and high-grade DCIS.

Almost all patients with microcalcification-associated DCIS are treated with surgery, the majority consisting of breast-conserving treatment. SCNB has replaced surgical excision biopsies and is currently the preferred technique for the diagnosis of microcalcification-associated breast abnormality (20–24). However, given the observed decrease in PPV for microcalcification recalls, one might question whether all recalled microcalcifications necessitate biopsy. Unfortunately, any underlying malignancy cannot always be ruled out with additional breast imaging, making SCNB the preferred minimally invasive technique to obtain tissue for pathologic examination.

Several studies showed that FFDM may be more effective than SFM for the detection of microcalcifications, as was also demonstrated in this study. Another explanation for the increased recall rate of microcalcifications may be the performance of routine two-view mammography (mediolateral oblique and craniocaudal views) of each breast since the start of FFDM in

2009–2010. In the SFM period, all women attending the program for the first time underwent two-view mammography, whereas subsequent screening examinations consisted of a routine mediolateral oblique view of each breast and additional craniocaudal views only if indicated (3,25,26). Depending on the characteristics of mammography screening programs, adding tomosynthesis to conventional FFDM may have a beneficial impact on recall rate and cancer detection (27). However, breast tomosynthesis has not yet been implemented in the Dutch screening program.

The increased detection of DCIS also resulted in a rise in the number of low-grade DCIS and thus probably some degree of overdiagnosis, which may pose therapeutic dilemmas for clinicians and may lead to overtreatment. As ongoing studies investigate the possibility of close surveillance of low-grade and even intermediate-grade DCIS (28,29), better discrimination of microcalcifications to prevent SCNB for suspected low-grade DCIS based on radiologic features might be a next step in de-escalating treatment. High-grade DCIS is more often associated with specific abnormal mammographic features, such as necrosis, rod and linear branch shapes, or coarse granular microcalcifications (5,30–32).

On the other hand, more than half of DCIS related to suspicious microcalcifications in our study showed high-grade histopathologic characteristics. As the detection and subsequent treatment of high-grade DCIS may reduce further development to high-grade invasive carcinoma, histologic analysis of screening-detected microcalcifications carries substantial clinical value (9). Consequently, SCNB is still considered mandatory in the work-up of these lesions to date because it is not yet clear to which degree histologic features of DCIS can be estimated by using the patterns of microcalcifications alone (30).

Surgery, including additional radiation therapy in case of breast-conserving surgery, remains the recommended choice for DCIS treatment (33). Most women with DCIS in our study were treated with breast-conserving surgery. In almost 5% of all surgically treated women, no residual DCIS was found in the surgical specimen, suggesting that all DCIS was removed at SCNB as has also been described in a recent study by Dubrovsky et al (34) in 14% of all surgically treated women. In the study by Dubrovsky et al, omission of additional radiation therapy did not alter the local recurrence rate. Vacuum-assisted excision biopsy devices can remove more tissue than SCNB and may be the future therapy for patients with small groups of clustered microcalcification-associated DCIS (35,36). A wait-and-see strategy for low-grade and intermediate-grade DCIS may be favored over surgical intervention if subsequent mammography shows no residual calcifications.

Our study has some limitations. Unfortunately, individual data at a woman level could not be retrieved for repeat screening examinations. Therefore, possible within-subject dependencies of the data in repeat screening examinations could not be modeled and estimated. Given these limitations in the data, we could only analyze all subsequent screening examinations (which may be dependent) as one group. Furthermore, we were not able to reliably retrieve the detailed radiologic features on morphologic characteristics and distribution of microcalcifications because they were not specified in our data.

Characterizing radiologic features of suspicious microcalcifications to search for patterns that correlate with more aggressive underlying disease might be a desirable next step toward a more selective use of SCNB in case of screening-detected microcalcifications. Because it is not yet clear to what extent histologic features of DCIS can be estimated by the patterns of microcalcifications alone, SCNB is still considered mandatory in the work-up of these lesions to date.

In conclusion, the recall rate of suspicious microcalcifications at mammographic screening significantly increased over the past 2 decades, while the ductal carcinoma in situ detection rate increased less rapidly at the expense of a lower positive predictive value of recall.

Author contributions: Guarantors of integrity of entire study, J.D.L., E.J.T.L., L.E.M.D.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, J.D.L., E.J.T.L., L.E.M.D.; clinical studies, E.J.T.L., L.E.M.D.; statistical analysis, J.D.L., M.J.M.B., K.C.B.R., L.E.M.D.; and manuscript editing, J.D.L., A.C.V., E.J.T.L., K.C.B.R., V.C.G.T., L.E.M.D.

Disclosures of Conflicts of Interest: J.D.L. disclosed no relevant relationships. **A.C.V.** disclosed no relevant relationships. **E.J.T.L.** disclosed no relevant relationships. **M.J.M.B.** disclosed no relevant relationships. **K.C.B.R.** disclosed no relevant relationships. **V.C.G.T.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: institution has grants/grants pending from Roche, Pfizer, Novartis, and Lilly; institution receives payment for lectures including service on speakers bureaus from Novartis; institution receives travel/accommodations/meeting expenses unrelated to activities listed from Roche, Pfizer, Novartis, and Lilly. Other relationships: disclosed no relevant relationships. **L.E.M.D.** disclosed no relevant relationships.

References

1. Knipe H, Kruger G. Albert Salomon. http://radiopaedia.org/articles/dr-albert-salomon-1. Accessed April 2019.
2. Castronovo V, Bellahcene A. Evidence that breast cancer associated microcalcifications are mineralized malignant cells. Int J Oncol 1998;12(2):305–308.
3. Del Turco MR, Mantellini P, Ciatto S, et al. Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. AJR Am J Roentgenol 2007;189(4):860–866.
4. Morgan MP, Cooke MM, McCarthy GM. Microcalcifications associated with breast cancer: an epiphenomenon or biologically significant feature of selected tumors? J Mammary Gland Biol Neoplasia 2005;10(2):181–187.
5. Rauch GM, Hobbs BP, Kuerer HM, et al. Microcalcifications in 1657 patients with pure ductal carcinoma in situ of the breast: correlation with clinical, histopathologic, biologic features, and local recurrence. Ann Surg Oncol 2016;23(2):482–489.
6. Hofvind S, Iversen BF, Eriksen L, Styr BM, Kjellevold K, Kurz KD. Mammographic morphology and distribution of calcifications in ductal carcinoma in situ diagnosed in organized screening. Acta Radiol 2011;52(5):481–487.
7. van Steenbergen LN, Voogd AC, Roukema JA, et al. Screening caused rising incidence rates of ductal carcinoma in situ of the breast. Breast Cancer Res Treat 2009;115(1):181–183.
8. Hong YK, McMasters KM, Egger ME, Ajkay N. Ductal carcinoma in situ current trends, controversies, and review of literature. Am J Surg 2018;216(5):998–1003.
9. Luiten JD, Voogd AC, Luiten EJT, Duijm LEM. Trends in incidence and tumour grade in screen-detected ductal carcinoma in situ and invasive breast cancer. Breast Cancer Res Treat 2017;166(1):307–314.
10. Luiten JD, Voogd AC, Tjan-Heijnen VCG, Wesseling J, Luiten EJT, Duijm LEM. Utility of diagnostic breast excision biopsies during two decades of screening mammography. Breast 2019;46:157–162. https://doi.org/10.1016/j.breast.2019.05.018.
11. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ; National Evaluation Team for Breast Cancer Screening in Netherlands Study Group (NETB). Nation-wide data on screening performance during

the transition to digital mammography: observations in 6 million screens. Eur J Cancer 2013;49(16):3517–3525.
12. Jing H, Yang Y, Wernick MN, Yarusso LM, Nishikawa RM. A comparison study of image features between FFDM and film mammogram images. Med Phys 2012;39(7):4386–4394.
13. Setz-Pels W, Duijm LE, Coebergh JW, Rutten M, Nederend J, Voogd AC. Re-attendance after false-positive screening mammography: a population-based study in the Netherlands. Br J Cancer 2013;109(8):2044–2050.
14. Nederend J, Duijm LE, Louwman MW, Groenewoud JH, Donkers-van Rossum AB, Voogd AC. Impact of transition from analog screening mammography to digital screening mammography on screening outcome in the Netherlands: a population-based study. Ann Oncol 2012;23(12):3098–3103.
15. BI-RADS Committee. ACR BI-RADS atlas: Breast Imaging Reporting And Data System. 5th ed. Reston, Va: American College of Radiology, 2013.
16. Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM, et al. The Breast Imaging Reporting and Data System (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. Eur Radiol 2012;22(8):1717–1723.
17. Klompenhouwer EG, Voogd AC, den Heeten GJ, et al. Blinded double reading yields a higher programme sensitivity than non-blinded double reading at digital screening mammography: a prospected population based study in the south of the Netherlands. Eur J Cancer 2015;51(3):391–399.
18. van Bommel RMG, Voogd AC, Nederend J, et al. Incidence and tumour characteristics of bilateral and unilateral interval breast cancers at screening mammography. Breast 2018;38:101–106.
19. Knol MJ, Pestman WR, Grobbee DE. The (mis)use of overlap of confidence intervals to assess effect modification. Eur J Epidemiol 2011;26(4):253–254.
20. Mikhail RA, Nathan RC, Weiss M, et al. Stereotactic core needle biopsy of mammographic breast lesions as a viable alternative to surgical biopsy. Ann Surg Oncol 1994;1(5):363–367.
21. Liberman L, Dershaw DD, Rosen PP, Cohen MA, Hann LE, Abramson AF. Stereotaxic core biopsy of impalpable spiculated breast masses. AJR Am J Roentgenol 1995;165(3):551–554.
22. Meyer JE, Smith DN, Lester SC, et al. Large-core needle biopsy of nonpalpable breast lesions. JAMA 1999;281(17):1638–1641.
23. Rakha EA, Ellis IO. An overview of assessment of prognostic and predictive factors in breast cancer needle core biopsy specimens. J Clin Pathol 2007;60(12):1300–1306.
24. van Breest Smallenburg V, Nederend J, Voogd AC, et al. Trends in breast biopsies for abnormalities detected at screening mammography: a population-based study in the Netherlands. Br J Cancer 2013;109(1):242–248.
25. Sala M, Domingo L, Macià F, Comas M, Burón A, Castells X. Does digital mammography suppose an advance in early diagnosis? Trends in performance indicators 6 years after digitalization. Eur Radiol 2015;25(3):850–859.
26. Bluekens AM, Holland R, Karssemeijer N, Broeders MJ, den Heeten GJ. Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a multicenter study. Radiology 2012;265(3):707–714.
27. Li T, Marinovich ML, Houssami N. Digital breast tomosynthesis (3D mammography) for breast cancer screening and for assessment of screen-recalled findings: review of the evidence. Expert Rev Anticancer Ther 2018;18(8):785–791.
28. Elshof LE, Tryfonidis K, Slaets L, et al. Feasibility of a prospective, randomised, 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):1497–1510.
29. Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. Eur J Cancer 2015;51(16):2296–2303.
30. Szynglarewicz B, Kasprzak P, Biecek P, Halon A, Matkowski R. Screen-detected ductal carcinoma in situ found on stereotactic vacuum-assisted biopsy of suspicious microcalcifications without mass: radiological-histological correlation. Radiol Oncol 2016;50(2):145–152.
31. Evans A, Clements K, Maxwell A, et al. Lesion size is a major determinant of the mammographic features of ductal carcinoma in situ: findings from the Sloane project. Clin Radiol 2010;65(3):181–184.
32. Tang X, Yamashita T, Hara M, Kumaki N, Tokuda Y, Masuda S. Histopathological characteristics of breast ductal carcinoma in situ and association with imaging findings. Breast Cancer 2016;23(3):491–498.
33. Richtlijn mammacarcinoom. Federatie Medisch Specialisten. https://richtlijnendatabase.nl/richtlijn/borstkanker/dcis.html. Published January 7, 2018. Accessed June 12, 2018.
34. Dubrovsky E, Nguyen P, Chun J, et al. Ductal carcinoma in situ on core needle biopsy only with no residual disease at surgery. Breast J 2018;24(6):971–975.
35. Rageth CJ, O'Flynn EA, Comstock C, et al. First International Consensus Conference on lesions of uncertain malignant potential in the breast (B3 lesions). Breast Cancer Res Treat 2016;159(2):203–213.
36. Rageth CJ, O'Flynn EAM, Pinker K, et al. Second International Consensus Conference on lesions of uncertain malignant potential in the breast (B3 lesions). Breast Cancer Res Treat 2019;174(2):279–296 [Published correction appears in Breast Cancer Res Treat 2019;176(2):481–482.].