European Journal of Radiology 136 (2021) 109517

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Trends in delayed breast cancer diagnosis after recall at screening mammography

J.R.C. Lameijer [a,g,*], A.C. Voogd [b,c,g], M.J.M. Broeders [d,e,g], R.M. Pijnappel [d,f,g], W. Setz-Pels [a,g], L.J. Strobbe [h,g], F.H. Jansen [a,g], V.C.G. Tjan-Heijnen [h,g], L.E.M. Duijm [d,i,g]

[a] Department of Radiology, Catharina Hospital Eindhoven, Michelangelolaan 2, 5623 EJ, Eindhoven, the Netherlands
[b] Department of Epidemiology, Maastricht University, GROW, P Debyelaan 1, 6229 HA, Maastricht, the Netherlands
[c] Department of Research, Netherlands Comprehensive Cancer Organization (IKNL), Godebaldkwartier 419, 3511 DT, Utrecht, the Netherlands
[d] Dutch Expert Centre for Screening, Wijchenseweg 101, 6538 SW, Nijmegen, the Netherlands
[e] Department for Health Evidence, Radboud University Medical Center, P.O. Box 9101, 6500 HB, Nijmegen, the Netherlands
[f] Department of Radiology and Nuclear Medicine, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX, Utrecht, the Netherlands
[g] Department of Surgical Oncology, Canisius Wilhelmina Hospital, Weg door Jonkerbos 100, 6532 SZ, Nijmegen, the Netherlands
[h] Department of Internal Medicine, Division of Medical Oncology, GROW, Maastricht University Medical Centre, P Debyelaan 1, 6229 HA, Maastricht, the Netherlands
[i] Department of Radiology, Canisius Wilhelmina Hospital, Weg door Jonkerbos 100, 6532 SZ, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
Epidemiology
Breast cancer
Screening
Malignancy
Delay

ABSTRACT

Objectives: To determine the extent and characteristics of delay in breast cancer diagnosis in women recalled at screening mammography.

Methods: We included a consecutive series of 817,656 screens of women who received biennial screening mammography in a Dutch breast cancer screening region between 1997 and 2016. During at least 3.5 years follow-up, radiological reports and biopsy reports were collected of all recalled women. The inclusion period was divided into four cohorts of four years each. We determined the number of screen-detected cancers and their characteristics, and assessed the proportion of recalled women who experienced a diagnostic delay of at least 4 months in breast cancer confirmation.

Results: The proportion of recalled women who experienced diagnostic delay decreased from 7.5 % in 1997−2001 (47/623) to 3.0 % in 2012−2016 (67/2223, $P < 0.001$). The proportion of women with a delay of at least two years increased from 27.7 % (13/47) in 1997−2001 to 75.7 % (53/70) in 2012−2016 ($P < 0.001$). Cancers with a diagnostic delay > 2 years were more frequently invasive ($P = 0.009$) than cancers with a diagnostic delay of 4−24 months. The most frequent cause of diagnostic delays was incorrect radiological classifications by clinical radiologists (55.2 % overall) after recall.

Conclusions: The proportion of recalled women with a delayed breast cancer diagnosis has more than halved during two decades of screening mammography. Delays in breast cancer diagnosis are characterized by longer delay intervals, although the proportion of these delays among all screen-detected cancers has not increased. Preventing longer delays in breast cancer confirmation may help improve breast cancer survival.

1. Introduction

In many countries breast cancer is one of the most commonly diagnosed malignancies in women [1]. Screening mammography programmes have been established with the aim to reduce breast cancer morbidity and mortality through early detection and treatment of the disease. In combination with improved therapy, this early detection has resulted in a significant reduction of breast cancer mortality over the past decades [2,3]. After recall for a suspicious abnormality at screening mammography, a timely confirmation of a breast malignancy is of

Abbreviations: BI-RADS, Breast Imaging Reporting and Data System; CCMO, Dutch Central Committee on Research involving Human Subjects; CEDM, contrast enhanced digital mammography; CNB, core needle biopsy; DBT, digital breast tomosynthesis; FFDM, full field digital mammography; FNAB, fine-needle aspiration biopsy; PPV, positive predictive value; SCNB, stereotactic core needle biopsy.

* Corresponding author at: Department of Radiology, Catharina Hospital Eindhoven, Michelangelolaan 2, 5623, EJ, Eindhoven, the Netherlands.
E-mail address: joost.lameijer@rmh.nhs.co.uk (J.R.C. Lameijer).

https://doi.org/10.1016/j.ejrad.2020.109517
Received 22 September 2020; Received in revised form 19 December 2020; Accepted 28 December 2020
Available online 2 January 2021

utmost importance in order to prevent a treatment delay. A diagnostic delay after recall may cause the tumour to be diagnosed in a more advanced stage, which could have a negative impact on breast cancer survival [4].

Several studies have shown that up to 5% of recalled women experience a delay in the confirmation of their breast cancer. [5,6] Most of these studies have been performed in the era of screen-film mammography (SFM). Currently, full-field digital mammography (FFDM) has replaced screen-film mammography in both the screening setting and in the setting of clinical mammography in most countries. Also, the management of breast disease has changed considerably over the years, with the introduction of new diagnostic modalities (digital breast tomosynthesis, 3D ultrasonography, spectral mammography, advanced magnetic resonance techniques), replacement of fine needle aspiration cytology by core biopsy, implementation of multidisciplinary meetings, and a further specialization of health care professionals in breast diseases [6]. For these reasons, one may assume that the proportion of women who face a delay in their breast cancer diagnosis after recall at screening mammography has decreased over time. At the same time, a reduction in length of delay is expected due to increased diagnostic accuracy preventing false positive recall and repeated recall, while also reducing the need for radiological follow-up in cases of equivocal findings at mammography or ultrasound. We therefore determined trends in the frequency of delayed breast cancer confirmation and lengths of delays in women recalled at screening mammography. We also assessed the causes of these delays and investigated tumour characteristics of these breast cancers with short and longer delay intervals in an observational follow-up study spanning two decades of screening mammography conducted in the south of the Netherlands.

2. Material and methods

2.1. Study population and screening procedure

We included a consecutive series of 817,656 screens (SFM: 49,318 initial screens and 328,637 subsequent screens; FFDM: 48,223 initial screens and 391,478 subsequent screens) of women aged 50–75 years who received biennial screening mammography in a southern region of the Netherlands between January 1, 1997 and December 31, 2016 (Table 1). Women participating in the screening programme can indicate that they do not give permission to use their data for quality assessment and scientific purposes. Three recalled women did not give this permission and they were excluded from analysis. Ethical approval was waived by the Dutch Central Committee on Research involving Human Subjects (CCMO).

All screening mammograms were obtained by certified radiographers at one of four dedicated screening units (one fixed unit and three mobile units). Screen-film mammography was replaced by full-field digital mammography in 2009/2010. The examinations were routinely and independently double read by a team of 17 certified screening radiologists. The examinations were routinely and independently double read by a team of 17 certified screening radiologists, each reading more than 10,000 screening mammograms yearly. The radiologists classified the mammographic abnormality in case of recall (suspicious mass, suspicious calcifications, suspicious mass in combination with calcifications, asymmetry, architectural distortion or other suspicious abnormality). Recalled women were referred by their general practitioner to a hospital breast unit for further analysis of the mammographic abnormality. In case of a false positive recall (i.e., no breast cancer at workup) women were encouraged to return to the screening programme.

2.2. Workup of recalled women and follow-up

A total of 36 hospitals were involved in the workup of recalled women, of which the majority (99.4 %, 18,476/18,592) was analysed in eight hospitals centrally located in our screening region. Recalled women first underwent physical examination by a surgical oncologist or dedicated breast nurse at the surgical department. The type of diagnostic workup was at the discretion of the clinical radiologist and could include additional mammographic views, breast tomosynthesis, breast ultrasonography (including 3D ultrasonography) and/or magnetic resonance imaging of the breasts. Clinical breast imaging was followed by percutaneous breast biopsy if indicated, and included fine needle aspiration biopsy (FNAB), ultrasound guided core needle biopsy (CNB, 14–18 Gauge) and/or vacuum assisted core needle biopsy (9–11 Gauge, either stereotactic guided or MRI guided). The use of surgical biopsy for diagnostic purposes sharply declined over the years and is currently reserved for those cases where malignancy has not been ruled out by percutaneous biopsy. [7] Recalled women were discussed at multidisciplinary meetings that were gradually implemented in the hospitals.

During a minimum of 3.5 years follow-up (until July 1, 2020), one of the screening radiologists (LD) and several radiology residents collected clinical data and data from diagnostic breast imaging, biopsy procedures and surgical interventions of all recalled women. The radiologist entered these data in a database which had been constructed for quality assurance of the regional screening programme. Breast cancers were categorized into ductal carcinoma in-situ (DCIS) and invasive cancers. Lobular carcinoma in-situ was considered a non-malignant lesion. The TNM classification (6th and 7th edition) was used for malignant lesions. [8,9] For women with a bilateral malignancy, the cancer with the highest TNM was retained and multiple foci of cancer in one breast were counted as one cancer.

2.3. Delayed breast cancer diagnosis

In the current study, we defined an interval of at least four months between the recall at screening mammography and the confirmation of breast cancer as a diagnostic delay. The 20-year inclusion period was divided into four cohorts of five years each: 1997−2001, 2002−2006, 2007−2011 and 2012−2016. For each cohort, we determined the number of and characteristics of detected cancers at recall, and assessed the proportion of recalled women who experienced a diagnostic delay of at least 4 months in breast cancer confirmation. The database for quality assurance was used to determine the interval between recall and time of diagnosis of breast cancer to determine whether an interval of at least four months between recall and diagnosis had occurred. The database for quality assurance was also used to identify women who had been

Table 1
Proportions of women with a diagnostic delay in breast cancer diagnosis after recall at screening mammography.

	Year of screening				Total
	1997−2001	2002−2006	2007−2011	2012−2016	
Screens, *N*	128081	155398	209523	324654	817616
Recalls, *N*	1324	2123	4972	10173	18592
Screen detected cancers, *N*	568	750	1227	2148	4693
Cancers with a diagnostic delay, *N* (%)	47 (7.6)	53 (6.6)	69 (5.3)	70 (3.1)	239
Delay 4−24 months	34 (72.3)	34 (64.2)	21 (30.4)	17 (24.3)	106 (44.4)
Delay ≥24 months	13 (27.7)	19 (35.8)	48 (69.6)	53 (75.7)	133 (55.6)

recalled twice for the same mammographic abnormality between January 1997 and January 2017. A team of three radiologists (WS, FJ, LD) then determined whether the second recall concerned the same mammographic abnormality for which a woman had been recalled previously. Each case was independently assessed by two radiologists and discrepant observations between them were solved by consensus. To determine whether an incorrect radiological assessment after recall had resulted in a diagnostic delay, the radiologists also independently and retrospectively reviewed the screening mammograms and diagnostic breast imaging and intervention reports of all women with a diagnostic delay. Each reviewer classified the lesions using the Breast Imaging Reporting and Data System (BI-RADS). [10,11] Finally, the radiologists reviewed the clinical data (including biopsy reports, discharge records, outcome of multi-disciplinary team meetings) to identify other causes than a false negative radiological assessment for a delayed breast cancer diagnosis, including false-negative biopsy results or patient related delays.

Parameters such as tumour histology, nodal involvement, histological grading and surgery performed were calculated for women with a delay of respectively 4−24 months or a delay of at least 24 months between recall and breast cancer diagnosis. We also compared these characteristics of breast cancers with a diagnostic delay with those of screen-detected cancers without this delay.

2.4. Statistical analysis

Descriptive statistics were performed using Statistical Package for Social Science 23.0 (SPSS Inc., IBM, Chicago, IL). The chi-square test was used to test for differences between women with a diagnostic delay of 4−24 months and women with a diagnostic delay of at least 24 months with respect to tumour characteristics i.e. histology, stage, nodal status, receptor status, and type of final surgical treatment. The proportions of women with diagnostic delays (4−24 months versus >24 months) in the four time periods (1997−2001, 2002−2006, 2007−2011, and 2012−2016) were tested using the chi-square test. Each delay was categorized according to causes of delayed diagnosis over the time periods as mentioned above, and differences in proportions were tested using the chi-square test. Her2/Neu-receptor status was routinely determined from 2008 onward and thus not available for the earlier inclusion periods. Missing data were excluded in the Chi-square test analyses when comparing tumour characteristics for the two delay cohorts. A *P*-value of less than 0.05 was considered to indicate a statistically significant difference. P-values were two-sided.

3. Results

3.1. Overall screening outcome

A total of 18,592 women were recalled at screening mammography (recall rate: 2.3 % (18,592/817,656)), of whom 4693 proved to have breast cancer at workup (6.0 cancers detected per 1000 screens (4693/817,656), positive predictive value of recall: 25.2 % (4693/18,592). The 4693 screen detected cancers comprised 907 DCIS (19.3 %) and 3786 (80.7 %) invasive cancers.

3.2. Women with a delayed diagnosis of breast cancer

A delay in breast cancer diagnosis of at least four months was present in 4.8 % of recalled women with breast cancer at workup (239/4932). This percentage gradually decreased from 7.6 % in 1997−2001 (47/615) to 3.1 % in 2012−2016 (70/2218, $P < 0.001$, Table 1).

Among all delays, the proportion of women with a delay of at least 24 months increased from 27.7 % (13/47) in 1997−2001 to 75.7 % (53/70) in 2012−2016 ($P < 0.001$). Fig. 1 shows the number of delayed cancer diagnosis per 1000 screen-detected cancers. A decline in the rate of cancers with a 4−24 months diagnostic delay is seen over the years, whereas the rate of cancers confirmed at least 24 months after recall remained stable (except for the cohort screened 2007−2011).

In 93.2 % (124/133) of the women with a delay of at least 24 months, the breast cancer was finally confirmed after a second recall for the same lesion for which she had been recalled previously. This second recall took place at the subsequent screen, two years after the initial recall, in 59.7 % of the women (74/124) and at a later subsequent screening round in 40.3 % (50/124). Six women, who preferred opportunistic screening rather than returning to the screening programme, proved to have breast cancer at one of the subsequent opportunistic screening mammograms. Two women with a palpable breast lump consulted a surgical oncologist more than two years after their false positive recall, with breast cancer confirmed at the site of the previous abnormality at screening mammography. One diagnostic delay was due to a patient's refusal to undergo biopsy. Cancers with a diagnostic delay of at least 2 years were more frequently invasive ($P = 0.009$), but showed a similar Bloom & Richardson grading ($P = 0.16$) compared to cancers with a diagnostic delay of 4−24 months (Table 2). Tumour size and lymph node status were comparable for both groups, as well as the type of surgical treatment (breast conserving surgery versus mastectomy).

Tumour characteristics of breast cancers with a delayed diagnosis were also compared to those of screen-detected cancers without a diagnostic delay, as shown in Table 3. The two groups differed in histological profile, with more ductal and ductolobular cancer among screen-detected breast cancers ($P < 0.001$). Screen-detected, invasive

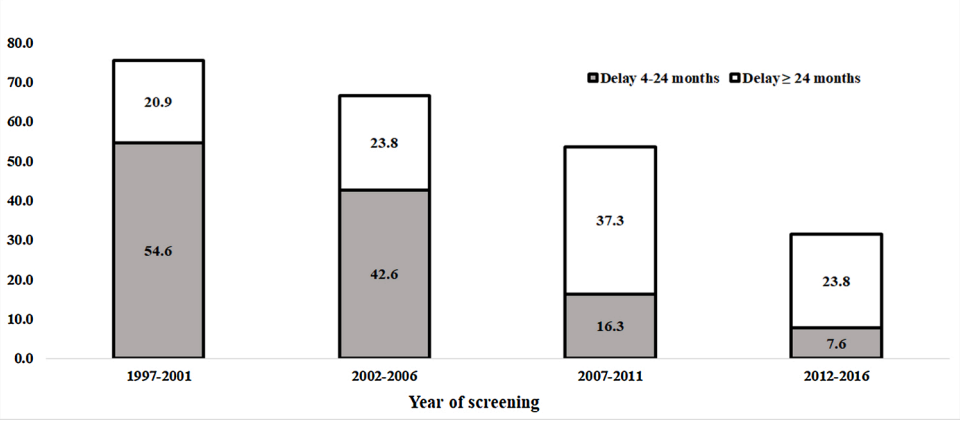

Fig. 1. Time trend in number of delays per 1000 screen detected cancers.

Table 2
Tumour characteristics and surgical therapy of breast cancers with a diagnostic delay of 4-24 months versus a diagnostic delay of at least 2 years.

	Delay in breast cancer diagnosis 4−24 months ($N = 106$)	Delay in breast cancer diagnosis ≥2 years ($N = 133$)	*P*
Type of cancer, N (%)			0.007
DCIS	26 (24.5)	15 (11.5)	
Invasive	80 (75.5)	118 (88.5)	
Histology of invasive cancers, *N* (%)			0.92
Ductal	58 (72.5)	88 (74.6)	
Lobular	11 (13.8)	14 (11.8)	
Ductolobular	1 (1.3)	1 (0.8)	
Other	8 (10.0)	15 (12.8)	
Unknown	2 (2.5)	0	
Tumour stage of invasive cancers, *N* (%)			0.40
T1a + b	33 (41.3)	39 (33.1)	
T1c	33 (41.3)	56 (47.5)	
T2+	12 (15.0)	23 (19.4)	
Unknown	2 (2.5)	0	
Lymph node status of invasive cancers, *N* (%)			0.08
N+	20 (25.0)	18 (15.3)	
N-	57 (71.3)	97 (82.2)	
Unknown	3 (3.8)	3 (2.5)	
Bloom & Richardson grade, *N* (%)			0.16
I	31 (38.8)	45 (38.1)	
II	22 (27.6)	56 (47.5)	
III	12 (15.0)	15 (12.8)	
Unknown	15 (18.8)	2 (1.6)	
Estrogen receptor status, No (%)			0.90
Positive	66 (82.5)	105 (88.9)	
Negative	8 (10.0)	12 (10.3)	
Unknown	6 (7.5)	1 (0.8)	
Progesterone receptor status, No (%)			0.08
Positive	58 (72.5)	77 (65.3)	
Negative	16 (20.0)	39 (33.1)	
Unknown	6 (7.5)	2 (1.6)	
Her2/Neu receptor status, No (%)			0.020*
Positive	9 (11.3)	9 (7.6)	
Negative	31 (38.8)	98 (83.1)	
Unknown	40 (50.0)	11 (9.3)	
Final surgical treatment, No (%)			0.13
Breast conserving surgery	74 (69.8)	106 (79.7)	
Mastectomy	27 (25.5)	25 (18.6)	
No surgery	5 (4.7)	2 (1.5)	

DCIS = Ductal carcinoma in-situ *In earlier cohorts (until 2008), Her2/Neu receptor status was not routinely reported, yielding a large number of missing data. Missing numbers were not incorporated in the Chi-square test.

Table 3
Tumour characteristics and surgical therapy of breast cancers with a diagnostic delay versus screen detected cancers without delay.

	Breast cancers with diagnostic delay ($N = 239$)	Screen detected cancers (no delay) ($N = 4693$)	*P*
Type of cancer, N (%)			0.41
DCIS	41 (17.2)	907 (19.3)	
Invasive	198 (82.8)	3786 (80.7)	
Histology of invasive cancers, *N* (%)			<0.001
Ductal	146 (73.7)	2943 (77.7)	
Lobular	25 (12.6)	453 (11.9)	
Ductolobular	2 (1.0)	159 (4.2)	
Other	23 (11.6)	214 (5.6)	
Unknown	2 (1.1)	17 (0.5)	
Tumour stage of invasive cancers, *N* (%)			0.47
T1a + b	72 (36.3)	1246 (32.9)	
T1c	89 (44.9)	1735 (45.8)	
T2+	35 (17.6)	786 (20.7)	
Unknown	2 (1.2)	19 (0.6)	
Lymph node status of invasive cancers, *N* (%)			0.24
N+	38 (19.1)	905 (23.9)	
N-	154 (77.7)	2801 (73.4)	
Unknown	6 (3.2)	80 (2.7)	
Bloom & Richardson grade, *N* (%)			0.58
I	76 (38.3)	1563 (41.3)	
II	78 (39.4)	1520 (40.1)	
III	27 (13.6)	453 (14.0)	
Unknown	17 (8.7)	250 (4.6)	
Estrogen receptor status, No (%)			0.99
Positive	171 (86.4)	3279 (86.6)	
Negative	20 (10.1)	376 (9.9)	
Unknown	7 (3.5)	131 (3.5)	
Progesterone receptor status, No (%)			0.74
Positive	135 (68.2)	2678 (70.7)	
Negative	55 (27.7)	968 (25.6)	
Unknown	8 (4.1)	140 (3.7)	
Her2/Neu receptor status, No (%)			0.38*
Positive	18 (9.1)	278 (7.3)	
Negative	129 (65.2)	2635 (69.6)	
Unknown	51 (25.7)	873 (23.1)	
Final surgical treatment, No (%)			<0.001
Breast conserving surgery	180 (75.3)	3966 (84.5)	
Mastectomy	52 (21.8)	673 (14.3)	
No surgery	7 (2.9)	54 (1.2)	

DCIS = Ductal carcinoma in-situ *In earlier cohorts (until 2008), Her2/Neu receptor status was not routinely reported, yielding a large number of missing data. Missing numbers were not incorporated in the Chi-square test.

cancers were more frequently lymph node positive (23.8 % (905/3786 versus 19.1 % (38/239), although this difference was not statistically significant (P = 0.24). The Bloom and Richardson tumour grading was comparable for both groups (P = 0.58). Patients with a diagnostic delay were more likely to be treated with mastectomy (21.8 % versus 14.3 %, P < 0.001).

3.3. Causes of diagnostic delays

In all 4 cohorts an incorrect radiological classification given by the clinical radiologists after recall was the most frequently encountered cause of a diagnostic delay (Table 4). These women received a BI-RADS 1 (no abnormalities) or BI-RADS 2 (benign) classification at workup, without additional biopsy. The reviewing study radiologists (LD, WS, FJ), however, considered the lesions suspicious or malignant in retrospect, with the necessity of subsequent biopsy. Examples of erroneous BI-RADS classifications are shown in Figs. 2 and 3.

Among all delays, the proportion of women with an erroneous radiological classification increased from 42.6 % in 1997/2001 (20/47) to 72.9 % in 2012−2016 (51/70, $P = 0.001$) and this increase was most prominent in women with diagnostic delays of at least 24 months (Fig. 4). False negative percutaneous biopsy was another major cause of a diagnostic delay and comprised 15.7 % (2012−2016) to 30.2 % (2002−2006) of all delays in the screened cohorts. The proportion of women who experienced a delayed breast cancer confirmation as a result of radiological surveillance of a probably benign lesion at radiology significantly decreased over the years, from 21.3 % (10/47) in 1997−2001 to 5.7 % (4/70) in 2012−2016 ($P = 0.01$). None of the

Table 4
Causes of diagnostic delay in breast cancer diagnosis following recall at screening mammography.

Cause of diagnostic delay, *N* (%)	Year of screening				Total
	1997–2001	2002–2006	2007–2011	2012–2016	
Incorrect radiological classification	20 (42.6)	21 (39.6)	40 (58.0)	51 (72.9)	132 (55.2)
False negative percutaneous biopsy	10 (21.3)	16 (30.2)	15 (21.7)	11 (15.7)	52 (21.8)
Probably benign lesion at radiology (BI-RADS 3) followed by radiological surveillance	10 (21.3)	12 (22.6)	12 (17.4)	4 (5.7)	38 (15.9)
False negative surgical (excision) biopsy	4 (8.5)	1 (1.9)	0 (0)	2 (2.9)	7 (2.9)
Other:					
Surgical oncologist does not follow radiologist's advice of (repeated) lesion biopsy	1 (2.1)	1 (1.9)	1 (1.4)	0 (0)	3 (1.3)
Surgical oncologist does not follow pathologist's advice of lesion excision	1 (2.1)	1 (1.9)	0 (0)	0 (0)	2 (0.8)
Surveillance of a radiologically suspicious lesion which is inaccessible to percutaneous biopsy	0 (0)	0 (0)	1 (1.4)	1 (1.4)	2 (0.8)
Patient refuses biopsy	1 (2.1)	1 (1.9)	0 (0)	1 (1.4)	3 (1.3)
Total	47	53	69	70	239

BI-RADS = Breast Imaging Reporting and Data System.

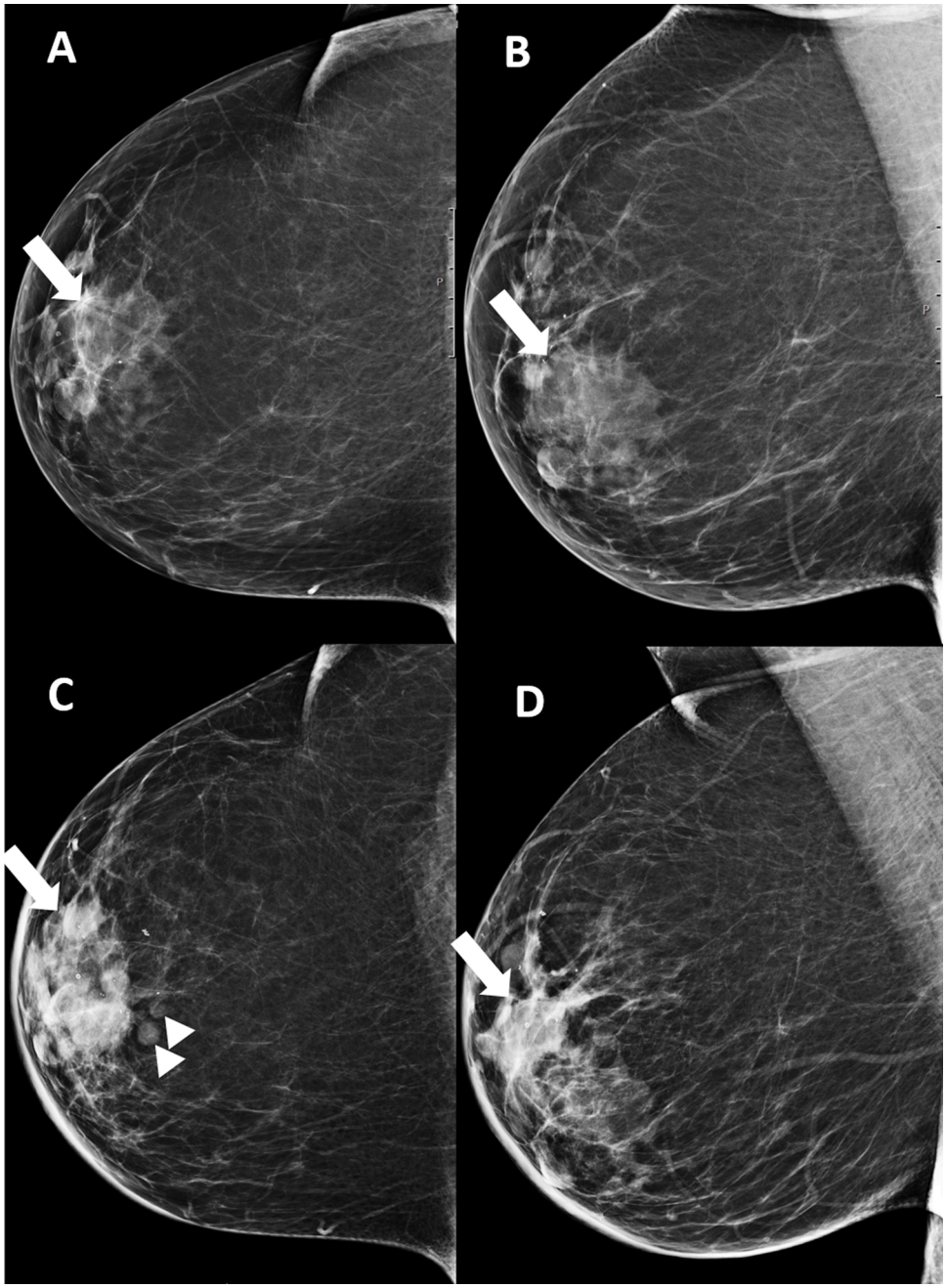

Fig. 2. Patient example of a 12-month delay due to erroneous BI-RADS classification by the clinical radiologist. Two-view screening mammograms (A and C, cranio-caudal (CC) view, and B and D, medio-lateral oblique (MLO) view) of the right breast at recall in 2013. Images A and B show a lesion in the retroareolar region of the breast (white arrows), classified as BI-RADS 4 by the screening radiologists. Digital breast tomosynthesis and ultrasound were performed, and the lesion was classified as BI-RADS 2 (benign lesion). No biopsy was performed. Twelve months later the patient reported to the surgical department with a growing lump in the right breast. Images C and D in 2014 show increased density of the retroareolar lesion in the right breast (white arrows), with possible satellite lesions (white arrowheads). Targeted ultrasound showed an irregular, hypoechogenic lesion with satellite lesions, which was classified as BI-RADS 5. Ultrasound guided true-cut biopsy was performed which revealed an invasive ductal carcinoma with axillary metastases (TNM: T3N2M0). The patient was treated with neo-adjuvant chemotherapy, mastectomy, radiotherapy, and hormonal therapy.

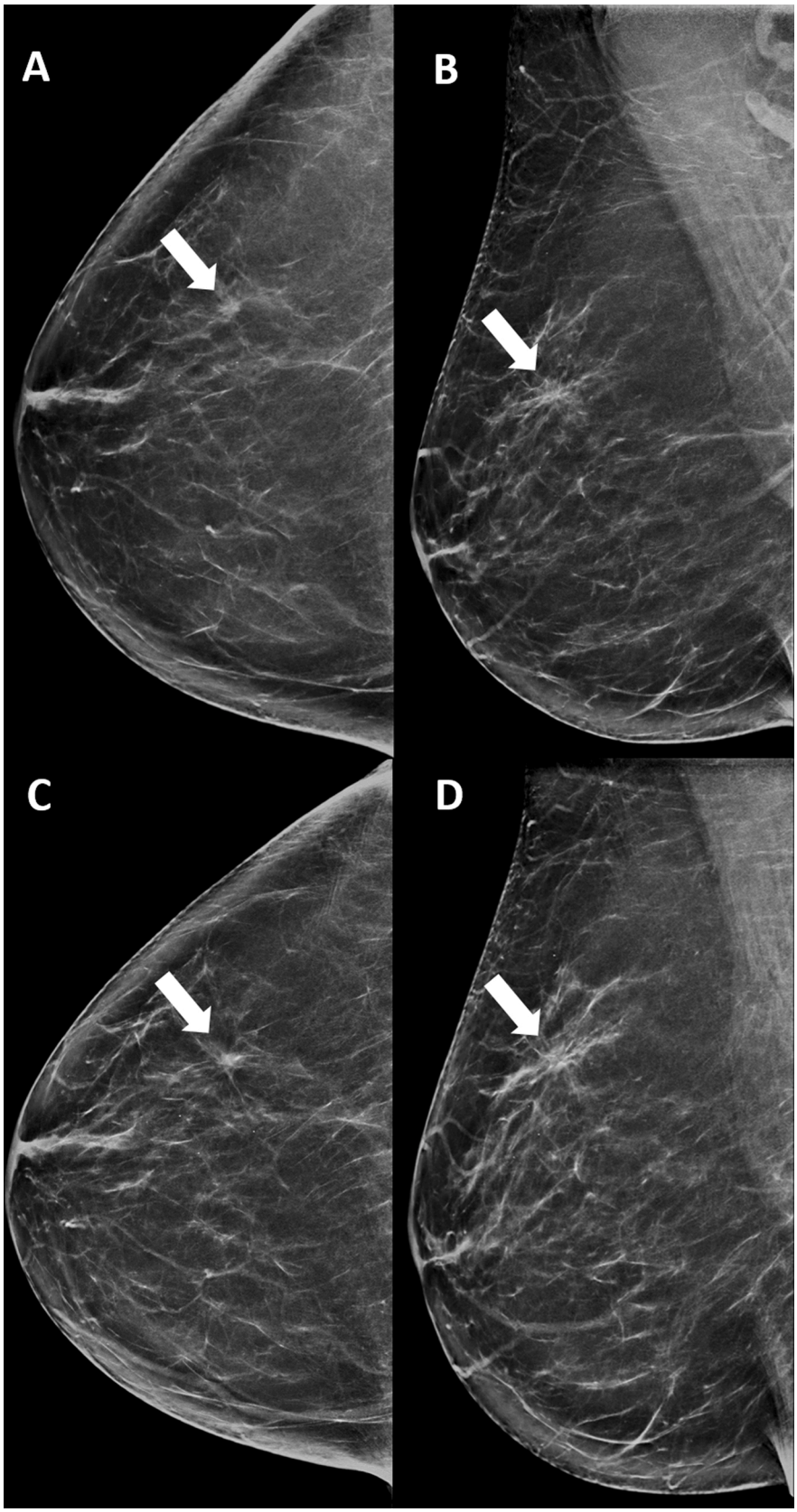

Fig. 3. Patient example of a 24-month delay due to erroneous BI-RADS classification by the clinical radiologist. Two-view screening mammograms (A and C, craniocaudal (CC) view, and B and D, medio-lateral oblique (MLO) view) of the right breast at recall in 2015 show a lesion in the upper lateral quadrant of the right breast. Images A and B show the lesion in 2015 (white arrows), classified as BI-RADS 4 by the screening radiologists. Digital breast tomosynthesis and ultrasound were performed, and the lesion was classified as BI-RADS 2 (benign lesion). No biopsy was performed. At repeated recall in 2017 for the same lesion (again characterized as BI-RADS 4 by the screening radiologists), increased density of the lesion was seen, with an increase in size (images C and D). Targeted ultrasound showed a spiculated, irregular and hypoechogenic lesion, which was classified as BI-RADS 5. Ultrasound guided true-cut biopsy was performed which revealed an invasive ductal carcinoma. There were no axillary metastases (TNM: T1bN0). The patient was treated with breast conserving surgery and adjuvant radiotherapy.

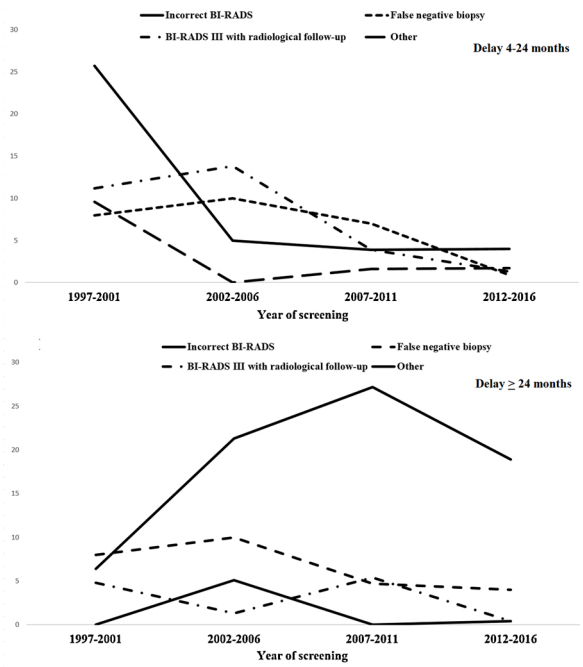

Fig. 4. Trends in causes of delayed breast cancer diagnosis per 1000 screen detected cancers.

women who underwent follow-up imaging for a BI-RADS 3 abnormality experienced a diagnostic delay as a result of non-adherence to the recommended follow-up examinations.

4. Discussion

In the current study we found that the proportion of recalled women with a delay of at least four months in their confirmation of breast cancer more than halved during two decades of screening mammography. This reduction was mainly due to the sharp decline in the number of women with a delay of 4−24 months in their breast cancer diagnosis. At the same time, the number of women who experienced a delay of at least 24 months in their breast cancer diagnosis, per 1000 screen-detected cancers, has remained stable over the years. Tumours that were diagnosed after a delay of at least 24 months were more often invasive and less frequently Her2/Neu positive compared to breast cancers that were confirmed after a delay of 4−24 months, but other tumour histology characteristics, tumour size and nodal status were comparable for both groups. Frequent causes of delay included incorrect radiological classification at recall and false negative percutaneous biopsy.

Several studies have shown that longer diagnostic and therapeutic intervals may have a detrimental effect on breast cancer survival. Therefore, a timely diagnosis of breast cancer not only requires the recognition of suspicious mammographic findings at screening, but it also necessitates a prompt confirmation of the malignancy after recall. Over the years, we observed a significantly reduced probability of a delayed breast cancer confirmation after recall. The last decade has seen substantial changes in the diagnostic management of (a)symptomatic breast disease, with the introduction of multi-disciplinary meetings and new imaging modalities (e.g., digital breast tomosynthesis), an intensified use of advanced imaging modalities (breast MRI), a further replacement of fine needle biopsy by (vacuum assisted) large core biopsy, and an increased sub-specialization of physicians in breast care [6,12,13]. These developments have likely resulted in the decline of women facing a delayed breast cancer confirmation. However, this improvement was mainly present for diagnostic delays of less than 24 months but not for longer delays. All Dutch screening radiologists are trained and certified by the Dutch Expert Centre for Screening (LRCB). This centre increasingly emphasizes a timely confirmation of the benign or malignant nature of a recalled mammographic abnormality [14,15]. The LRCB therefore discourages abnormalities to be classified as probably benign at workup, if possible, as these lesions require follow-up for 6–24 months. Although this policy shortens the period of uncertainty among recalled women and helps to reduce short delays after recall, it may at the same time result in longer diagnostic delays, i.e. when radiological re-evaluation of a probably benign breast lesion at follow-up would have resulted in additional biopsy rather than a second recall for the same mammographic abnormality at a later screening round. The shift from shorter to longer diagnostic delays is related to the finding that the most frequent cause of a diagnostic delay in this study was an incorrect radiological classification at workup, with a proportional increase of these delays over the years. At the same time, the probability to experience a delayed breast cancer confirmation as a result of radiological surveillance of a probably benign lesion decreased, as a decrease in the amount of BI-RADS 3 classifications after radiological recall was seen (from 21.3 % in 1997−2001 to 5.7 % in 2012−2016, $P = 0.01$). Diagnostic improvements at screening mammography result in an increased lesion detection, reflected by increased recall rates and cancer detection as shown previously [16–19]. Subtle lesions may have a larger potential to receive an incorrect workup after recall and may also lead to equivocal findings necessitating additional follow-up. It is subject of debate whether a delayed diagnosis of small lesions significantly influences prognosis and further studies are needed to address this topic [4,13]. We found that the size of the cancers with a diagnostic delay was comparable to those of screen-detected cancers, with a similar grading and less lymph node involvement. Therefore, the prognosis of both groups will most likely be comparable, whereas the diagnostic delay will have a certain negative impact on survival due to increase in tumour size and the development of nodal involvement during the diagnostic delay [13]. Several studies have shown that interval cancers have a significantly worse prognosis than screen-detected cancers [20,21]. However, our paper focuses on cancers diagnosed in recalled women and we did not evaluate interval cancers. Nevertheless, it will be very interesting and clinically valuable to determine the contribution of delayed cancer diagnoses in recalled women among all potentially preventable delayed cancer diagnoses. This requires the determination of the proportion and tumour characteristics of “missed” cancers among all interval cancers, as well as the determination of these parameters for screen-detected cancers that were missed at the previous screening round. This issue is definitely a topic for further research.

The breast cancers in the two groups of diagnostic delays showed comparable tumour characteristics except of fewer DCIS and more Her2/Neu negative invasive cancers among women with a longer delay. DCIS is usually detected through the presence of calcifications at mammography. The availability of stereotactic core needle biopsy as an alternative to surgical biopsy has gradually resulted in a routine biopsy of BI-RADS 3 calcifications rather than follow-up, which may explain the low proportion of DCIS among women experiencing a long delay in their breast cancer confirmation. To improve the communication between screening radiologists and clinicians, the BI-RADS classification for mammographic abnormalities was implemented in the Dutch screening mammography programme in 2009, along with the replacement of screen-film by digital screening mammography [15,22]. The BI-RADS classification system increases inter-observer agreement between breast radiologists and helps to guide the diagnostic pathways at recall. A further reduction in the number of women with a delayed breast cancer confirmation may be obtained by restricting the workup to dedicated breast radiologists in accredited hospitals. The Dutch screening programme differs from those in several other European countries and the USA as screening and recall are not integrated. Hospitals handling recalled women are separated from the screening units and quality control of the workup is not mandatory [15,23]. The Dutch

screening programme, however, is subject to quality assurance, consistent with recommendations of the European Society of Breast Imaging (EUSOBI) and International Agency for Research on Cancer (IARC) [23–26]. The introduction of quality assurance and assessment sessions may prove beneficial for hospitals handling recall, reducing the number of recalled women facing a delayed cancer diagnosis.

Our study has several strengths and limitations. To our knowledge, this is the first study focusing on time trends in delayed breast cancer confirmation after recall for a screen detected mammographic abnormality. A large series of consecutive screening mammograms with almost complete follow-up was analysed during two decades of screening mammography. On the other hand, the relatively short follow-up period of at least 3.5 years may not have enabled us to identify all cases with a longer delay in the most recently screened cohort. Also, screening mammography programmes are continuously subject to changes and the design of these programmes differs among countries. Assessment after recall may be performed in hospitals that are not linked to the screening programme, like in the Netherlands, or may be embedded in the screening procedure itself. The missing data of the Her2/Neu-receptor status in patients with a 4−24 months of delay may have hampered statistical analysis. Unfortunately, we do not have data available on the amount of clinical mammograms read by screening radiologists and non-screening radiologists. It will be interesting to investigate whether the quality of the breast imaging assessments after recall differs between screening and non-screening radiologists but we currently lack the data to analyse this issue.

We conclude that the proportion of recalled women with a delay in their confirmation of breast cancer has significantly decreased during two decades of screening mammography, but these delays are now characterized by longer delay intervals. Quality assurance should not only cover the screening mammography programmes, but it should also focus on the hospitals handling the workup of recalled women, in order to prevent a worse breast cancer survival due to a delayed breast cancer confirmation. Currently, none of the Dutch hospitals are involved in a mandatory or non-mandatory quality assurance programme concerning the handling of women recalled at screening mammography, although the hospitals can receive a yearly distinction on breast cancer care in general. With the results of our study we hope to be able to start a discussion about the necessity to start a quality assurance programme for hospitals involved in the evaluation of women recalled at screening mammography.

CRediT authorship contribution statement

J.R.C. Lameijer: Methodology, Software, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **A.C. Voogd:** Methodology, Formal analysis, Investigation, Writing - review & editing, Visualization, Supervision. **M. J.M. Broeders:** Methodology, Writing - review & editing. **R.M. Pijnappel:** Methodology, Writing - review & editing. **W. Setz-Pels:** Investigation, Data curation, Writing - review & editing. **L.J. Strobbe:** Writing - review & editing. **F.H. Jansen:** Investigation, Data curation, Writing - review & editing. **V.C.G. Tjan-Heijnen:** Methodology, Writing - review & editing, Supervision. **L.E.M. Duijm:** Conceptualization, Methodology, Software, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors report no declarations of interest.

References

[1] Global Cancer Observatory, International Agency for Research on Cancer, 2019 (Accessed 15th of September, 2020, at http://gco.iarc.fr.).
[2] E.R. Myers, P. Moorman, J.M. Gierisch, L.J. Havrilesky, L.J. Grimm, S. Ghate, B. Davidson, R.C. Mongtomery, M.J. Crowley, D.C. McCrory, A. Kendrick, G. D. Sanders, Benefits and harms of breast Cancer screening: a systematic review, JAMA 314 (2015) 1615–1634, https://doi.org/10.1001/jama.2015.13183.
[3] V.D.V. Sankatsing, N.T. van Ravesteyn, E.A.M. Heijnsdijk, C.W.N. Looman, P. A. van Luijt, J. Fracheboud, G.J. den Heeten, M.J.M. Broeders, H.J. de Koning, The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up, Int. J. Cancer 141 (2017) 671–677, https://doi.org/10.1002/ijc.30754.
[4] M.J. Broeders, N.C. Onland-Moret, H.J. Rijken, J.H. Hendriks, A.L. Verbeek, R. Holland, Use of previous screening mammograms to identify features indicating cases that would have a possible gain in prognosis following earlier detection, Eur. J. Cancer 39 (2003) 1770–1775, https://doi.org/10.1016/s0959-8049(03)00311-3.
[5] S. Ciatto, N. Houssami, D. Ambrogetti, R. Bonardi, G. Collini, M.R. Del Turco, Minority report - false negative breast assessment in women recalled for suspicious screening mammography: imaging and pathological features, and associated delay in diagnosis, Breast Cancer Res. Treat. 105 (2007) 37–43, https://doi.org/10.1007/s10549-006-9425-3.
[6] L.E. Duijm, J.H. Groenewoud, H.J. de Koning, J.W. Coebergh, M. van Beek, M. J. Hooijen, L.V. van de Poll-Franse, Delayed diagnosis of breast cancer in women recalled for suspicious screening mammography, Eur. J. Cancer 45 (2009) 774–781, https://doi.org/10.1016/j.ejca.2008.10.020.
[7] J.D. Luiten, A.C. Voogd, V.C.G. Tjan-Heijnen, J. Wesseling, E.J.T. Luiten, L.E. M. Duijm, Utility of diagnostic breast excision biopsies during two decades of screening mammography, Breast. 46 (2019) 157–162, https://doi.org/10.1016/j.breast.2019.05.018.
[8] L.H. Sobin, C. Wittekind, International Union against cancer. TNM Classification of Malignant Tumours, 6th ed., Wiley-Liss, New York, 2002.
[9] L.H. Sobin, M.K. Gospodarowicz, C. Wittekind, International Union against cancer., Ebrary inc. TNM classification of malignant tumours. Chichester, West Sussex, 7th ed., Wiley-Blackwell, UK ; Hoboken, NJ, 2009, p. 310, p. xx.
[10] American College of Radiology, BI-RADS Committee. ACR BI-RADS atlas: Breast Imaging Reporting and Data System, 4th ed., American College of Radiology, Reston, VA, 2003.
[11] American College of Radiology, . BI-RADS Committee. ACR BI-RADS atlas: Breast Imaging Reporting and Data System, 5th ed., American College of Radiology, Reston, VA, 2013.
[12] L.E. Duijm, J.H. Groenewoud, R.M. Roumen, H.J. de Koning, M.L. Plaisier, J. Fracheboud, A decade of breast cancer screening in The Netherlands: trends in the preoperative diagnosis of breast cancer, Breast Cancer Res. Treat. 106 (2007) 113–119, https://doi.org/10.1007/s10549-006-9468-5.
[13] J.R.C. Lameijer, A.C. Voogd, R.M. Pijnappel, W. Setz-Pels, M.J. Broeders, V.C. G. Tjan-Heijnen, L.E.M. Duijm, Delayed breast cancer diagnosis after repeated recall at biennial screening mammography: an observational follow-up study from the Netherlands, Br. J. Cancer (2020), https://doi.org/10.1038/s41416-020-0870-2.
[14] E. Lazarus, M.B. Mainiero, B. Schepps, S.L. Koelliker, L.S. Livingston, BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value, Radiology. 239 (2006) 385–391, https://doi.org/10.1148/radiol.2392042127.
[15] J.M. Timmers, H.J. van Doorne-Nagtegaal, H.M. Zonderland, H. van Tinteren, O. Visser, A.L. Verbeek, G.J. den Heeten, M.J. Broeders, 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. 22 (2012) 1717–1723, https://doi.org/10.1007/s00330-012-2409-2.
[16] J. Nederend, L.E. Duijm, M.W. Louwman, J.H. Groenewoud, A.B. Donkers-van Rossum, A.C. Voogd, Impact of transition from analog screening mammography to digital screening mammography on screening outcome in the Netherlands: a population-based study, Ann. Oncol. 23 (2012) 3098–3103, https://doi.org/10.1093/annonc/mds146.
[17] C.G. Glynn, D.M. Farria, B.S. Monsees, J.T. Salcman, K.N. Wiele, C.F. Hildebolt, Effect of transition to digital mammography on clinical outcomes, Radiology 260 (2011) 664–670, https://doi.org/10.1148/radiol.11110159.
[18] N.M. Hambly, M.M. McNicholas, N. Phelan, G.C. Hargaden, A. O'Doherty, F. L. Flanagan, Comparison of digital mammography and screen-film mammography in breast cancer screening: a review in the Irish breast screening program, AJR Am. J. Roentgenol. 193 (2009) 1010–1018, https://doi.org/10.2214/AJR.08.2157.
[19] N. Karssemeijer, A.M. Bluekens, D. Beijerinck, J.J. Deurenberg, M. Beekman, R. Visser, R. van Engen, A. Bartels-Kortland, M.J. Broeders, Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program, Radiology 253 (2009) 353–358, https://doi.org/10.1148/radiol.2532090225.
[20] V.L. Irvin, Z. Zhang, M.S. Simon, R.T. Chlebowski, S.W. Luoh, A.H. Shadyab, J. L. Krok-Schoen, F.K. Tabung, L. Qi, M.L. Stefanick, P. Schedin, S. Jindal, Comparison of mortality among participants of women's health initiative trials with screening-detected breast cancers vs interval breast cancers, JAMA Netw Open. 3 (2020), e207227, https://doi.org/10.1001/jamanetworkopen.2020.7227.
[21] N. Houssami, K. Hunter, The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening, NPJ Breast Cancer 3 (2017) 12, https://doi.org/10.1038/s41523-017-0014-x.
[22] Richtlijn Mammacarcinoom, Het Nationaal Borstkanker Overleg Nederland (NABON), 2012 (Accessed 23rd of May, 2020, at http://www.oncoline.nl/mammacarcinoom.).
[23] F. Sardanelli, H.S. Aase, M. Alvarez, E. Azavedo, H.J. Baarslag, C. Balleyguier, P. A. Baltzer, V. Beslagic, U. Bick, D. Bogdanovic-Stojanovic, R. Briediene,

B. Brkljacic, J. Camps Herrero, C. Colin, E. Cornford, J. Danes, G. de Geer, G. Esen, A. Evans, M.H. Fuchsjaeger, F.J. Gilbert, O. Graf, G. Hargaden, T.H. Helbich, S. H. Heywang-Kobrunner, V. Ivanov, A. Jonsson, C.K. Kuhl, E.C. Lisencu, E. Luczynska, R.M. Mann, J.C. Marques, L. Martincich, M. Mortier, M. Muller-Schimpfle, K. Ormandi, P. Panizza, F. Pediconi, R.M. Pijnappel, K. Pinker, T. Rissanen, N. Rotaru, G. Saguatti, T. Sella, J. Slobodnikova, M. Talk, P. Taourel, R.M. Trimboli, I. Vejborg, A. Vourtsis, G. Forrai, Position paper on screening for breast cancer by the European Society of Breast Imaging (EUSOBI) and 30 national breast radiology bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, The Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey, Eur. Radiol. 27 (2017) 2737–2743, https://doi.org/10.1007/s00330-016-4612-z.

[24] M. Arbyn, A. Anttila, J. Jordan, G. Ronco, U. Schenck, N. Segnan, H. Wiener, A. Herbert, L. von Karsa, European guidelines for quality assurance in cervical Cancer screening. Second edition–summary document, Ann. Oncol. 21 (2010) 448–458, https://doi.org/10.1093/annonc/mdp471.

[25] B. Lauby-Secretan, C. Scoccianti, D. Loomis, L. Benbrahim-Tallaa, V. Bouvard, F. Bianchini, K. Straif, International Agency for Research on Cancer Handbook Working G, Breast-cancer screening–viewpoint of the IARC working group, N. Engl. J. Med. 372 (2015) 2353–2358, https://doi.org/10.1056/NEJMsr1504363.

[26] N. Perry, M. Broeders, C. de Wolf, S. Tornberg, R. Holland, L. von Karsa, European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition–summary document, Ann. Oncol. 19 (2008) 614–622, https://doi.org/10.1093/annonc/mdm481.