EUROPEAN JOURNAL OF CANCER 45 (2009) 774–781

available at www.sciencedirect.com

journal homepage: www.ejconline.com

Delayed diagnosis of breast cancer in women recalled for suspicious screening mammography

Lucien E.M. Duijm[a,*], *Johanna H. Groenewoud*[b], *Harry J. de Koning*[c], *Jan Willem Coebergh*[c], *Mike van Beek*[d], *Marianne J.H.H. Hooijen*[e], *Lonneke V. van de Poll-Franse*[f]

[a]*Department of Radiology, Catharina Hospital, P.O. Box 1350, 5602 ZA, Eindhoven, The Netherlands*
[b]*Centre of Expertise Transitions in Care, University of Applied Sciences, P.O. Box 25035, 3001 HA, Rotterdam, The Netherlands*
[c]*Department of Public Health, Erasmus MC University Medical Centre, P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands*
[d]*Regional Laboratory for Pathology, PAMM Laboratories, P.O. Box 1350, 5602 ZA, Eindhoven, The Netherlands*
[e]*Department of Radiology, Saint Anna Hospital, P.O. Box 90, 5660 AB, Geldrop, The Netherlands*
[f]*Comprehensive Cancer Centre South (IKZ)/Eindhoven Cancer Registry, P.O. Box 231, 5600 AE, Eindhoven, The Netherlands*

ARTICLE INFO

Article history:
Received 22 June 2008
Received in revised form
19 October 2008
Accepted 24 October 2008
Available online 4 December 2008

Keywords:
Delay in diagnosis
False-negative assessment
Mammography
Screening

ABSTRACT

Purpose: To determine the frequency, pathology and causes of a delay in cancer diagnosis in women recalled for suspicious screening mammography.

Methods: We included all 290,943 screening mammograms of women aged 50–75 years, who underwent biennial screening mammography between 1st January 1995 and 1st January 2006. During a follow-up period of at least 2 years, clinical data, breast imaging reports, biopsy results and breast surgery reports were collected of all 3513 women with a positive screening result. Tumour stages of breast cancers with a diagnostic delay (defined as breast cancer confirmation more than 3 months following a positive mammography screen) were compared with those of cancers diagnosed within 3 months following referral and with interval cancers.

Results: A diagnostic delay occurred in 97 (6.5%) of 1503 screen-detected cancers. These 97 false-negative assessments comprised significantly more ductal cancers *in situ* (26.8%) than did cancers with an adequate assessment after recall (15.5%, $p = 0.004$) or interval cancers (3.7%, $p < 0.001$). Compared with interval cancers, cancers with a false-negative assessment had a more favourable tumour size (T1a–c, 87.3% versus T1a–c, 46.4%; $p < 0.001$) and showed significantly fewer cases with axillary lymph node metastases (22.5% versus 48.2%; $p < 0.001$). Between hospitals having performed the workup of at least 500 referred women each, the percentage of women with a false-negative assessment varied from 5.0% to 9.1% ($p = 0.03$). In these hospitals, improper classification of lesions at diagnostic mammography comprised 64.4% of false-negative assessments.

Conclusion: We found that 6.5% of recalled women experienced a delay in breast cancer diagnosis, with significant performance variations between hospitals.

* *Corresponding author:* Address: Department of Radiology, Catharina Hospital, Michelangelolaan 2, 5623 EJ, Eindhoven, The Netherlands. Tel.: +31 40 2398565; fax: +31 40 2398567.
E-mail address: LEMDuijm@hotmail.com (L.E.M. Duijm).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2008.10.020

1. Introduction

Population-based studies have shown that breast cancer screening with mammography is effective for the early detection of breast cancer and reduction of breast cancer mortality.[1–3] In the Netherlands, the nation-wide breast cancer screening programme provides biennial screening mammography for all women aged 50–75 years.[4] An early diagnosis of breast cancer not only requires the recognition of suspicious mammographic findings at screening, but also necessitates a prompt confirmation of the malignancy after recall. Some women, who are recalled for further evaluation of a screen-detected abnormality, experience a delay in breast cancer diagnosis due to an inadequate diagnostic assessment. This delay may result in worsening of the cancer stage, and may therefore have a negative impact on prognostic outcome. Several studies report on the diagnostic delay in breast cancer diagnosis in symptomatic patients,[5,6] but very limited data are available regarding asymptomatic, screened women. In a retrospective study, Ciatto et al. found a diagnostic delay of more than 3 months in 4.1% of cancers occurring in women recalled after a positive screen.[7]

In the current study, we prospectively identified women with a diagnostic delay after recall and compared the radiological and tumour characteristics of these malignancies with those of true interval cancers and screen-detected cancers without a diagnostic delay. Furthermore, we examined whether the proportion of women with a false-negative assessment changed over time, and we determined variations in diagnostic accuracy between hospitals, where the evaluation of recalled women was performed.

2. Material and methods

2.1. Screening procedure and referral

We included all 290,943 screening mammograms of women aged 50–75 years, who underwent screening mammography at one of the two specialised conventional mammography screening units (one fixed and one mobile) in the southern breast cancer screening region of the Netherlands (Bevolkings Onderzoek Borstkanker Zuid, BOBZ) between 1st January 1995, and 1st January 2006. Details of the Dutch nation-wide breast cancer screening programme have been described previously.[4] In brief, breast screening was gradually implemented in the BOBZ screening region between 1992 and 1995. The programme initially offered biennial screening mammography to women aged 50–69 years; in 1998 the upper age limit was extended to 75. In initial screens, i.e. the first time women are screened within the screening programme, two-view (medio-lateral-oblique and cranio-caudal) mammography of each breast is performed. Subsequent screens commonly include one-view mammography only (medio-lateral-oblique). Additional cranio-caudal views are obtained in about 50% of the subsequent screens. Indications for this two-view mammography include complicated judgement due to breast surgery or dense fibroglandular tissue; any changes in mammographical findings, such as new or increased microcalcifications or densities; and a longer-than-two-year interval since the previous screen. All screening mammograms are assessed independently by two certified screening radiologists. If the two radiologists disagree about referral, they usually discuss the mammogram together to reach consensus about referral. After consensus reading, women are referred for further diagnostics if at least one of the radiologists considers this necessary. From July 1998 till January 2001, discrepant readings were presented to a third radiologist or to a panel of radiologists to reach a decision about referral.[8] From 2003, the mammograms of women screened in the BOBZ region were independently read by two mammography screening technologists in addition to independent double reading by two screening radiologists.[9] In the BOBZ screening region, the mammographic findings are classified according to one of the five categories of abnormal findings: suspicious high density (e.g. spiculated density or density with indistinct borders), suspicious microcalcifications (e.g. pleomorphic, branching or amorphous/indistinct microcalcifications) and high density in combination with microcalcifications, architectural distortion or asymmetry.

2.2. Diagnostic workup

If further assessment of a mammographic abnormality is indicated, the woman is referred to a surgical oncologist of a general hospital or university hospital. After physical examination by the surgeon, two-view mammography (medio-lateral-oblique and cranio-caudal view) of each breast is obtained according to the Dutch guidelines on breast cancer screening and diagnosis[10]; local compression or magnification mammograms are performed if necessary. At diagnostic workup, radiologists classify the radiological findings according to the American College of Radiology BI-RADS.[11] Depending on the findings at physical examination and mammography and depending on the diagnostic workup protocols and hospital facilities available, further diagnostic evaluation may include breast ultrasonography, magnetic resonance mammography, percutaneous fine needle aspiration biopsy (FNAB, cytology) or core biopsy (usually image-guided) or open surgical biopsy. After having been referred for further evaluation, patients with benign breast imaging or biopsy results routinely undergo a first follow-up mammography at 6–12 months. Depending on the findings at follow-up mammography, a repeated mammogram at a later stage may be obtained to exclude malignancy.

2.3. Follow-up procedure

The follow-up period for all screened women included the time through the next screening round (the screening interval was approximately 2 years). For all women with a positive screening mammogram, we collected data on diagnostic procedures undertaken, breast cancer diagnosis, histopathology and TNM (tumour-node-metastases) classification to identify screen-detected cancers.[12] The date of a definitive diagnosis of malignancy was defined as the date of any percutaneous biopsy yielding a malignant lesion, or the date of open surgical biopsy if the diagnosis of breast cancer had not been established preoperatively. We excluded breast malignancies other than primary breast cancers from this analysis, and consid-

ered lobular carcinoma *in situ* to be a benign lesion. For women with a bilateral disease, the cancer with the highest stage was retained; multiple foci of cancer in one breast were counted as one cancer.

2.4. Detection of interval cancers

Most interval cancers were identified by linking the screening records to the regional cancer registry (Eindhoven Cancer Registry). Interval cancers are breast cancers that are diagnosed in women after a negative screening test (defined as no recommendation for referral). To trace interval cancers to the maximum, we also obtained the following information: (a) we received all radiotherapy reports from regional radiotherapy institutes of women who underwent radiation treatment for breast malignancy and who had participated in the screening programme; (b) we inquired about pathology specimens at the various regional pathology laboratories, some months after any hospital had requested the screening mammograms of a participant who had not been referred for further analysis; (c) we obtained the pathology records of women who had cancelled a call for subsequent screening because breast cancer had been diagnosed after a previous negative screen, and (d) we received occasional reports on interval cancers provided by general practitioners or medical specialists to the screening centre.

2.5. Delay in breast cancer diagnosis

We considered a definite diagnosis – e.g. confirmation or exclusion of malignancy – that was obtained within 3 months after a positive screen to be acceptable as a minor delay. A definite diagnosis of malignancy after 3 months was regarded as a serious delay with a risk of a worse prognosis. We determined the causes of these delays through review of all diagnostic procedures that had been performed. To find out whether a delay in breast cancer diagnosis could be attributed to the radiological assessment, two breast radiologists independently and retrospectively reviewed the diagnostic breast images of all women who had breast cancer that was pathologically confirmed in more than 3 months following a positive screen. Each reviewer classified the lesions according to the American College of Radiology BI-RADS.[11,13] To determine whether a delay in cancer diagnosis could be attributed to a false-negative pathology report, a pathologist reviewed the specimen of those women who had undergone more than one breast biopsy procedure needed for breast cancer confirmation. At review, both the radiologists and the pathologist knew that they reassessed cases characterised by a delay in cancer diagnosis, but they were blinded to the name of the hospital where the diagnostic procedures had been performed.

2.6. Workup facilities at hospitals

To assess the availability of diagnostic workup facilities and changes in workup procedures over time, one of the authors (L.D.) inquired at hospitals involved in the evaluation of screen-positive women, how many diagnostic mammograms they performed on a yearly basis, whether an out-patient breast clinic was present, whether the follow-up results of women with a positive screen were regularly discussed by a multidisciplinary team and when advanced diagnostic modalities such as magnetic resonance mammography and stereotactic core needle biopsy had been introduced.

2.7. Statistical analysis

The main outcome measures were the determination of the proportion of screen-positive women with a more than 3-month delay in breast cancer diagnosis, and determination of possible fluctuations in delay percentages through the years. Screening characteristics (initial screen versus subsequent screen; distribution of lesion type at screening) and tumour stages of cancers with a diagnostic delay were compared with those of cancers diagnosed within 3 months following referral and with interval cancers. Finally, we determined the reasons for a delay in breast cancer diagnosis. The chi-square test was used to test the differences in delay percentages between the groups for a statistical significance. The significance level was set at $p = 0.05$.

3. Results

3.1. Referral and diagnostic follow-up examinations

Of the 290,943 mammographic screening examinations, 66,925 (23.0%) were initial screens and 224,018 (77.0%) were subsequent screens. The mean age of the screened women was 59 years (age range 50–75 years). Altogether, 3513 (1.2%) women were referred for further diagnostic examination (Table 1). Of the 17 hospitals that were involved in the assessment of screen-positive women, 4 hospitals evaluated 500 women or more. The mammographic features of the 3513 screen-positive women were 2489 densities (70.9%), 621 microcalcification abnormalities (17.7%), 257 densities with microcalcifications (7.3%), 93 architectural distortions (2.6%) and 53 asymmetries of breast parenchyma (1.5%). Seven hundred and thirty two of the 1156 women referred at first screens (63.3%) and 1537 of the 2357 women referred at subsequent screens (65.2%) underwent percutaneous or open surgical biopsy.

3.2. Breast cancers diagnosed within 3 months after a positive screen

In total, 1503 histologically proven breast cancers were diagnosed, yielding an overall cancer detection rate of 5.2 per 1000 women screened and a true-positive referral rate of 42.8% (Table 1). Breast cancer was diagnosed within 3 months following a positive screen in 1406 women (93.5%).

3.3. False-negative assessments after a positive screen and interval cancers

A diagnostic delay in breast cancer diagnosis of more than 3 months occurred in 97 recalled women, comprising 6.5% of all screen-detected cancers (Table 1). These 97 cancers were diagnosed 4–6 months (23 cases), 7–12 months (42 cases) or 13–24 months (23 cases) after recall, or at the next biennial

Table 1 – Breast cancers detected by mammography screening from 1995 to 2006 in the BOBZ screening region, the Netherlands.

Year	Screens, No	Referral, No	Breast cancers, No	Cancer detection rate per 1000 women screened	Diagnostic delay > 3 months, No (%)
1995	15,008	200 (1.3)	85	5.7	8 (9.4)
1996	26,713	290 (1.1)	156	5.8	16 (10.3)
1997	24,154	293 (1.2)	109	4.5	8 (7.3)
1998	24,567	244 (1.0)	115	4.7	7 (6.1)
1999	26,535	214 (0.8)	120	4.5	6 (5.0)
2000	27,183	285 (1.0)	155	5.7	11 (7.1)
2001	25,642	289 (1.1)	124	4.8	9 (7.3)
2002	27,847	264 (0.9)	130	4.7	8 (6.2)
2003	29,278	469 (1.6)	161	5.5	5 (3.1)
2004	31,973	516 (1.6)	184	5.8	12 (6.5)
2005	32,868	449 (1.4)	164	5.0	7 (4.3)
Total	290,943	3513 (1.2)	1503	5.2	97 (6.5)

Percentages are given in parentheses.

screen (9 cases). The mean delay in diagnosis of the 97 cancers was 11 months (range, 4–28 months). The 6–12 months follow-up of 1615 women with benign or probably findings at initial workup resulted in a 2.6% cancer detection rate among these women (42/1615). Cancers characterised by a diagnostic delay of 13 months or more comprised fewer ductal carcinomas *in situ* (21.9%; $n = 7$) than breast malignancies diagnosed after a delay of 4–12 months (29.2%; $n = 19$; $p = 0.4$), fewer small invasive T1a–b tumours (15.6% ($n = 5$) versus. 36.9% ($n = 24$); $p = 0.3$) and more cancers with axillary lymph node metastases (25.0% ($n = 8$) versus. 12.3% ($n = 8$); $p = 0.1$). The overall proportion of diagnostic delays showed a non-significant decrease from 7.4% (50 of 672 screen-detected cancers) in the 1st 5.5-year screening period (1st January 1995–30th June 2000) to 5.7% (47 of 831 screen-detected cancers) in the 2nd 5.5-year screening period (from 1st July 2000; $p = 0.2$). There were no significant differences in the diagnostic delay according to the mammographic lesion types between the two periods. For DCIS tumours, however, the proportion of diagnostic delays significantly decreased from 15.6% (17 of 109) in the 1st 5.5-year period to 6.7% (9 of 135) in the 2nd 5.5-year period ($p = 0.02$).

The characteristics of all screen-detected cancers and interval cancers are presented in Table 2. Cancers with a diagnostic delay greater than 3 months due to the assessment failure more frequently comprised initial screens (37.1%) compared with cancers without a diagnostic delay (29.0%) or interval cancers (24.0%), but the difference was statistically significant for interval cancers only ($p = 0.007$). The proportion of ductal carcinoma *in situ* of 26.8% for cancers with a diagnostic delay was significantly greater than for cancers with an adequate assessment after recall (15.5%, $p = 0.004$) or interval cancers (3.7%, $p < 0.001$). Moreover, invasive cancers of the first group tended to have a more favourable tumour size (T1a–c, 87.3%) than invasive cancers diagnosed within 3 months after a positive screen (T1a–c, 78.5%; $p = 0.07$), or interval cancers (T1a–c, 46.4%; $p < 0.001$). The proportions of invasive tumours with positive axillary lymph nodes of screen-detected cancers, diagnosed with or without delay, were similar (22.5% versus 26.7%; $p = 0.4$). Compared with interval cancers, patients with a diagnostic delay showed significantly fewer invasive cancers with axillary lymph node metastases (22.5% versus 48.2%; $p < 0.001$).

3.4. *Differences in diagnostic workup accuracy between hospitals and reasons for false-negative assessment after recall*

The proportion of women with breast cancer and a diagnostic delay of more than 3 months varied from 3.7% (1 of 27 cancers) to 10.7% (6 of 56 cancers; $p = 0.3$) in the seven hospitals that each evaluated 50 referred women or more. For those four hospitals having performed the workup of at least 500 referred women each, the percentage of false-negative assessments ranged from 5.0% (hospital B, 20/401) to 9.1% (hospital C, 29/320; $p = 0.03$; Table 3). An incorrect assessment was observed more frequently for microcalcifications than for densities (10.5% (26/247) versus 5.4% (50/932); $p = 0.003$), with variations from 5.1% (2/39, hospital A) to 22.0% (13/59, hospital C; $p = 0.02$) for microcalcifications and from 3.6% (10/278, hospital B) to 7.0% (15/213, hospital C; $p = 0.09$) for densities. An incorrect BI-RADS classification, especially categorisation of suspicious lesions (BI-RADS 4) or malignant lesion (BI-RADS 5) as probably benign lesions (BI-RADS 3), was the major determinant of false-negative assessments in all four hospitals, comprising 45.0% (hospital B) to 75.9% (hospital C; $p = 0.03$) of the patients with a diagnostic delay (Table 4).

At review, the pathologist did not encounter any incorrectly classified biopsy reports. However, false-negative percutaneous biopsy results, i.e. retrieval of non-suspicious cells from a malignant lesion, accounted for 19.5% of diagnostic delays, varying from 6.9% (hospital C) to 35.0% (hospital B; $p = 0.01$).

3.5. *Workup facilities and changes in workup over time*

The four centrally located hospitals performed between 2000 and 3500 diagnostic mammographic examinations yearly. Out-patient breast clinics became available at these hospitals between 1999 and 2007, and a systematic discussion of positive screens by a multidisciplinary team of physicians had been established in the hospitals between 2002 and 2007.

Table 2 – Characteristics of screen-detected cancers and interval cancers.

	Screen-detected breast cancers			Interval cancers
	No diagnostic delay N = 1406	Diagnostic delay >3 months N = 97	*p*-Value	N = 537
Initial screen, No (%)	408 (29.0)	36 (37.1)	0.09	129 (24.0)
Subsequent screen, No (%)	998 (71.0)	61 (62.9)		408 (76.0)
Mean age at screening, years	62.0	61.2		59.5
Lesion type on screening mammography, No (%)				
Density	951 (67.6)	58 (59.8)	0.1	N/A
Microcalcifications	246 (17.5)	28 (28.9)	0.005	N/A
Density+microcalcifications	155 (11.0)	8 (8.2)	0.4	N/A
Architectural distortion	45 (3.2)	3 (3.1)	1.0	N/A
Asymmetry	9 (0.6)	0 (0)	0.4	N/A
Tumour stage (tumour size and nodal status), No (%)				
DCIS	218 (15.5)	26 (26.8)	0.004	20 (3.7)
Invasive	1188 (84.5)	71 (73.2)		517 (96.3)
T1a/b/c	932 (78.5)	62 (87.3)		240 (46.4)
T2+	249 (21.0)	9 (12.7)		261 (50.5)
Unknown	7 (0.6)	0 (0)		16 (3.1)
Lymph-node status of invasive cancers, No (%)				
Negative	840 (70.7)	53 (74.6)	0.5	247 (47.8)
Positive	317 (26.7)	16 (22.5)		249 (48.2)
Unknown	31 (2.6)	2 (2.8)		21 (4.1)

The *p*-value concerns statistical comparison between screen-detected cancers with and without diagnostic delay using the Chi-square test.

Table 3 – Variations in diagnostic delay in breast cancer diagnosis between hospitals.

	Hospital A		Hospital B		Hospital C		Hospital D	
Delay in breast cancer diagnosis	No	Yes	No	Yes	No	Yes	No	Yes
Characteristics of referred women								
Referral, No (%)	206 (94.9)	11 (5.1)	381 (95.0)	20 (5.0)	291 (90.9)	29 (9.1)	415 (93.9)	27 (6.1)
Initial screens, No (%)	63 (29.0)	3 (1.4)	103 (25.7)	6 (1.5)	95 (29.7)	14 (4.4)	89 (20.1)	7 (1.6)
Subsequent screens, No (%)	143 (65.9)	8 (3.7)	278 (69.3)	14 (3.5)	196 (61.3)	15 (4.7)	326 (73.8)	20 (4.5)
Mammographic abnormality, No (%)								
Density	141 (65.0)	7 (3.2)	268 (66.8)	10 (2.5)	198 (61.9)	15 (4.7)	275 (62.2)	18 (4.1)
Microcalcifications	37 (17.1)	2 (0.9)	60 (15.0)	4 (1.0)	46 (14.4)	13 (4.1)	78 (17.6)	7 (1.6)
Density + microcalcifications	22 (10.1)	2 (0.9)	42 (10.5)	3 (0.7)	35 (10.9)	1 (0.3)	43 (9.7)	2 (0.5)
Asymmetry	1 (0.5)	–	3 (0.7)	–	1 (0.3)	–	2 (0.5)	–
Architectural distortion	5 (2.3)	–	8 (2.0)	3 (0.7)	11 (3.4)	–	17 (3.8)	–
Mean age at referral, years	61.6	61.4	62.3	61.1	62.2	61.9	62.0	61.0
Type of breast cancer, No (%)								
DCIS	33 (15.2)	1 (0.5)	50 (12.5)	4 (1.0)	42 (13.1)	12 (3.8)	76 (17.2)	5 (1.1)
Invasive	173 (79.7)	10 (4.6)	331 (82.5)	16 (4.0)	249 (77.8)	17 (5.3)	339 (76.7)	22 (5.0)
T1a/b/c	148 (80.9)	10 (5.5)	278 (80.1)	15 (4.3)	191 (71.8)	13 (4.9)	249 (69.0)	21 (5.8)
T2+	24 (13.1)	–	49 (14.1)	1 (0.3)	58 (21.8)	4 (1.5)	88 (24.4)	1 (0.3)
Unknown	1 (0.5)	–	4 (1.2)	–	–	–	2 (0.6)	–
Lymph–node status of invasive cancers, No (%)								
Positive	44 (24.0)	–	76 (21.9)	5 (1.4)	69 (25.9)	6 (2.3)	113 (31.3)	3 (0.8)
Negative	126 (68.9)	10 (5.5)	245 (70.6)	11 (3.2)	175 (65.8)	11 (4.1)	216 (59.8)	18 (5.0)
Unknown	3 (1.6)	–	10 (2.9)	–	5 (1.9)	–	10 (2.8)	1 (0.3)

Percentages are given in parentheses
Percentages in the first row (referral) are horizontal percentages; all other percentages are vertical percentages.

Magnetic resonance mammography, stereotactic core needle biopsy (14 Gauge) and axillary ultrasound with lymph node sampling were introduced, respectively, between 2000–2004, 2000–2007 and 1998–2000. Stereotactic vacuum-assisted core biopsy (either 9 or 10 Gauge) became available between 2004 and 2007 at the four sites, and is currently used as the stan-

Table 4 – Reasons of false–negative assessment (longer-than-3-month diagnostic delay) of 87 cancers at four hospitals.

	Hospital				
	A	B	C	D	Total
Reason of false-negative assessment at workup					
Incorrect BI-RADS classification, No	8 (72.7)	9 (45.0)	22 (75.9)	17 (63.0)	56 (64.4)
– Lesion classified probably benign (BI-RADS 3) at workup, but 'suspicious' (BI-RADS 4) or 'malignant' (BI-RADS 5) in retrospect	5 (45.4)	4 (20.0)	17 (58.6)	13 (48.1)	39 (44.8)
– Lesion classified 'normal' (BI-RADS 1) or 'benign' (BI-RADS 2) at workup, but BI-RADS 4 or BI-RADS 5 in retrospect	3 (27.3)	5 (25.0)	5 (17.2)	4 (14.8)	17 (19.5)
Correct BI-RADS 3 classification in retrospect	–	1 (5.0)	–	–	1 (1.1)
Benign open (surgical) biopsy from BI-RADS 4 or BI-RADS 5 lesion	–	1 (5.0)	3 (10.3)	1 (3.7)	5 (5.7)
Benign percutaneous biopsy from BI-RADS 4 or BI-RADS 5 lesion	2 (18.2)	7 (35.0)	2 (6.9)	6 (22.2)	17 (19.5)
– Fine needle aspiration biopsy	1 (9.1)	1 (5.0)	1 (3.5)	–	3 (3.4)
– Core biopsy	1 (9.1)	4 (20.0)	–	2 (7.4)	7 (8.0)
– Stereotactic core needle biopsy	–	2 (10.0)	1 (3.5)	4 (14.8)	7 (8.0)
Combination of benign percutaneous biopsy + benign open (surgical) biopsy from BI-RADS 4 or BI-RADS 5 lesion	1 (9.0)	–	–	2 (7.4)	3 (3.4)
Surgical oncologist did not comply with the pathologist's advise to excise a probably malignant lesion at percutaneous biopsy	–	2 (10.0)	2 (6.9)	1 (3.7)	5 (5.7)

Percentages are given in parentheses.

dard biopsy technique for clustered microcalcifications. Through the years, one hospital (hospital C) mainly performed ultrasound-guided fine needle aspiration cytology of solid breast lesions, whereas the other three hospitals gradually replaced cytology by core biopsy.

4. Discussion

Delay in the diagnosis of breast cancer is not infrequent, and is the most common clinical scenario resulting in malpractice in the United States.[14] In symptomatic women, the leading cause of physician delay is the inappropriate reassurance that a palpable mass is benign, without radiological evaluation or biopsy being performed.[15] Although Barber and colleagues found that only 1.4% of symptomatic patients in a British breast clinic experienced a diagnostic delay exceeding 2 months,[16] other studies report higher delay percentages in the range of 4–39%.[5,6,17] Very few studies report on the delay in diagnosis in women recalled from screening (positive screen), but subsequently misdiagnosed (false negative) at the assessment. The delay percentage of 6.5% in our series is somewhat higher than the 4.1% and 5.2% of delays reported in two retrospective studies.[7,18] This difference may partly be explained by the prospective nature of our study. Warren and colleagues observed a decrease in the proportion of inadequately assessed women through the years.[18] Although the highest percentage of false-negative assessments in our study occurred in the initial two years of inclusion, the proportion of diagnostic delays stabilised later on. Despite the introduction of new diagnostic modalities such as magnetic resonance mammography and the increased use of core biopsy instead of cytology, we found no significant decrease in the percentage of false-negative assessments in the second screening period (from July 2000) as compared with the first screening period. Studies suggest that the establishment of breast care units and multidisciplinary teams improves the assessment of symptomatic breast disease.[19,20] As these modalities have only recently become fully available at the hospitals involved in the workup of screen-positive women, we were not able to determine a possible benefit of these developments on diagnostic accuracy.

Contrary to the observations of Ciatto and et al.,[7] the tumour stages in women who experienced a more than 3-month diagnostic delay in our study were more favourable than those of cancers diagnosed within 3 months of a positive screen. However, the mean diagnostic delay in their series was almost as twice as long as ours. Our finding that women with screen-detected breast cancers diagnosed between 13 and 28 months after mammographic screening had an increased risk of larger tumours and axillary lymph node metastases, compared with women in whom breast cancer was diagnosed within 4–12 months of an abnormal screen, is in line with the results of a Canadian study.[21] Isolated microcalcification abnormalities were more likely to have been inadequately assessed than other types of lesions, a finding which is consistent with the results of a UK study.[18]

The tumour sizes and axillary lymph node stages were much better for cancers with a diagnostic delay as compared with true interval cancers. Moreover, the number of interval cancers by far exceeded that of women experiencing a delay in cancer diagnosis after recall. Efforts at improving the impact of screening should thus be especially directed at the reduction of interval cancers.

Conflicting reports are available as to whether a delayed diagnosis is associated with a lower survival in symptomatic patients. Afzelius and colleagues found that physician delay of more than 60 days was not associated with an unfavourable outcome,[22] but a systematic review study indicated that a delay of 3–6 months was associated with lower survival.[23] Allgood and colleagues recently estimated that unsatisfactory assessments of recalled women in a screened population may increase breast cancer mortality within this special tumour population by 14–18%.[24]

We found a worrisome and considerable variation in the workup performance across hospitals of recalled women, with a proportion of false-negative assessments ranging from 3.7% to 10.7%. An improper use of the BI-RADS 3 category at diagnostic mammography was the major determinant of diagnostic delay in the four hospitals that performed the majority of workup procedures. Studies have consistently found that no more than 2% of lesions characterised as 'probably benign' actually turn out to be cancerous, but only when strict diagnostic criteria are applied.[25,26] For this type of lesions, periodic mammographic surveillance is indicated as an alternative to open surgical biopsy or percutaneous imaging-guided tissue sampling. Our findings suggest that the use of the probably benign category was not always based on well-studied, documented criteria, but on the individual radiologist's instinct.

We recently started to give radiologists and surgical oncologists feedback on their workup performance of recalled women, and we hope that this quality assurance procedure will decrease the number of women experiencing a delay in breast cancer diagnosis.

We were informed about the annual number of diagnostic mammograms performed at the four centrally located hospitals in our study. We did not have this information from the other hospitals, but it is likely that these hospitals, located at the periphery of our screening region or even at a further distance, assessed women from other screening regions as well. In addition, the hospitals are not only involved in the diagnostic workup of (asymptomatic) women with positive screening mammograms, but are also involved in the diagnostic workup of symptomatic women, including other age groups as well. Limited experience in mammography reading may lead to a higher number of false-negative assessments, but whether this was the case in our study should be addressed in further research. If limited experience accounts for higher numbers of false-negative assessments, centralisation of recalls in a high performing centre might improve the outcome.

The decreased diagnostic delay in DCIS tumours over time may suggest that a wait-and-see approach with follow-up mammography was preferred to open surgical biopsy in the first years of screening, when stereotactic core biopsy of microcalcifications was not available.[27] The numbers in our study, however, are too small to provide further statistical evidence in support of this hypothesis.

In summary, we found that an overall 6.5% of women experienced a diagnostic delay in breast cancer diagnosis after recall, with large variations between the hospitals. This percentage only slightly decreased in the recent years, and our results stress that a strict adherence to the established mammographic criteria for lesion categorisation, especially for microcalcifications, is needed to decrease the number of women with a false-negative assessment. Although cancers with a diagnostic delay show a more favourable tumour stage than cancers without a diagnostic delay or interval cancers, a delay in breast cancer diagnosis may impact long-term outcomes.

Conflict of interest statement

None declared.

Funding

None.

REFERENCES

1. Otto SJ, Fracheboud J, Looman CWN, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;**361**:1411–7.
2. Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* 2003;**361**:1405–10.
3. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *New Engl J Med* 2005;**353**:1784–92.
4. Fracheboud J, de Koning HJ, Boer R, et al. Nationwide breast cancer screening programme fully implemented in the Netherlands. *Breast* 2001;**10**:6–11.
5. Jenner DC, Middleton A, Webb WM, Bates T. In-hospital delay in the diagnosis of breast cancer. *Brit J Surg* 2000;**87**:914–9.
6. Goodson WH, Moore DH. Causes of physician delay in the diagnosis of breast cancer. *Arch Intern Med* 2002;**162**:1343–8.
7. Ciatto S, Houssami N, Ambrogetti D, Bonardi R, Collini G, Rosselli Del Turco M. 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* 2007;**105**:37–43.
8. Duijm LEM, Groenewoud JH, Hendriks JHCL, de Koning HJ. Independent double reading of screening mammograms in the Netherlands: effect of arbitration following reader disagreements. *Radiology* 2004;**231**:564–70.
9. Duijm LEM, Groenewoud JH, Fracheboud J, de Koning HJ. Additional double reading of screening mammograms by radiologic technologists: impact on screening performance parameters. *J Natl Cancer Inst* 2007;**99**:1162–70.
10. Kwaliteitsinstituut voor de Gezondheidszorg CBO. *Mammacarcinoom: screening en diagnostiek*. Alphen aan den Rijn: Van Zuiden Communications B.V.; 2000.
11. American College of Radiology. *Breast imaging reporting and data system (BI-RADS)*. 2nd ed. Reston, VA: American College of Radiology; 1995.
12. UICC (International Union against Cancer). *TNM classification of malignant tumours*. 4th rev ed. Berlin: Springer Verlag; 1987.
13. American College of Radiology. *Breast imaging reporting and data system (BI-RADS)*. 4th ed. Reston, VA: American College of Radiology; 2003.
14. Physicians Insurers Association of America. *Breast cancer report*. Rockville, MD: Physicians Insurers Association of America; 2003.
15. Tartter PI, Pace D, Frost M, Bernstein JL. Delay in diagnosis of breast cancer. *Ann Surg* 1999;**229**:91–6.
16. Barber MD, Jack W, Dixon JM. Diagnostic delay in breast cancer. *Brit J Surg* 2004;**91**:49–53.
17. Montella M, D'Aiuto D, de Marco M, et al. Determinant factors for diagnostic delay in operable breast cancer patients. *Eur J Cancer Prev* 2001;**10**:53–9.
18. Warren R, Allgood P, Hunnam G, Godward S, Duffy S. East Anglian Breast Screening Programme. An audit of assessment procedures in women who develop breast cancer after a negative result. *J Med Screen* 2004;**11**:180–6.
19. Purushotham AD, Pain SJ, Miles D, Harnett A. Variations in treatment and survival in breast cancer. *Lancet Oncol* 2001;**2**:719–25.

20. Haward R, Amir Z, Borrill C, et al. Breast cancer teams: the impact of constitution, new cancer workload, and methods of operation on their effectiveness. *Brit J Cancer* 2003;**89**:15–22.
21. Olivotto IA, Gomi A, Bancej C, et al. Influence of delay to diagnosis on prognostic indicators of screen-detected breast carcinoma. *Cancer* 2002;**94**:2143–50.
22. Afzelius P, Zedeler K, Sommer H, Mouridsen HT, Blichert-Toft M. Patient's and doctor's delay in primary breast cancer. Prognostic implications. *Acta Oncol* 1994;**33**:345–51.
23. Richards MA, Westcombe AM, Love SB, Littlejohns P, Ramirez AJ. Influence of delay on survival in patients with breast cancer: a systematic review. *Lancet* 1999;**353**:1119–26.
24. Allgood PC, Duffy SW, Warren R, Hunman G. Audit of negative assessments in a breast-screening programme in women who later develop breast cancer – implications for survival. *Breast* 2006;**15**:503–9.
25. Vizcaíno I, Gadea L, Andreo L, et al. Short-term follow-up results in 795 nonpalpable probably benign lesions detected at screening mammography. *Radiology* 2001;**219**:475–83.
26. Yasmeen S, Romano PS, Pettinger M, et al. Frequency and predictive value of a mammographic recommendation for short-interval follow-up. *J Natl Cancer Inst* 2003;**95**:429–36.
27. Duijm LEM, Groenewoud JH, Roumen RMH, de Koning HJ, Plaisier ML, Fracheboud J. A decade of breast cancer screening in The Netherlands: trends in the preoperative diagnosis of breast cancer. *Breast Cancer Res Treat* 2007;**106**:113–9.