

Keywords: breast cancer; screen-detected; interval cancer; DCIS; program sensitivity; program specificity; tumour characteristics

Digital vs screen-film mammography in population-based breast cancer screening: performance indicators and tumour characteristics of screen-detected and interval cancers

Linda de Munck^{*,1}, Geertruida H de Bock², Renée Otter^{1,3}, Dick Reiding⁴, Mireille JM Broeders^{5,6}, Pax HB Willemse³ and Sabine Siesling^{1,7}

¹Department of Research, Netherlands Comprehensive Cancer Organisation, PO Box 19079, Utrecht 3501 DB, The Netherlands;

²Department of Epidemiology, University of Groningen, University Medical Center Groningen, PO Box 30001, Groningen 9700 RB, The Netherlands;

³Department of Medical Oncology, University of Groningen, University Medical Center Groningen, PO Box 30001, Groningen 9700 RB, The Netherlands;

⁴National Cancer Screening Programme, region North, PO Box 425, Groningen 9700 AK, The Netherlands;

⁵Radboud Institute for Health Sciences, Radboud University Medical Center, PO Box 9101, Nijmegen 6500 HB, The Netherlands;

⁶Dutch Reference Centre for Screening, PO Box 6873, Nijmegen 6503 GJ, The Netherlands and

⁷Department of Health Technology & Services Research, University of Twente, PO Box 217, Enschede 7500 AE, The Netherlands

Background: Full-field digital mammography (FFDM) has replaced screen-film mammography (SFM) in most breast cancer screening programs due to technological advantages such as possibilities to adjust contrast, better image quality and transfer capabilities. This study describes the performance indicators during the transition from SFM to FFDM and the characteristics of screen-detected and interval cancers.

Methods: Data of the Dutch breast cancer screening program, region North from 2004 to 2010 were linked to The Netherlands Cancer Registry ($N=902\,868$). Performance indicators and tumour characteristics of screen-detected and interval cancers were compared between FFDM and SFM.

Results: After initial screens, recall rates were 2.1% (SFM) and 3.0% (FFDM; $P<0.001$). The positive predictive values (PPV) were 25.6% (SFM) and 19.9% (FFDM; $P=0.002$). Detection rates were similar, as were all performance indicators after subsequent screens. Similar percentages of low-grade ductal carcinoma *in situ* (DCIS) were found for SFM and FFDM. Invasive cancers diagnosed after subsequent screens with FFDM were more often of high-grade ($P=0.024$) and ductal type ($P=0.030$). The incidence rates of interval cancers were similar for SFM and FFDM after initial (2.69/1000 vs 2.51/1000; $P=0.787$) and subsequent screens (2.30 vs 2.41; $P=0.652$), with similar tumour characteristics.

Conclusions: FFDM resulted in similar rates of screen-detected and interval cancers, indicating that FFDM performs as well as SFM in a breast cancer screening program. No signs of an increase in low-grade DCIS (which might connote possible overdiagnosis) were seen. Nonetheless, after initial screening, which accounts for 12% of all screens, FFDM resulted in higher recall rate and lower PPV that requires attention.

*Correspondence: L de Munck; E-mail: l.demunck@iknl.nl

Received 10 February 2016; revised 13 May 2016; accepted 3 July 2016; published online 4 August 2016

© 2016 Cancer Research UK. All rights reserved 0007–0920/16

In recent years, screen-film mammography (SFM) has been replaced by full-field digital mammography (FFDM) in most Western breast cancer screening programs. The main reasons are a better image quality and dose optimisation, and easier and more efficient archiving and image transfer capabilities (Obenauer *et al*, 2002; Pisano and Yaffe, 2005).

Studies have shown that FFDM does not result in differences in cancer detection rate (Vinnicombe *et al*, 2009; Juel *et al*, 2010; Domingo *et al*, 2011) or higher overall breast cancer detection compared with SFM in the clinical as well as in a population-based setting (Lewin *et al*, 2001; Pisano *et al*, 2005; Del Turco *et al*, 2007; Glynn *et al*, 2011). However, others have shown a higher detection rate only for ductal carcinoma *in situ* (DCIS) (Vigeland *et al*, 2008; Karssemeijer *et al*, 2009). Furthermore, FFDM results in a higher recall rate, in combination with higher detection rates and decreased positive predictive values (PPV) of recall (Skaane *et al*, 2007; Bluekens *et al*, 2012). Nevertheless, the opposite effect on recall rate and PPV has also been described (Hofvind *et al*, 2014). The recall and detection rates for SFM vary between different screening programs. These differences in 'baseline' measurements might partly explain the different outcomes for the introduction of FFDM when comparing with SFM.

FFDM seems associated with an increase in smaller cancers and DCIS (Del Turco *et al*, 2007; Vigeland *et al*, 2008; Bluekens *et al*, 2012; van Luijt *et al*, 2013). It remains uncertain to which extent this might be associated with overdiagnosis and a wide range of overdiagnosis estimates exist (Jørgensen and Gøtzsche, 2009). Furthermore, some studies comparing FFDM with SFM have analysed tumour characteristics other than DCIS and report no differences in tumour characteristics such as T- or N-stage (Vigeland *et al*, 2008; Juel *et al*, 2010; Domingo *et al*, 2011; Bluekens *et al*, 2012). However, one study found that FFDM detected more lymph node-negative and hormone receptor-positive cancers (Nederend *et al*, 2012). Studies about the incidence of interval cancers so far report no differences in interval cancer rates between SFM and FFDM (Juel *et al*, 2010; Hofvind *et al*, 2014). As the recall rate in The Netherlands is lower than in other countries with population-based screening programs, the transition from SFM to FFDM might lead to different conclusions regarding interval cancer rates. However, data regarding interval cancers in the Dutch situation, with a national screening program, is still scarce (Nederend *et al*, 2013).

The objective of this study was to evaluate whether the transition from SFM to FFDM resulted in changes in the recall rate, PPV, the detection rate of DCIS and invasive cancers and tumour characteristics of screen-detected cancers in the Dutch screening program. Moreover, the rate and tumour characteristics of interval cancers were compared, as were program sensitivity and specificity.

MATERIALS AND METHODS

Setting. In The Netherlands, a national population-based screening program is operational since 1990, inviting women 50–69 years of age for a biennial screening examination (Fracheboud *et al*, 2001). Women 70–75 years of age are included from 1999 onwards. All mammographic examinations are performed by specialised radiographers. Recall for further diagnostic work-up was indicated in case of incomplete examination (BI-RADS 0) or suspicious or malignant findings (BI-RADS 4 or 5) (D'Orsi *et al*, 2003; Timmers *et al*, 2012). The screening program started in 1990 using screen-film mammography. In 2003 the first digital mammography unit was introduced and in 2010 all screening examinations were performed using digital mammography. By participating in the screening program, women consent to the use

of their data to evaluate the program. Information about the use of their data is provided in a flyer accompanying the letter of invitation. If a woman does not want her data being used for this purpose, she can return a signed form to the screening organisation. Only a minor fraction of screened women (0.01%) uses this option (Fracheboud *et al*, 2014).

Data on cancer is registered in The Netherlands Cancer Registry (NCR). The NCR is a population-based database, which contains data on patient, tumour and treatment characteristics of all *in situ* and invasive malignancies diagnosed in The Netherlands. This study was approved by the Committee of Privacy of The Netherlands Comprehensive Cancer Organisation (IKNL), which hosts the NCR.

Women and data. All women (50–75 years) who underwent a screening examination between 2004 and 2010 in the Dutch Breast Cancer Screening Program, region North-Netherlands were included. The target population of this area comprised ~300 000 women, of whom about 150 000 were invited each year. The region North-Netherlands followed the national program. It started in 1991 with SFM and implemented the first digital mammography in 2003. During 2009, the remaining seven screen-film mammography units were replaced by digital mammography units. The screening data included all dates and results of the screening examinations. These screening data were linked to data of the NCR. The main source of notification for the NCR is PALGA, the nationwide Dutch network and registry of histo and cytological pathology. After notification to the NCR, specially trained registration clerks visit the hospitals to collect information on patient and tumour characteristics, stage and treatment from patient files. Coding of the items is based on international coding rules. For the staging of cancers, TNM classification is used (Wittekind *et al*, 2004).

Image acquisition and interpretation. SFM images were acquired with two types of systems (GE 600/800T, GE Healthcare, Buc, France). Digital mammograms were acquired by using Embrace DR (Agfa-Gevaert, Mortsel, Belgium) or Lorad Selenia FFDM systems (Hologic, Danbury, CT, USA). Initial screening examinations performed with SFM or FFDM always included two standard views, craniocaudal and mediolateral oblique. At subsequent screening examinations, mediolateral oblique views were routinely acquired, craniocaudal views were acquired if indicated by criteria based on, among others, breast density and visible abnormality. Radiologists interpreted mammograms made by SFM as well as FFDM. Mammograms were interpreted in batch mode and were independently read by two radiologists. Differences in opinion were resolved by consensus. In subsequent screens with SFM as well as FFDM, prior mammograms were available for comparison. To facilitate soft-copy reading of subsequent screening examinations with FFDM, the most recent screen-film mediolateral oblique and craniocaudal views were digitised by using film scanners and archiver designed for mammography (IMPAX TS5 transmit scanner, Agfa-Gevaert; DigitalNow, Hologic).

Definitions. Women with a 'screen-detected cancer' were defined as women having a positive screening examination at the breast cancer screening program, which was subsequently followed by a diagnosis of breast cancer within 12 months. The second group comprised women with an 'interval cancer', that is, women who, despite a negative examination at the breast cancer screening, were diagnosed with breast cancer within 24 months. In the analysis, women with screen-detected breast cancer and women with interval cancer were considered separately. Recall rate was defined as the proportion of women with a positive screening examination out of all screened women. Detection rate was defined as the proportion of women diagnosed with breast cancer out of all screened women and PPV was calculated as the percentage of

women with screen-detected breast cancer out of all recalled women. All screening examinations from 2004 to 2010 were considered for the calculation of recall rate, detection rate and PPV ($n = 902\,868$). Program sensitivity was defined as the proportion of screen-detected breast cancers out of all breast cancers (interval cancers plus screen-detected cancers). Program specificity was defined as the proportion of women without breast cancer with a negative examination (true negatives) out of all women without breast cancer (true negatives plus false positives). Program sensitivity, program specificity and the incidence of interval cancers were calculated for women screened from 2004 to 2008. As follow up was complete until 2010, the complete number of breast cancer cases, including interval cancers can be diagnosed up to 24 months after the last screening examination ($n = 638\,663$). For women with bilateral disease, the tumour with the highest stage was selected. Tumour stage was defined according to the greatest dimension of the largest tumour size (T1a: ≤ 0.5 cm, T1b: > 0.5 cm and ≤ 1 cm, T1c: > 1 cm and ≤ 2 cm, T2+: > 2 cm).

Statistical analysis. SFM and FFDM were compared regarding the performance indicators (recall rate, detection rate and PPV), tumour characteristics of screen-detected cancers, incidence rate and tumour characteristics of interval cancers, program sensitivity and program specificity. The comparison was stratified by initial and subsequent screening examinations. Pearson χ^2 -tests or Fisher exact tests were applied. Missing values in tumour characteristics were excluded in the statistical comparison. Variation in recall and detection rate per year was graphically visualised for SFM and FFDM, stratified by initial and subsequent screening examinations. The statistical significance level was set at a P -value < 0.05 . Analyses were performed using the STATA software package, version 13.1 for Windows (Stata Corporation LP, College Station, TX, USA).

RESULTS

In the period 2004–2010 in total 902 868 screening examinations were performed by the Dutch Breast Cancer Screening Program, region North-Netherlands, of which 255 633 (28%) by FFDM. Twelve percent of the 647.235 examinations by SFM were initial examinations; for FFDM, this was 11% (Table 1).

Recall rate, detection rate and positive predictive value. The recall rate after initial examination of 2.1% for SFM remained stable during the study period (Table 1). For FFDM the recall rate of 3.0% proved to be significantly higher than for SFM ($P < 0.001$). After the complete introduction of FFDM, the recall rate decreased to 2.8% in 2010 and was still significantly higher than the recall rate of 2.1% for SFM ($P < 0.001$) (Figure 1A).

After initial examination, the detection rate for DCIS was 0.86 per 1000 women for SFM compared with 1.18 per 1000 women for FFDM ($P = 0.137$). For invasive cancers, detection rates were 4.42 per 1000 women for SFM and 4.83 per 1000 women for FFDM ($P = 0.338$). During the total period, the overall detection rate for SFM remained fairly constant. For FFDM the detection rate fluctuated markedly, due to small numbers (Figure 1B). The PPV after initial examination was 25.6% for SFM and 19.9% for FFDM ($P = 0.002$).

After subsequent examinations, the overall recall rate for SFM was 1.2% and for FFDM, after an initially high rate, 1.1% ($P = 0.532$; Table 1; Figure 1C.) The overall detection rates were 5.28 per 1000 women for SFM and 5.14 per 1000 women for FFDM ($P = 0.638$; Table 1). No differences were found in detection rates for DCIS (0.74 per 1000 women for SFM vs 0.81 per 1000 women for FFDM; $P = 0.298$), nor for invasive cancers (4.54 per 1000 women for SFM vs 4.33 per 1000 women for FFDM; $P = 0.210$). The detection rate remained stable over the years for SFM as well as FFDM (Figure 1D). Positive predictive values were 45.7 and 45.2% for SFM and FFDM, respectively ($P = 0.638$).

Tumour characteristics of screen-detected cancers. After initial examinations, no differences were found in tumour characteristics (Table 2), but a non-significant trend towards more T1a cancers was seen (4.7% for SFM and 10.6% after FFDM, $P = 0.138$). In 16.3% of all cancers a DCIS was diagnosed after SFM, compared with 19.6% after FFDM ($P = 0.335$). After subsequent examinations, DCIS was diagnosed in 14.0% of all cancers after SFM, 15.8% after FFDM ($P = 0.139$). After FFDM a significantly larger part of invasive cancers were ductal cancers (SFM: 81.6%, FFDM: 85.1%; $P = 0.030$) and the cancers were more often high-grade (SFM: 9.0%, FFDM: 23.4%; $P = 0.024$).

Incidence rate of interval cancers, program sensitivity and specificity. Overall incidence rates of interval cancers were similar with 2.35 per 1000 women for SFM and 2.42 per 1000 women for FFDM ($P = 0.748$). After initial screens, no difference was found in the incidence of interval cancers with 2.69 per 1000 women for SFM and 2.51 per 1000 women for FFDM ($P = 0.787$, Table 3). The sensitivity after initial examinations was 66.1% for SFM and 69.1% for FFDM ($P = 0.657$), specificity was 98.5% and 96.9% for SFM and FFDM, respectively ($P < 0.001$). After subsequent examinations, the overall incidence rate of all interval cancers was 2.30 per 1000 women for SFM and 2.41 per 1000 women for FFDM ($P = 0.652$), the sensitivity was 69.7% for SFM and 66.7% for FFDM ($P = 0.232$). Specificity was 99.4% SFM and 99.2% for FFDM ($P < 0.001$).

Tumour characteristics of interval cancers. No differences were found in the tumour characteristics of interval cancers diagnosed after FFDM compared with SFM after initial or after subsequent examinations (Table 4).

DISCUSSION

After initial screens with FFDM, we found a higher recall rate and a lower PPV. No difference was found in the proportion and malignancy grade of DCIS. Tumour characteristics of invasive cancers and the incidence rate and tumour characteristics of interval cancers were similar. After subsequent screens the proportion of DCIS was similar for SFM and FFDM. After FFDM, a significantly larger part of invasive cancers were ductal cancers and high-grade. No differences were found in the incidence of interval cancers or their tumour characteristics between SFM and FFDM.

In our study a higher recall rate and lower PPV were found after initial screens with FFDM, which accounts for 12% of all screens, but no differences were found after subsequent screens. A systematic review by Vinnicombe *et al* (2009) showed that recall rates varied greatly with some studies showing lower and others showing higher recall rates for FFDM. We did not find a higher detection rate, which is in line with some other studies (Del Turco *et al*, 2007; Vinnicombe *et al*, 2009; Juel *et al*, 2010; Domingo *et al*, 2011). In contrast, other studies reported higher detection rates for FFDM (Glynn *et al*, 2011; van Luijt *et al*, 2013), including a Dutch study. The latter study also describes some variation in detection rates between regions within The Netherlands. As our study only included data from one region and the study could not interfere with the ongoing screening program, existing variation between regions and countries before the introduction of FFDM might explain the different results and might have an effect on the generalisability of our results. In the first years using FFDM, recall and detection rates fluctuated. This might be due to a learning curve, which has been described before (Bluekens *et al*, 2010), in combination with the relative low number of FFDM screens performed in comparison with SFM. As SFM was already in use since the start of the program, recall and detection rates remained stable during this study period as expected. The overall increase in recall rate after FFDM is probably due to a higher resolution and the possibilities

Table 1. Performance indicators of women (50–75 years) screened, recalled and positive predictive value for SFM and FFDM (2004–2010)

2004–2010	SFM			FFDM			P-value	
	N		%	N		%		
Screened women	647 235			255 633				
Initial screens	76 642		11.8	27 947		10.9		
Subsequent screens	570 593		88.2	227 686		89.1		
Recalled women	8174		1.26	3436		1.34	0.002	
Initial screens	1584		2.07	844		3.02	<0.001	
Subsequent screens	6590		1.15	2592		1.14	0.532	
Screen-detected	N	Rate^a	PPV	N	Rate^a	PPV	P-value^b	P-value^c
Breast cancers	3418	5.28	41.8	1339	5.24	39.0	0.800	0.004
Initial screens	405	5.28	25.6	168	6.01	19.9	0.159	0.002
Subsequent screens	3013	5.28	45.7	1171	5.14	45.2	0.445	0.638
DCIS	489	0.76	6.0	218	0.85	6.3	0.137	0.456
Initial screens	66	0.86	4.2	33	1.18	3.9	0.137	0.761
Subsequent screens	423	0.74	6.4	185	0.81	7.1	0.298	0.213
Invasive cancers	2929	4.53	35.8	1121	4.39	32.6	0.369	0.001
Initial screens	339	4.42	21.4	135	4.83	16.0	0.385	0.001
Subsequent screens	2590	4.54	39.3	986	4.33	38.0	0.208	0.264

Abbreviations: DCIS = ductal carcinoma *in situ*; FFDM = full-field digital mammography; PPV = positive predictive values; SFM = screen-film mammography.

^aDetection rates per 1000 women screened.

^bEstimated p-values are based on comparison between rate.

^cEstimated p-values are based on comparison between PPV.

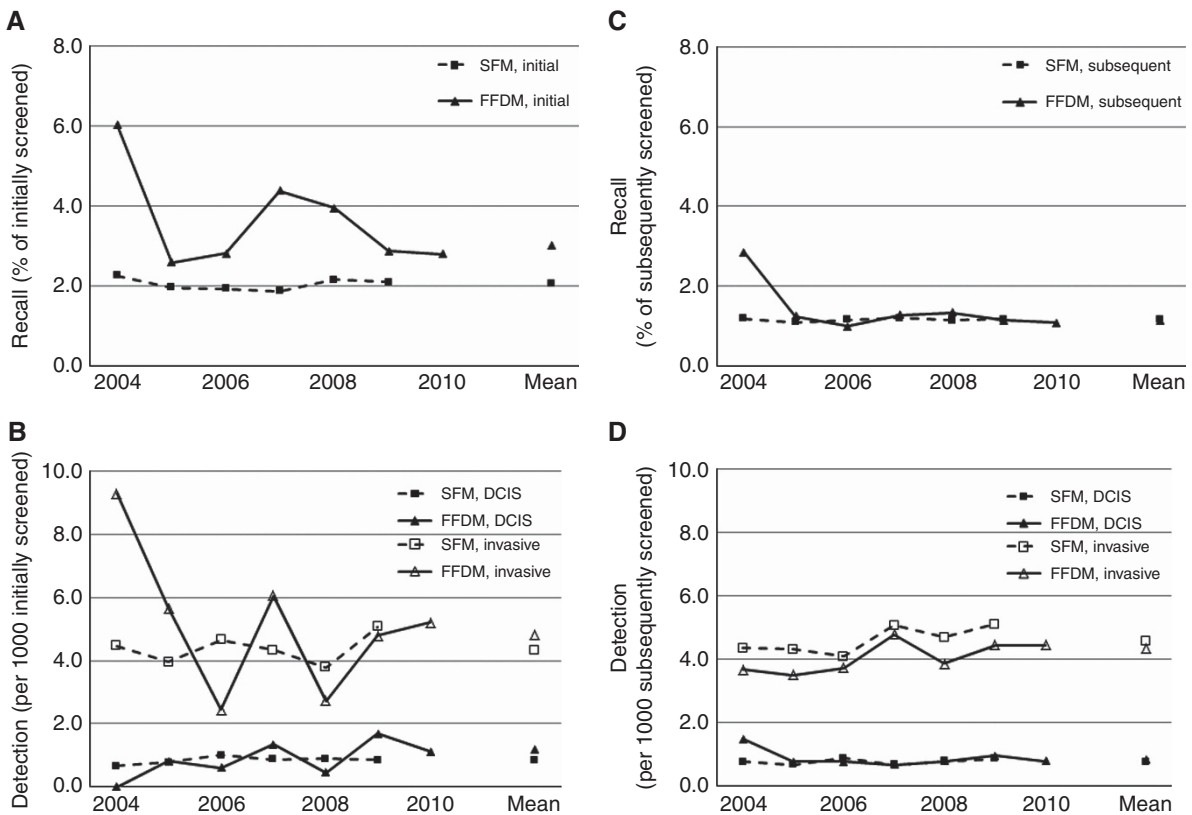


Figure 1. Recall and detection rates. (A) Percentage of women recalled after initial screens. (B) Detection rate per 1000 screened women after initial screens. (C) The percentage of women recalled after subsequent screens. (D) Detection rate per 1000 screened women after subsequent screens. Of the women with SFM as initial examination, 92% were aged 50–54 years, compared with 95% of women with FFDM as initial examination. For subsequent examinations, 18 and 17% were aged 50–54 years for SFM and FFDM, respectively; 16% were aged 70–75 years for SFM as well as FFDM.

to adjust contrast. After FFDM, women are more often recalled due to microcalcifications, which might be the result of this contrast adjustment (Bluekens *et al*, 2010; Nederend *et al*, 2012).

Unfortunately, we did not have any information if the higher recall rate after initial screens with FFDM lead to an increased number of biopsies performed in patients with a false-positive result.

Table 2. Tumour characteristics of screen-detected cancers in women 50–75 years of age for SFM and FFDM (2004–2010)

	Initial screen					Subsequent screen				
	SFM		FFDM		P-value	SFM		FFDM		P-value
	N	%	N	%		N	%	N	%	
DCIS										
Grade										
Low	11	17.5	4	13.8	0.860	47	11.7	15	8.6	0.524
Intermediate	25	39.7	13	44.8		132	32.9	61	34.9	
High	27	42.9	12	41.4		222	55.4	99	56.6	
Unknown	3		4			22		10		
DCIS, total	66		33			423		185		
Invasive cancers										
Tumour size										
T1a	16	4.7	14	10.6	0.138	137	5.3	72	7.4	0.126
T1b	57	16.9	22	16.7		645	25.2	237	24.3	
T1c	171	50.7	63	47.7		1296	50.6	476	48.9	
T2+	93	27.6	33	25.0		483	18.9	189	19.4	
Unknown	2		3			29		12		
Morphology										
Ductal	293	86.4	115	85.2	0.089	2114	81.6	839	85.1	0.030
Lobular	37	10.9	11	8.1		345	13.3	113	11.5	
Other	9	2.7	9	6.7		131	5.1	34	3.4	
Grade										
Low	105	32.3	46	37.7	0.469	813	32.5	268	28.5	0.024
Intermediate	151	46.5	55	45.1		1188	47.5	451	48.0	
High	69	21.2	21	17.2		498	19.9	220	23.4	
Unknown	14		13			91		47		
Lymph node status										
Negative	208	62.7	84	62.7	0.994	1859	73.1	717	73.5	0.794
Positive	124	37.3	50	37.3		684	26.9	258	26.5	
Unknown	7		1			47		11		
Invasive cancers, total	339		135			2590		986		

Abbreviations: DCIS = ductal carcinoma *in situ*; FFDM = full-field digital mammography; SFM = screen-film mammography.

Table 3. Women (50–75 years) screened, recalled and the number of women with screen-detected or interval cancers (2004–2008)

	SFM		FFDM		P-value
	N	%	N	%	
2004–2008					
Screened women	582 869		55794		
Initial screens	69 529	11.9	6768	12.1	
Subsequent screens	51 334	88.1	49 026	87.9	
Screen-detected		Rate^a		Rate^a	
Breast cancers	3089	5.30	274	4.91	0.226
Initial screens	365	5.25	38	5.61	0.692
Subsequent screens	2724	5.31	236	4.81	0.150
Interval					
Breast cancers	1370	2.35	135	2.42	0.748
Initial screens	187	2.69	17	2.51	0.787
Subsequent screens	1183	2.30	118	2.41	0.652
DCIS	67	0.11	4	0.07	0.525
Initial screens	8	0.12	1	0.15	0.567
Subsequent screens	59	0.11	3	0.06	0.370
Invasive cancers	1303	2.24	131	2.35	0.592
Initial screens	179	2.57	16	2.36	0.743
Subsequent screens	1124	2.19	115	2.35	0.481

Abbreviations: DCIS = ductal carcinoma *in situ*; FFDM = full-field digital mammography; SFM = screen-film mammography.

^aRates per 1000 women screened.

Table 4. Tumour characteristics of interval cancers in women 50–75 years of age for SFM and FFDM (2004–2008)

	Initial screen					Subsequent screen				
	SFM		FFDM		P-value	SFM		FFDM		P-value
	N	%	N	%		N	%	N	%	
DCIS										
Grade										
Low	1	12.5	0	0		11	22.4	1	33.3	
Intermediate	2	25	1	100		11	22.4	0	0	
High	5	62.5	0	0		27	55.1	2	66.7	
Unknown						10		0		
DCIS, total	8		1			59		3		
Invasive cancers										
Tumour size										
T1a	3	1.7	0	0	0.288	27	2.5	2	1.8	0.457
T1b	13	7.3	2	12.5		80	7.4	13	11.7	
T1c	67	37.9	9	56.3		419	38.7	41	36.9	
T2+	94	53.1	5	31.3		557	51.4	55	49.5	
Unknown	2		0			41		4		
Morphology										
Ductal	143	79.9	11	68.8	0.325	850	75.6	85	73.9	0.868
Lobular	31	17.3	5	31.3		196	17.4	22	19.1	
Other	5	2.8	0	0		78	6.9	8	7	
Grade										
Low	34	20.2	1	7.1	0.384	178	18.1	14	13.6	0.523
Intermediate	67	39.9	5	35.7		417	42.4	46	44.7	
High	67	39.9	8	57.1		389	39.5	43	41.7	
Unknown	11		2			140		12		
Lymph node status										
Negative	94	52.8	9	56.3	0.792	575	52.6	66	58.4	0.240
Positive	84	46.9	7	43.8		518	47.4	47	41.6	
Unknown	1		0			31		2		
Invasive, cancers total	179		16			1124		115		

Abbreviations: DCIS = ductal carcinoma *in situ*; FFDM = full-field digital mammography; SFM = screen-film mammography.

After initial screens, no differences in tumour characteristics between SFM and FFDM were found, which might indicate that FFDM performs similar for the entire spectrum of breast cancer. However, a non-significant trend towards more T1a cancers was seen after initial screens with FFDM. To further analyse this, the percentages of T1a cancers of all invasive cancers diagnosed after initial FFDM (10.6%) and SFM (4.7%) in this regional study were compared with the percentage of T1a cancers in initially screened women in the total Dutch population. For the period 1990–2007 (when most women were screened with SFM) this percentage was around 4% (Fracheboud *et al*, 2009). As we found a similar rate for SFM, this indicates that our data can be representative of the Dutch data and the trend towards more T1a cancers might be of interest. In literature a non-significant increase in T1a-c cancers was described, but the T1a cancers were not studied separately (Nederend *et al*, 2012).

After subsequent screens with FFDM, more ductal and high-grade invasive cancers were found, whereas other characteristics were similar. Another Dutch, multicentre study found no differences in grade in invasive cancers after subsequent screens (Bluekens *et al*, 2012). Our findings also differ from two Norwegian studies reporting no differences in tumour characteristics in the total screened population (Hofvind *et al*, 2014). These differences are interesting as they may reflect possible underlying differences in screening education or performance.

Overdiagnosis will always remain an issue in screening programs. It is accepted as an inevitable side effect of

the screening program, but the debate to which extent overdiagnosis is acceptable still continues (Independent UK Panel on Breast Cancer Screening, 2012a,b). Overdiagnosis can be defined as a malignancy that, if left undetected and untreated, never would have surfaced in a person's lifetime. An increase in low-grade DCIS might connote possible overdiagnosis. However, in our study we did not find a higher incidence of DCIS nor did we find more low-grade DCIS after FFDM compared with SFM. Another study did report a higher incidence of DCIS, but no differences in the distribution of grade (Bluekens *et al*, 2012). In contrast to our study, others did report that FFDM resulted in differences in grade. However, these studies reported opposite effects either towards more high-grade DCIS (Vigeland *et al*, 2008) or towards more low-grade DCIS (Nederend *et al*, 2012). Furthermore, any effect on breast cancer mortality is still unknown.

We found no decrease in the incidence rate of interval cancers after FFDM. The incidence of interval cancers for SFM and FFDM was similar, which is in concordance with a randomised trial comparing SFM and FFDM (Skaane *et al*, 2007). The rates of interval cancers in the Norwegian population-based screening program were also similar to our study (Hofvind *et al*, 2014). In addition, tumour characteristics did not differ between interval cancers diagnosed after SFM and FFDM, which is in concordance with other studies (Nederend *et al*, 2013; Hofvind *et al*, 2014).

The program sensitivity of SFM and FFDM found in the present study were both lower than in the study by Nederend

et al (2013), performed in another Dutch region. A Norwegian study found a slightly higher sensitivity for FFDM and lower for SFM than in our study (Skaane *et al*, 2007). Both studies did not discriminate between initial and subsequent screens, which might explain the differences. Compared with the program sensitivity of the complete Dutch program of 70% after initial screens and 71% after subsequent screens (Fracheboud *et al*, 2014), our study found a slightly lower sensitivity. The program sensitivity has been stable at this level since at least 2004 and the Health Council of The Netherlands advised to continue the organised screening program (Health Council of The Netherlands, 2014).

Our study has several strengths and limitations. In the current study, initial and subsequent screening examinations were analysed separately, increasing the insight of the effects of FFDM. Furthermore, this is a large population-based study using standardised collected data from the NCR, including tumour characteristics on tumour size, morphology, grade and lymph node status. Data on oestrogen or progesterone receptor or HER2neu status were not available for this cohort. A limitation in comparing results with other studies might lie in the fact that the Dutch screening program invites women 50–75 years of age, which differs from other screening programs mostly offering screening to women 50–69 years old. This might limit the generalisability of our results. Second, in The Netherlands, screening examinations are independently read by two radiologists and – particularly in the Northern region – recall rates are relatively low. Comparing FFDM with SFM might therefore lead to different conclusions in comparison with other studies. Furthermore, a Dutch study found variation in recall rate, with less variation in detection rate, between regions in The Netherlands (van Luijt *et al*, 2013). Finally, during this study period a policy change towards making standard cranio-caudal views at subsequent screening examinations started in The Netherlands. Also, during this period some radiologists synchronous read mammograms made using FFDM and SFM. Both effects could have influenced recall or detection rate for both SFM and FFDM. However, in our data we did not find an increased recall or detection rate for SFM after subsequent screening examination, indicating that these effects were negligible.

These data show that FFDM can be safely used for screening purposes. The easier image transfer capabilities of FFDM may prevent repeating the mammography during diagnostic work-up in the hospital. Further research on possible effects of the higher recall rate on the number of biopsies performed in patients with a false-positive result might add to the comparison of FFDM and SFM, as can future research of the proportion of T1a tumours, hormone receptor and HER2 status.

In conclusion, FFDM resulted in similar rates of screen-detected and interval cancers compared with SFM. This indicates that FFDM performs as well as SFM in a breast cancer screening program, with more ductal and high-grade invasive cancers found after subsequent screens. No signs of an increase in low-grade DCIS (which might connote possible overdiagnosis) were seen. Nonetheless, after initial screening, which accounts for 12% of all screens, FFDM resulted in higher recall rate and lower PPV that requires attention.

ACKNOWLEDGEMENTS

We thank the Dutch National Cancer Screening Program, region North (Bevolkingsonderzoek Noord) in Groningen for providing data on mammography. Furthermore, the registration teams of The Netherlands Comprehensive Cancer Organisation are thanked

for their effort in gathering the data in The Netherlands Cancer Registry, and the scientific staff of The Netherlands Cancer Registry.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Bluekens AM, Holland R, Karssemeijer N, Broeders MJM, den Heeten GJ (2012) Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a Multicenter Study. *Radiology* **265**(3): 707–714.
- Bluekens AM, Karssemeijer N, Beijerinck D, Deurenberg JJ, van Engen RE, Broeders MJM, den Heeten GJ (2010) Consequences of digital mammography in population-based breast cancer screening: initial changes and long-term impact on referral rates. *Eur Radiol* **20**(9): 2067–2073.
- D'Orsi CJB, Bassett LW, Berg W, Feig SA, Jackson VP, Kopans DB, Linver MN, Mendelson EB, Moss LJ, Sickles EA (2003) *Breast Imaging Reporting and Data System: ACR BI-RADS Mammography*, 4th edition. pp 7–175. American College of Radiology: Reston, VA, USA.
- Domingo L, Romero A, Belvis F, Sánchez M, Ferrer J, Salas D, Ibáñez J, Vega A, Ferrer F, Laso MS, Maciá F, Castells X, Sala M (2011) Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain. *Eur Radiol* **21**(9): 2020–2028.
- Del Turco MR, Mantellini P, Ciatto S, Bonardi R, Martinelli F, Lazzari B, Houssami N (2007) Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. *AJR Am J Roentgenol* **189**(4): 860–866.
- Fracheboud J, de Gelder R, Otto SJ, van Ineveld BM, Otten JDM, Broeders MJM, Verbeek ALM, den Heeten G, Holland R, de Bruijn AE, de Koning HJ (2009) *National Evaluation of Breast Cancer Screening in the Netherlands 1990–2007 (XII)*. Vol. 1st edn. p 94. Optima Grafische Communicatie: Rotterdam.
- Fracheboud J, van Luijt PA, Sankatsing VDV, Ripping TM, Broeders MJM, Otten JDM, van Ineveld BM, Heijnsdijk EAM, Verbeek ALM, Holland R, den Heeten GJ, de Bruijn AE, de Koning HJ (2014) *National Evaluation of Breast Cancer Screening in the Netherlands 1990–2011/2012 (XIII)*. pp 20–34. Optima Grafische Communicatie: Rotterdam.
- Fracheboud J, de Koning HJ, Boer R, Groenewoud JH, Verbeek AL, Broeders MJM, van Ineveld BM, Hendriks JH, de Bruyn AE, Holland R, van der Maas PJ. National Evaluation Team for Breast cancer screening in The N (2001) Nationwide breast cancer screening programme fully implemented in The Netherlands. *Breast* **10**(1): 6–11.
- Glynn CG, Farria DM, Monsees BS, Salzman JT, Wiele KN, Hildebolt CF (2011) Effect of transition to digital mammography on clinical outcomes. *Radiology* **260**(3): 664–670.
- Hofvind S, Skaane P, Elmore JG, Sebuødegård S, Hoff SR, Lee CI (2014) Mammographic Performance in a Population-based Screening Program: before, during, and after the Transition from Screen-Film to Full-Field Digital Mammography. *Radiology* **272**(1): 52–62.
- Health Council of the Netherlands (2014) *Population Screening for Breast Cancer: Expectations and Developments*. Health Council of the Netherlands: The Hague, publication no. 2014/01. ISBN 978-90-5549-991-5.
- Independent Uk Panel on Breast Cancer Screening (2012a) The benefits and harms of breast cancer screening: an independent review. *Lancet* **380**(9855): 1778–1786.
- Juel I-M, Skaane P, Hoff SR, Johannessen G, Hofvind S (2010) Screen-film mammography versus full-field digital mammography in a population-based screening program: the Sogn and Fjordane study. *Acta radiol* **51**(9): 962–968.
- Independent Uk Panel on Breast Cancer Screening (2012b) The breast cancer screening debate: closing a chapter? *Lancet* **380**(9855): 1714.
- Jørgensen KJ, Gøtzsche PC (2009) Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* **339**(b2587): 206–209.

- Karssemeijer N, Bluekens AM, Beijerinck D, Deurenberg JJ, Beekman M, Visser R, van Engen R, Bartels-Kortland A, Broeders MJM (2009) Breast cancer screening results 5 years after introduction of digital mammography in a Population-based Screening Program. *Radiology* **253**(2): 353–358.
- Lewin JM, Hendrick RE, D’Orsi CJ, Isaacs PK, Moss LJ, Karellas A, Sisney GA, Kuni CC, Cutter GR (2001) Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4945 paired examinations. *Radiology* **218**(3): 873–880.
- Nederend J, Duijm LEM, Louwman MWJ, Coebergh JW, Roumen RMH, Lohle PN, Roukema JA, Rutten MJCM, van Steenberg LN, Ernst MF, Jansen FH, Plaisier ML, Hooijen MJHH, Voogd AC (2013) Impact of the transition from screen-film to digital screening mammography on interval cancer characteristics and treatment—a population based study from the Netherlands. *Eur J Cancer* **50**(1): 31–39.
- Obenauer S, Luftner-Nagel S, von Heyden D, Munzel U, Baum F, Grabbe E (2002) Screen film vs full-field digital mammography: image quality, detectability and characterization of lesions. *Eur Radiol* **12**(7): 1697–1702.
- Nederend J, Duijm LEM, Louwman MWJ, Groenewoud JH, Donkers-van Rossum AB, Voogd AC (2012) Impact of transition from analog screening mammography to digital screening mammography on screening outcome in The Netherlands: a population-based study. *Ann Oncol* **23**(12): 3098–3103.
- Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D’Orsi C, Jong R, Rebner M (2005) Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* **353**(17): 1773–1783.
- Pisano ED, Yaffe MJ (2005) Digital mammography. *Radiology* **234**(2): 353–362.
- Skaane P, Hofvind S, Skjennald A (2007) Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in Population-based Screening Program: follow-up and final results of Oslo II Study. *Radiology* **244**(3): 708–717.
- Timmers JMH, van Doorne-Nagtegaal HJ, Zonderland HM, van Tinteren H, Visser O, Verbeek ALM, den Heeten GJ, Broeders MJM (2012) 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**(8): 1717–1723.
- Vigeland E, Klaasen H, Klingen T, Hofvind S, Skaane P (2008) Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. *Eur Radiol* **18**: 183–191.
- van Luijt PA, Fracheboud J, Heijnsdijk EAM, den Heeten G, de Koning HJ (2013) Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. *Eur J Cancer* **49**(16): 3517–3525.
- Wittekind C, Greene FL, Hutter RVP, Klimpfinger M, Sobin LH (2004) *TNM Atlas*. 5th edn. pp 207–223. Springer-Verlag: Berlin.
- Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, dos Santos Silva IM (2009) Full-field digital versus screen-film mammography: comparison within the UK Breast Screening Program and Systematic Review of Published Data. *Radiology* **251**(2): 347–358.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.