BREAST

Adriana M. J. Bluekens
Nico Karssemeijer
David Beijerinck
Jan J. M. Deurenberg
Ruben E. van Engen
Mireille J. M. Broeders
Gerard J. den Heeten

Consequences of digital mammography in population-based breast cancer screening: initial changes and long-term impact on referral rates

A. M. J. Bluekens (✉) · R. E. van Engen ·
M. J. M. Broeders · G. J. den Heeten
National Expert and Training Centre for
Breast Cancer Screening,
P.O. Box 6873, 6503 GJ, Nijmegen,
the Netherlands
e-mail: a.bluekens@lrcb.umcn.nl
Tel.: +31-24-3655155
Fax: +31-24-3655160

A. M. J. Bluekens
Department of Radiology,
St. Elisabeth Hospital, P.O. Box 90151,
5000 LC, Tilburg, the Netherlands

N. Karssemeijer
Department of Radiology,
Radboud University Nijmegen Medical
Centre, P.O. Box 9101, 6500 HB,
Nijmegen, the Netherlands

D. Beijerinck · J. J. M. Deurenberg
Preventicon Screening Centre/Mid-West,
Stationplein 91, 3511 ED, Utrecht,
the Netherlands

M. J. M. Broeders
Department of Epidemiology,
Biostatistics and HTA,
Radboud University Nijmegen Medical
Centre, P.O. Box 9101, 6500 HB,
Nijmegen,
the Netherlands

G. J. den Heeten
Department of Radiology,
Academic Medical Centre (AMC),
P.O. Box 22660, 1100 DD, Amsterdam,
the Netherlands

Abstract *Objectives:* To investigate the referral pattern after the transition to full-field digital mammography (FFDM) in a population-based breast cancer screening programme. *Methods:* Preceding the nationwide digitalisation of the Dutch screening programme, an FFDM feasibility study was conducted. Detection and referral rates for FFDM and screen-film mammography (SFM) were compared for first and subsequent screens. Furthermore, radiological characteristics of referrals in digital screening were assessed. *Results:* A total of 312,414 screening mammograms were performed (43,913 digital and 268,501 conventional), with 4,473 consecutive referrals (966 following FFDM). Initially the FFDM referral rate peaked, and many false-positive results were noted as a consequence of pseudolesions and increased detection of (benign) microcalcifications. A higher overall referral rate was observed in FFDM screening in both first and subsequent examinations ($p<.001$), with a significant increase in cancer detection ($p=.010$). *Conclusion:* As a result of initial inexperience with digital screening images implementing FFDM in a population-based breast cancer screening programme may lead to a strong, but temporary increase in referral. Dedicated training in digital screening for radiographers and screening radiologists is therefore recommended. Referral rates decrease and stabilise (learning curve effect) at a higher level than in conventional screening, yet with significantly enhanced cancer detection.

Keywords Breast cancer ·
Digital mammography ·
Mass screening · Referral rate ·
Training

Introduction

Screening mammography has been shown to be effective in reducing breast cancer mortality [1, 2]. Full-field digital mammography (FFDM) is increasingly used in the clinical setting, with a number of advantages resulting in better detection of (early) breast cancer and improvements in workflow [3].

Because of innovations, fast technical developments and emerging evidence that FFDM as a screening test is at least as effective as conventional screen-film mammography (SFM), the national coordinating team of the Dutch breast cancer screening decided to start digital pilot projects in 2003. Their main aim was to gain experience and to test the feasibility of digitalisation of the nationwide Dutch screening programme.

More recent studies concerning diagnostic accuracy confirmed that digital mammography as a screening technique in a population-based screening programme is at least as accurate as SFM [4–11]. In several studies an increase in referral rate following the implementation of FFDM has been reported [4, 5, 8], a tendency we also recognised in our previous study [11]. Nevertheless, this effect on referral rate has never been analysed in detail. Since changes in referral rate directly affect the daily screening practice and work-up, understanding this increase in referral rate is of crucial importance, and analysis could be useful to design a dedicated training programme.

The focus of this study, as part of the Dutch digital pilot projects, was to investigate the referral pattern after the introduction of digital mammography in the national screening programme. We provide a comprehensive evaluation of the changes in referral rates over time, reported separately for first and subsequent screening examinations compared with those in conventional mammography. Additionally, a detailed analysis of the radiographic characteristics of the referred lesions will be given.

Materials and methods

Setting

The current observational study encompasses the results of the first 4 years of the digital mammography pilot, which started in October 2003 at Preventicon screening centre (Utrecht, the Netherlands). At that time digital mammography was introduced by replacing one of the conventional units by an FFDM system. The pilot was part of the national screening programme, which offers mammography to all women aged 50 to 75 years every 2 years. Participation is on the basis of a written invitation by mail based on information provided by the national population registry. Details concerning the national screening programme have been published elsewhere [12]. When further assessment is indicated, women are referred to a breast team of a nearby hospital.

Study population

All women participating in the screening programme at Preventicon screening centre between October 2003 and December 2007 were included in the study. In the first year after the introduction of the FFDM system only women attending their first screening examination were offered digital mammography. One year after the introduction, FFDM also became available for women attending subsequent screens. Assignment of women to FFDM or SFM was based on the random availability of the units when participants presented at the screening centre. However, women who already had a previous digital screening mammogram were always offered FFDM.

Participants were informed in writing about the possibility of having a digital mammogram. They had the right to refuse this offer and undergo conventional mammography. To comply with privacy regulations, participants signed a general informed consent form which permits use of data from the screening programme for evaluation and scientific research. Refusal of sharing data for these purposes was extremely rare. The study was approved by the regional medical ethics review board. Specific written informed consent for this study was not required.

Image acquisition and interpretation

Screen-film mammograms were acquired on two types of systems (GE 600/800T, GE Healthcare, Buc, France). All digital mammograms were acquired using a Lorad Selenia FFDM system (Hologic Inc., Danbury, CT) with a 70-μm pixel size and a 232×286-mm field of view. First screening examinations always included mediolateral-oblique (MLO) and craniocaudal (CC) views. At subsequent screening examinations MLO views of each breast were routinely acquired and, when indicated, CC views were also obtained. Radiographers are extensively trained to decide on obtaining CC views in subsequent screening examinations (FFDM as well as SFM) based on prespecified criteria for breast density and visible abnormality. In the case of a subsequent screening examination (digitised) prior mammograms were available at all times. Computer aided detection (CAD) was freely available for FFDM (ImageChecker, Hologic/R2, Santa Clara, CA). Digital mammograms were interpreted using soft-copy reading. The same team of radiologists was involved in the double-reading of FFDM and SFM, with differences of opinion resolved by consensus. One of two radiologists (D.B., J.D.), both with more than 15 years of experience in mammography screening, was involved in each screening examination performed during the entire study period. They were complemented by a team of six screening radiologists, each performing over 5,000 readings per year. They all had at least 2 years' experience with working in a digital radiology environment before the study started. None of the readers had experience with the use of FFDM in screening.

More detailed descriptions of image acquisition and interpretation, reading conditions, and diagnostic work-up used in the Preventicon pilot are published elsewhere [11].

Data collection

Retrospectively we collected screening outcomes and demographic features of all women included in the study. Normal findings at assessment or benign breast disease (including lobular carcinoma in situ) were reported as a negative (or false positive) screening outcome, whereas a positive screening outcome indicated malignancy (including ductal carcinoma in situ). Radiological characteristics

Table 1 Screening performance of SFM compared with FFDM, for first and subsequent individual screening examinations

Figures in parentheses are rates per 1,000 women screened

[a] Result referral, missing values: 2 analogue screens (both subsequent examinations), 13 digital screens (5 first and 8 subsequent examinations)

[b] Per 24 months (regular screening interval) in subsequent screens

Variable	SFM	FFDM	P (χ^2)
Screened women (n)			
Total	268,501	43,914	
First screening examination	33,361	8,577	
Subsequent screening examination	235,140	35,337	
Referred women[a] (n, rate)			
Total	3,507 (13.1)	966 (22.0)	.000
First screening examination	1,136 (34.1)	368 (42.9)	.000
Subsequent screening examination	2,371 (10.1)	598 (16.9)	.000
Breast cancers detected (n, rate)			
Total	1,369 (5.0[b])	250 (5.9[b])	.010[b]
First screening examination	199 (6.0)	65 (7.6)	.092
Subsequent screening examination	1,170 (4.9[b])	185 (5.5[b])	.085[b]
PPV of referral (%)			
Total	39.7	26.2	.000
First screening examination	25.9	17.9	.147
Subsequent screening examination	43.7	31.4	.002

of all referred women screened digitally during the study period were reviewed.

Study data were collected from the Preventicon archive and digital database, which include number of women screened and referred, radiological reports and final diagnoses. A search of the nationwide network and archive of histocytopathology (PALGA) completed the data.

Abnormalities seen on mammograms followed by referral were categorised as: mass, calcifications and other.

Data analysis

All performance indicators were computed separately for first and subsequent screening examinations. Screening intervals were calculated for subsequent screening examinations by taking the period between the current and previous screen. Because screening intervals where somewhat different in the two populations we computed detection per 24 months, by multiplying the observed rates by $24/T$, with T denoting the median screening interval. The difference occurred due to different logistics in the permanent facility where FFDM was installed, whereas most SFM units were functional in mobile facilities.

We compared referral rate, breast cancer detection and positive predictive value (PPV) for FFDM and SFM separately for first and subsequent screening examinations. In addition, radiological characteristics of referrals in digital screening were assessed.

Statistical software was used for data management and analysis (SPSS, version 15.0 for Windows; SPSS Inc., Chicago, IL, USA). Results were compared using the Chi-square test. A p value of less than 0.05 was considered statistically significant. Dot graphs were drawn with a Loess fitted curve to evaluate trends in referral rates for both modalities. Bar graphs were used to present changes in radiological characteristics of referred lesions and the concurrent screening results. For all graphs variables were assessed per 2 months.
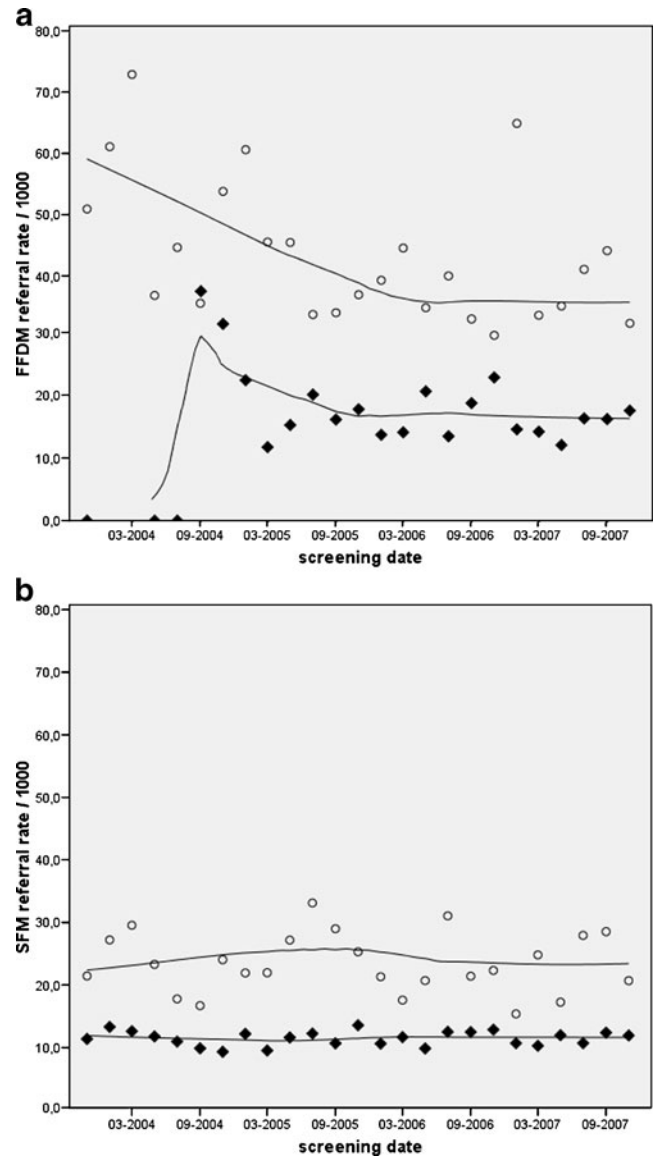

Fig. 1 Time course of referral in **a** FFDM and **b** SFM screening for first (○) and subsequent (♦) screening examinations

Results

During the 4-year study period a total of 312,415 screens were performed, consisting of 43,914 digital and 268,501 conventional mammography procedures. In the FFDM group a total of 966 women were referred, in the SFM group 3507. The mean age of referred women was 58.5 years for conventional screening and 57.4 years for digital screening (range 50–75 years). The median screening interval was 22.7 months in the group of referrals screened with FFDM and 24.6 months for referrals after SFM screening. Breast cancer was detected in 1,619 women of whom 250 had digital mammograms.

Screening performance

The screening performance is summarised in Table 1. Significantly more cancers were detected with FFDM, corrected for screening interval ($p=0.010$). In initial screening examinations the detection rate was 7.6‰ with FFDM and 6.0‰ with SFM ($p=0.092$). In subsequent screens detection rates were 5.5‰ and 4.9‰, respectively ($p=0.085$). Referral rate was significantly higher with FFDM, both in initial screens (from 3.4% to 4.3% with FFDM ($p<0.001$)), and in subsequent screens (from 1.0% to 1.7% ($p<0.001$)). As a result of the increase in referrals the overall PPV of referral decreased significantly with digital mammography. For first screening examinations the PPV decreased from 25.9% to 17.9%. For subsequent screens the PPV went from 43.7% to 31.4%.

Trends in referral and detection rate in conventional and digital screening

In Fig. 1 the referral rates for first and subsequent screening examinations with FFDM and SFM respectively are presented separately. The implementation of FFDM induced an increase in referral in first screens (73‰). With the introduction of subsequent screens in July 2004 another peak in referral was noted. This second peak in referral (37‰) was significantly higher ($p<0.001$) compared with its future baseline and relatively much higher than the decreasing, though still elevated, referral rate in first screens during that same period.

Within a few months the rate started to decline in both groups. Referral stabilised after a little more than 1 year of digital screening, however at a higher level than in conventional screening.

The high referral rate was reflected in the detection. In the first year the mean detection rate in digital screening was 7.8 per 1,000 women screened. Detection stabilised with decreasing referral rate, but overall remained higher than the detection in conventional screening: 5.9‰ versus 5.0‰, respectively. Referral and detection rates with SFM remain relatively constant over the entire study period.

Trends in digitally referred lesions

To evaluate the initial peak and subsequent changes in the referral rate in digital screening mammography we analysed the radiological characteristics of the referred lesions (Fig. 2). In the first months masses and densities were largely responsible for all referrals. Asymmetry and distortion (elements of the rest category 'other') were also observed more frequently. Half a year later the relatively high rates of referral were due to microcalcifications. Both masses and suspect calcifications appeared to be evenly responsible for referral at the time the mean referral rate stabilised.

Figure 3 depicts the share of detected cancers in all referred lesions, separately for masses and microcalcifications. The elevated referral rate seen in the first period of digital mammography led to a disproportionately high

Fig. 2 Radiological characteristics of referred lesions in digital mammography (rates per 1,000)
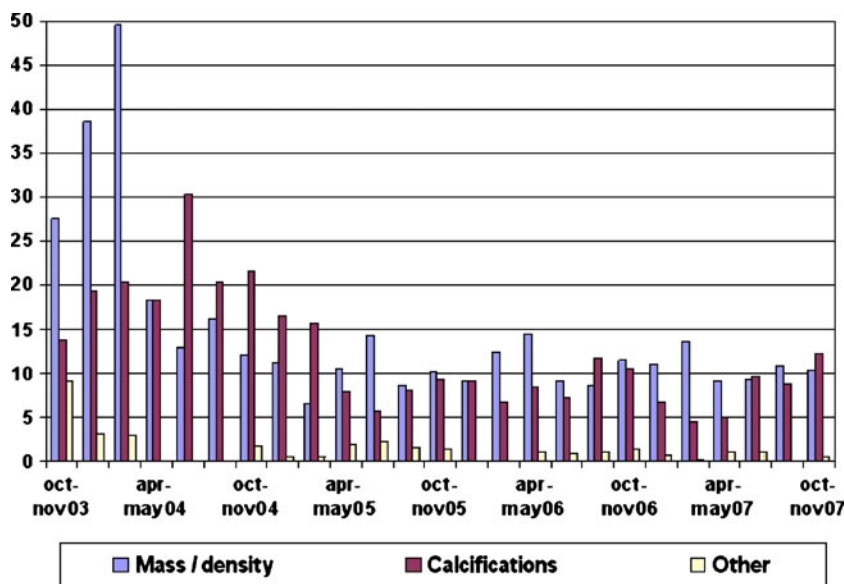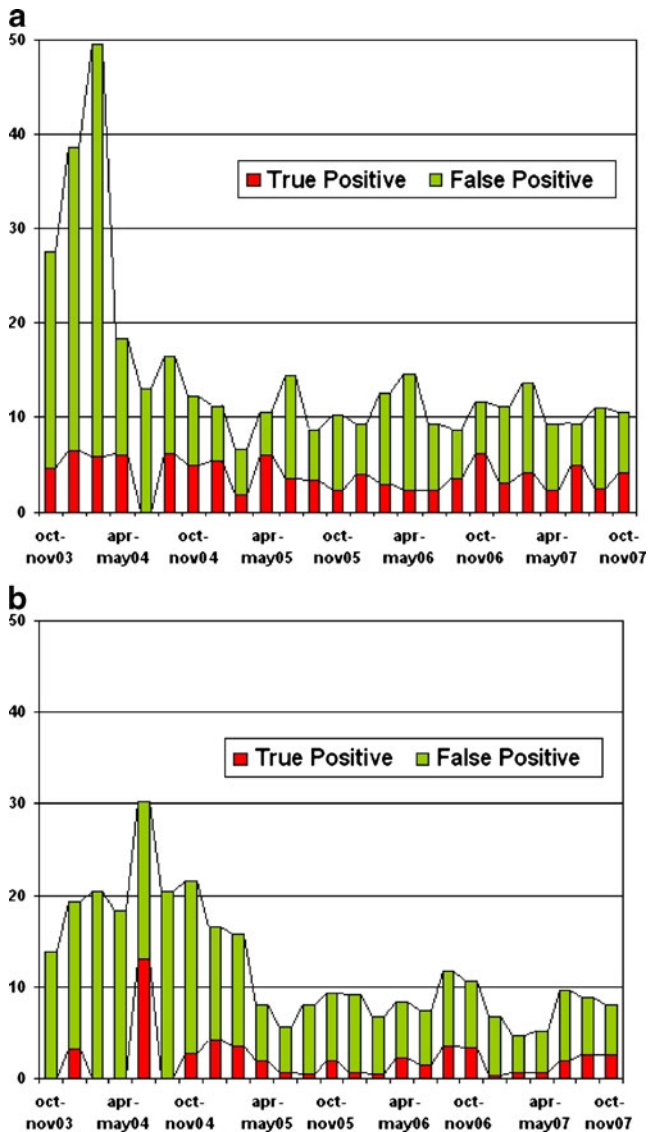
Fig. 3 Results of **a** referred masses and **b** referred microcalcifications (rates per 1,000)

amount of false-positive results (up to 88%), with masses referred twice as often as microcalcifications. Cancer detection in masses remained at a relatively high and constant level. In the first 6 months nearly all referred microcalcification clusters turned out to be benign after histological assessment. The subsequent increase in referrals based on detected microcalcifications coincided with a transient higher cancer incidence. When referral and detection stabilised roughly 1 in 4 referrals led to the detection of cancer (PPV 25%).

Discussion

In accordance with previous European studies [4–10] we found an improved detection with digital mammography, although this time significantly. In first screening exami-

nations, i.e. the youngest women of the screening population, the relative increase in detection was most prominent. This finding is in line with the results of the DMIST trial that first showed that FFDM performed markedly better in younger women than SFM [13]. In our study, the overall referral rate was higher in digital screening mammography. Referral rates may fluctuate in time, and an increase may sometimes be observed, for example after a course or meeting is attended. Since the referral rate in conventional mammography did not change during that same period, such a temporary trend is highly unlikely. Therefore, the high referral rate may be ascribed to the introduction of digital screening. The increase in referral was most distinct in the first period when only first screens were performed. However, a following substantial increase was present when subsequent examinations were introduced while a part of the learning curve had already been passed. The increase in referral in turn led to a lower PPV of referral.

An increase in referral in FFDM has been reported by previous studies evaluating screening programmes resembling the Dutch one [4, 5, 8]. On the other hand, two studies describing digital screening mammography in Scandinavia, found the highest referral rate in SFM [6, 7]. Both Scandinavian studies involved a subset of the screening population, concerning subsequent screens or first screens only. Furthermore, the study by Heddson et al. [6] was performed after a 6-month learning period, which might have lowered the overall referral rate. Two studies conducted in Spain and the UK reported no differences in referral after transition to FFDM in their screening programme [9, 10]. The relatively high baseline referral rate in Spain (11.5% in first screens, and 3.6% in subsequent screens) and the UK (overall 4.4%) might explain this outcome.

While former studies [4–11] focus on the aggregate screening results of FFDM compared with SFM, this study, to our knowledge, is the first to analyse the effects over time after the transition to digital screening mammography. Referral peaks were seen in the first months, followed by a rapid decrease. The overall referral rate stabilised at a higher level than in screen-film mammography. The significant rise in referral might have affected the cancer detection rate as the overall detection rate tends to be higher in FFDM. Our previous study showed this to be largely attributable to relatively more detected cancers in first screens [11].

The introduction of FFDM in the screening programme initially caused a disproportionate increase in referral with subsequent false-positive results. Three reasons for this effect can be distinguished; all related to the use of post-processed high contrast images. First of all, high contrast resolution accentuates not only lesions, but normal architecture as well. Therefore, simulated lesions, or pseudolesions, could be more often encountered in FFDM (Fig. 4). Although true lesions should be better distinguishable, the reading and interpretation of digital high volume screening images is a learning process. This might explain the high referral rate caused by densities and

Fig. 4 Pseudolesion seen with high contrast resolution in FFDM. **a** Prior screening mammogram (SFM), **b** screening mammogram followed by referral (FFDM), **c** mammogram in subsequent clinical assessment (FFDM)
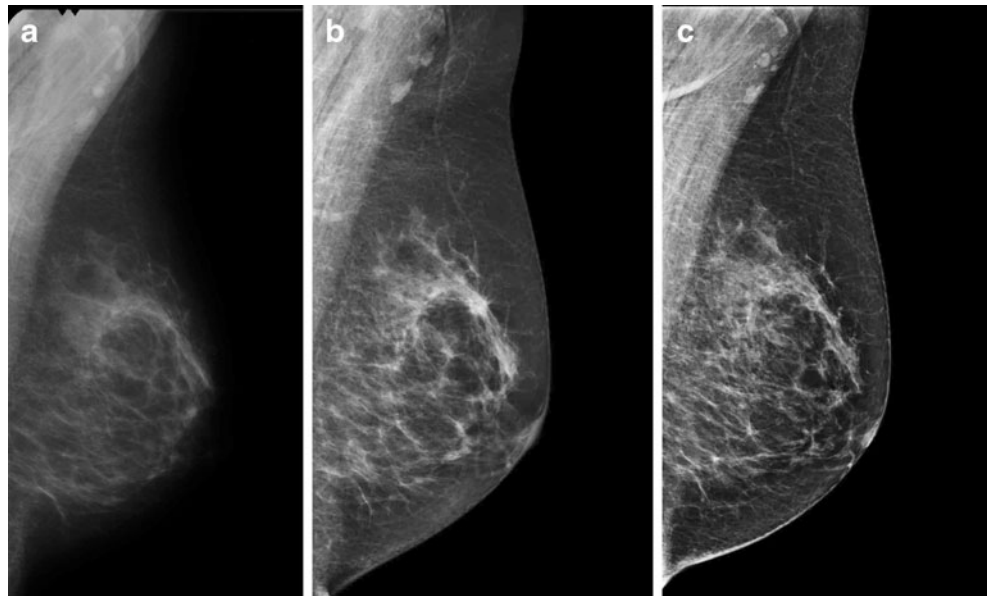
distortions, accompanied by an increase in the number of false positives in the first months.

Secondly, higher contrast resolution improves the detection of microcalcifications [14]. This explains the increased share of referrals on the basis of micro-calcifications and is in line with the significant increased detection of ductal carcinoma (both in situ and invasive) found in our previous study [11]. Since most clusters of microcalcifications do not represent malignancy, interpreting calcification type is another challenge in FFDM. Therefore, it is not surprising that in the first few months of digital screening we have seen an exceedingly high proportion of false-positive referrals due to microcalcifications.

The superior contrast resolution of digital mammography had yet a third effect. Compared with a previous conventional image, a digital mammogram can falsely give the impression that a lesion or cluster of microcalcifications is new or has grown. This doubt about progression in the presence of conventional prior mammograms can cause more women to be referred. While in other studies prior mammograms were offered only on request [5] or not until consensus meeting [4], we had (digitised) prior mammograms present at all times. Prior mammograms increase the specificity of screening [15] and not presenting them instantaneously was considered unethical. Accordingly, the introduction of subsequent screens a few months after the pilot started resulted in a second rise in referral.

Thus, while the higher overall referral rate can be justified by the improved cancer detection, the peak in referral and the high proportion of false-positive results just after implementation of FFDM primarily reflected inexperience with the interpretation of digital images and the comparison with digitised conventional prior mammograms in the screening setting.

The design of our study included the use of CAD, because we believe that with digital mammography this will become standard practice in screening programs. The fact that previous studies without CAD also reported increased detection of microcalcifications suggests that this result should not be attributed to CAD alone. On the other hand, CAD tends to increase referral rate as well [16]. However, this should not have influenced the course of referral over time.

Measurement of the effect of CAD as a separate variable was not a subject of this study. Most reports in the literature demonstrate a benefit of CAD when single reading is practised. However, double reading with consensus may yield comparable results [16]. Our study shows a combined effect of double reading and CAD. Double reading minimises perception errors and improves decisions, but most likely reduces the incremental benefit of CAD. We therefore reckon the possible confounding effect of CAD in this study to be marginal.

Two weaknesses of this study need to be addressed. This study was performed on one type of mammography unit (i.e. Lorad Selenia) with its proprietary image processing and display. It is, therefore, not clear whether these results can be generalised to other manufacturers as well. A comparative multi-centre study is needed to evaluate the applicability of these results in all mammography units.

Secondly, a pilot setting in which screening radiologists read conventional as well as digital mammograms, may not be entirely representative for the screening programme setting in which all mammograms have the same imaging characteristics. The mixture of conventional and digital mammograms in the pilot setting may have contributed to the increase in referral.

Screening programmes around the world vary in many different aspects. For instance, compared with screening on the basis of self-referral as performed in the USA, population-based screening programmes are character-ised by a relatively low baseline referral rate. The Netherlands in particular is known for its low referral

rate. Generalisation of our findings to other screening programmes is therefore not straightforward. However, our study results might still be of interest to those programmes with a similar set-up, planning a transition to digital mammography.

In the Netherlands, as a consequence of our findings we developed a dedicated training in digital screening mammography for radiographers and screening radiologists. Preliminary reports suggest this training is effective in the Netherlands since a similar effect in newly digitalised screening centres is not observed. We therefore recommend dedicated training in digital mass screening when converting to FFDM in a population-based breast cancer screening programme, especially when characterised by a relatively low referral rate.

References

1. Tabar L, Yen M, Vitak B, Chen H, Smith R, Duffy S (2003) Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. Lancet 361:1405–1410

2. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH et al (2003) Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 361:1411–1417

3. Bick U, Diekmann F (2007) Digital mammography: what do we and what don't we know? Eur Radiol 17:1931–1942

4. Skaane P, Skjennald A (2004) Screen-film mammography versus full-field digital mammography with soft-copy reading: randomised trial in a population-based screening program—Oslo II study. Radiology 232:197–204

5. Rosselli Del Turco M, Mantellini P, Ciatto S, Bonardi R, Martinelli F, Lazzari B et al (2007) Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. AJR 189:860–866

6. Heddson B, Ronnow K, Olsson M, Miller D (2007) Digital versus screen-film mammography: a retrospective comparison in a population-based screening program. Eur J Radiol 64:419–425

7. Vigeland E, Klaasen H, Klingen TA, 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

8. Hambly N, Phelan N, Hargaden G, O'Doherty A, Flanagan F (2008) Impact of digital mammography in breast cancer screening: initial experience in a national breast screening program. In: Krupinski EA (ed) IWDM 2008. Lecture notes in computer scince , vol 5116. Springer, Berlin, pp 55–60

9. Sala M, Comas M, Macià F, Martinez J, Casamitjana M, Castells X (2009) Implementation of digital mammography in a population-based breast cancer screening program: effect of screening round on recall rate and cancer detection. Radiology 252:31–39. doi:10.1148/radiol.2521080696

10. 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 251:347–358

11. Karssemeijer N, Bluekens A, Broeders M, Deurenberg J, Beekman M, Visser R, Van Engen R, Bartels-Kortland A, Beijerinck D (2009) Breast cancer screening results five years after introduction of digital mammography in a population-based screening program. Radiology 253:353–358

12. Holland R, Rijken H, Hendriks J (2007) The Dutch population-based mammography screening: 30-year experience. Breast Care 2:12–18

13. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S et al (2005) Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med 353:1773–1783

14. Fischer U, Baum F, Obenauer S, Luftner-Nagel S, von Heyden D, Vosshenrich R, Grabbe E (2002) Comparative study in patients with microcalcifications: full-field digital mammography vs screen-film mammography. Eur Radiol 12:2679–2683

15. Roelofs AA, Karssemeijer N, Wedekind N, Beck C, van Woudenberg S, Snoeren PR, Hendriks JH, Rosselli del Turco M, Bjurstam N, Junkermann H, Beijerinck D, Séradour B, Evertsz CJ (2007) Importance of comparison of current and prior mammograms in breast cancer screening. Radiology 242:70–77

16. Taylor P, Potts HW (2008) Computer aids and human second reading as interventions in screening mammography: two systematic reviews to compare effects on cancer detection and recall rate. Eur J Cancer 44:798–807