

# Breast Cancer Screening Results 5 Years after Introduction of Digital Mammography in a Population-based Screening Program<sup>1</sup>

Nico Karssemeijer, PhD  
Adriana M. Bluekens, MD  
David Beijerinck, MD  
Jan J. Deurenberg, MD  
Matthijs Beekman  
Roelant Visser, MS  
Ruben van Engen  
Annemieke Bartels-Kortland, MD  
Mireille J. Broeders, PhD

## Purpose:

To compare full-field digital mammography (FFDM) using computer-aided diagnosis (CAD) with screen-film mammography (SFM) in a population-based breast cancer screening program for initial and subsequent screening examinations.

## Materials and Methods:

The study was approved by the regional medical ethics review board. Informed consent was not required. In a breast cancer screening facility, two of seven conventional mammography units were replaced with FFDM units. Digital mammograms were interpreted by using soft-copy reading with CAD. The same team of radiologists was involved in the double reading of FFDM and SFM images, with differences of opinion resolved in consensus. After 5 years, screening outcomes obtained with both modalities were compared for initial and subsequent screening examination findings.

## Results:

A total of 367 600 screening examinations were performed, of which 56 518 were digital. Breast cancer was detected in 1927 women (317 with FFDM). At initial screenings, the cancer detection rate was .77% with FFDM and .62% with SFM. At subsequent screenings, detection rates were .55% and .49%, respectively. Differences were not statistically significant. Recalls based on microcalcifications alone doubled with FFDM. A significant increase in the detection of ductal carcinoma in situ was found with FFDM ( $P < .01$ ). The fraction of invasive cancers with microcalcifications as the only sign of malignancy increased significantly, from 8.1% to 15.8% ( $P < .001$ ). Recall rates were significantly higher with FFDM in the initial round (4.4% vs 2.3%,  $P < .001$ ) and in the subsequent round (1.7% vs 1.2%,  $P < .001$ ).

## Conclusion:

With the FFDM-CAD combination, detection performance is at least as good as that with SFM. The detection of ductal carcinoma in situ and microcalcification clusters improved with FFDM using CAD, while the recall rate increased.

© RSNA, 2009

<sup>1</sup> From the Department of Radiology, Radboud University Nijmegen Medical Centre, Geert Grootteplein 10, 6500 HB Nijmegen, the Netherlands (N.K.); National Expert and Training Centre for Breast Cancer Screening, Nijmegen, the Netherlands (N.K., A.B., R.V., R.v.E., M. Broeders); and Preventicon, Utrecht, the Netherlands (D.B., J.D., M. Beekman, A.B.K.). Received February 5, 2009; revision requested March 4; revision received April 3; accepted April 30; final version accepted May 19. Supported in part by a grant from the European Community in the 5th Framework Information Society Technologies program (IST-2001-33439, SCREEN-TRIAL). Address correspondence to N.K. (e-mail: [n.karssemeijer@rad.umcn.nl](mailto:n.karssemeijer@rad.umcn.nl)).

**S**creen-film mammography (SFM) is increasingly being replaced with digital systems because of their consistent image quality, the ability of postprocessing, and improved storage and communication capabilities. To benefit effectively from the new technology, screening organizations have to make a transition that goes far beyond replacement of mammography units, because a new infrastructure has to be implemented for archiving, soft-copy reading, and reporting. In screening organizations that operate nationwide, the scale at which digital technology is to be implemented is much larger than in clinical environments. This requires careful planning and may partly explain the relatively slow uptake of digital mammography in these programs.

Some large-scale studies have been conducted to date to compare digital with conventional mammography. Results suggest that digital mammography is at least as good as SFM in the clinical screening setting (1,2) and in population-based screening practice (3–7). A review of studies comparing digital with SFM was presented by Skaane (8).

In preparation of digitization of the nationwide breast cancer screening program in the Netherlands, digital mammography was installed in 2003 in a project at the Preventicon screening center in Utrecht. The purpose of the project was to demonstrate the effectiveness of digital breast cancer screening using soft-copy reading with computer-aided diagnosis (CAD) and to study problems related to the transition, such as dealing with prior SFM images (9). During this project, the majority of the screening examina-

tions performed at the center remained film based.

The purpose of this study was to compare full-field digital mammography (FFDM) using CAD with SFM in a population-based breast cancer screening program for initial and subsequent screening examination findings.

### Materials and Methods

MeVis Medical Solutions (Bremen, Germany) was a participant in the European-funded project in which this study was initiated. N.K. was a scientific consultant to R2/Hologic (Santa Clara, Calif) during part of the study period. Non-consultant authors had full control of the data and the information submitted for publication.

### Study Population

This study was conducted within the context of an ongoing population-based breast cancer screening program for asymptomatic women aged 50–75 years at the Preventicon screening center (Utrecht, the Netherlands). In this program, screening is conducted at a regular 2-year interval involving only mammography. Participation is on the basis of a written invitation by mail according to information provided by the national population registry. There are no exclusion criteria. Details concerning the program have been described previously (10,11). Digital mammography was introduced at Preventicon in 2003, with the replacement of one of two mammography units with a FFDM system in one facility. Five other conventional units were kept operational at other locations. In the 1st year after the introduction, only women attending their first screening examination were offered digital mammography. From 1 year after the introduction, women attending subsequent screenings were

also included. Assignment of women to FFDM or SFM was based on the availability of the units when participants presented at the screening center. However, women who already had a previous digital screening mammogram were always offered FFDM. In 2007, a second FFDM system was installed at the study location, and after July of that year almost all mammograms at this facility were digital.

Participants were informed in writing about the possibility of undergoing digital mammography, and they had the right to refuse and undergo conventional mammography. To comply with privacy regulation, they signed a general informed consent that permits use of data from the screening program for evaluation and scientific research. 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

SFM images were acquired with two types of systems: one using a molybdenum target and filter (600T; GE Healthcare, Buc, France) and one using a molybdenum target and molybdenum and rhodium filter (800T; GE Healthcare). Both systems used a Min-R 2000/Min-R 2190 (Kodak, Rochester, NY) screen-film

### Advances in Knowledge

- Full-field digital mammography (FFDM) with computer-aided diagnosis (CAD) improved detection of microcalcifications compared with screen-film mammography.
- There is a significant increase in the fraction of invasive cancers only visible because of microcalcifications with use of FFDM.

### Implication for Patient Care

- FFDM with CAD demonstrates advantages for screening younger women and better detection of calcifications associated with breast cancer.

### Published online before print

10.1148/radiol.2532090225

Radiology 2009; 253:353–358

### Abbreviations:

CAD = computer-aided diagnosis  
DCIS = ductal carcinoma in situ  
FFDM = full-field digital mammography  
PPV = positive predictive value  
SFM = screen-film mammography

### Author contributions:

Guarantors of integrity of entire study, N.K., M. Broeders; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, N.K., A.B., D.B., J.D., A.B.K., M. Broeders; clinical studies, J.D., R.v.E.; experimental studies, N.K., J.D.; statistical analysis, N.K., A.B., J.D., M. Beekman, M. Broeders; and manuscript editing, N.K., A.B., J.D., A.B.K., M. Broeders

See Materials and Methods for pertinent disclosures.

combination. All digital mammograms were acquired by using Lorad Selenia FFDM systems (Hologic, Danbury, Conn). Technique factors and breast doses for the FFDM and SFM units were monitored and found to be in compliance with the national and, where applicable, European guidelines. Mammograms were processed with commercially released, proprietary imaging processing algorithms. During the course of the study, imaging processing algorithms were regularly updated.

Initial screening examinations performed with FFDM or SFM always included the two standard views, craniocaudal and mediolateral oblique. At subsequent screening examinations, mediolateral oblique views of each breast were routinely acquired and, when indicated, craniocaudal views were also obtained by using criteria based on breast density and visible abnormality. The radiographers involved in the study received extensive training in the use of FFDM. They were instructed to obtain the best possible positioning and compression with each modality and used the same protocol to determine whether to acquire craniocaudal views at subsequent screening examinations. To this end, a dedicated workstation with a high-resolution monitor was installed in their work area to allow proper viewing of digital mammograms.

Mammograms were interpreted in a batch mode within 2 days of acquisition. All mammograms were read independently, with final decisions about recall resolved by consensus. Decisions did not include recommendations for biopsy or short-term follow-up. Diagnostic assessment was performed in nearby hospitals without involvement of the screening center. One of two radiologists (J.D., D.B., each with more than 15 years of experience in mammography screening) was involved in each screening examination performed during the entire study period. In total, they performed approximately 75% of the readings. The rest of the readings were performed by a team of six, and later seven, radiologists, each performing more than 5000 screening examinations per year. Of these radiologists, five

were involved during the whole study period. All radiologists were involved in SFM and FFDM screening, and they all had more than 2 years experience with working in a digital radiology environment before the study started. None of the readers had experience with use of FFDM in screening or with the type of processing implemented in the FFDM system used in the study. All radiologists had extensive experience with clinical use of digital mammography with a computed radiography detector.

Conventional mammograms were read in a darkened room by using mammogram alternators with a luminance of at least 2500 cd/m<sup>2</sup>. In subsequent screenings, the most recent prior mammograms were always mounted with the current screening mammograms. FFDM cases were interpreted in a separate room, with reading conditions optimized for soft-copy reading. A dedicated mammography workstation equipped with two 5-megapixel displays (Mevis Medical Solutions) was used. To facilitate soft-copy reading of subsequent screening examinations, the most recent prior screening mammograms of women who underwent FFDM were digitized by using a film scanner and archiver designed for mammography (DigitalNow; R2/Hologic). Original prior screening mammograms were also available for viewing.

A default protocol for presentation of mammograms was installed on the workstation. First, the current mammogram was displayed along with available prior mammograms. Next, all views were inspected in full-screen mode, where readers could use quadrant roaming and/or zooming for full resolution. Image manipulation tools could be used and included contrast manipulation and image inversion. For making comparisons with prior images, most readers used toggling. CAD was available for FFDM (ImageChecker; R2/Hologic), with software upgraded to the most recent versions as they became available. CAD was not available for SFM.

#### Data Collection and Analysis

In this study, we included all screening examinations performed within 5 years after the start of the program in September

2003. We collected data from all participants who were recalled after screening, as well as the total number of women screened per unit per month. For recalled women, the collected data included patient-related information, date of the examination (and for subsequent screening examinations, the date of the previous screening examination), and reports from the screening radiologists that included mammographic lesion characterization and assessment. If recall led to biopsy, results of histologic examination were included. Cases that were recalled were grouped in three categories on the basis of the reported abnormality: (a) mass or architectural distortion, (b) clustered microcalcifications as only sign, and (c) other.

All performance indicators were computed separately for initial and subsequent screening examinations. The recall rate was computed by dividing the number of recalls by the number of screening examinations. Detection rates were computed by dividing the number of recalled woman in whom cancer was detected by the number of screening examinations. Screening intervals were computed for subsequent screening examinations by taking the period between the current and the previous screening examination. Because screening intervals were somewhat different in the two populations, we computed detection and recall rates 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 and the other facilities that were all mobile.

We compared the breast cancer detection rate, recall rate, and positive predictive value (PPV) for the two screening modalities. Differences in radiologic characteristics of lesions and tumor type (invasive vs ductal carcinoma in situ [DCIS]) were evaluated. Statistical software was used for data analysis (R, version 2.3.1 for Linux; R Foundation for Statistical Computing, Vienna, Austria). Screening outcomes were compared by using Pearson  $\chi^2$  tests. A *P* value of less than .05 was considered to indicate a statistically significant difference. For comparisons of detection performance, a Bonferroni correction was applied, because a

total of six tests were performed to evaluate detection of all cancers, invasive cancers, and DCIS, for initial as well as subsequent screening examinations. A *P* value less than .008 was considered to indicate a significant difference in these comparisons. For testing age and screening interval differences, the independent two-sample *t* test was used.

## Results

During the study period, a total of 367 600 screening procedures were performed, of which 56 518 were con-

ducted with FFDM and 311 082 were conducted with SFM. Of these, 10 307 initial procedures were FFDM examinations and 38 754 were SFM examinations. Refusal of FFDM was extremely rare (less than one per 1000) and could be neglected. A total of 1239 women were recalled in the FFDM group and 4071 in the SFM group. Age of the participants ranged from 50 to 75 years in both groups. The median screening interval was 22.7 months in the group of recalls screened with FFDM and 24.6 months in the group of recalls screened

with SFM. The difference in the interval was significant ( $P < .001$ ). The mean age of recalled participants in the first screening examination was 51.3 years for FFDM and 51.9 years for SFM ( $P < .001$ ). In subsequent screening examinations, the mean age in the two groups was 61.6 and 62.7 years, respectively, ( $P < .001$ ).

Breast cancer was detected in 1927 women, 317 of whom had digital mammograms (Table 1). Cancer detection rates per 1000 women standardized to a 24-month interval were higher with FFDM. In initial screening examinations, the detection rate was .77% with FFDM and .62% with SFM ( $P = .11$ ). In subsequent screening examinations, the respective detection rates were .54% and .49% ( $P = .46$ ).

Film-based screening detection of DCIS was .12% in initial and .08% in subsequent screening examinations. With digital mammography, detection of DCIS increased to .22% ( $P = .015$ ) and .12% ( $P = .007$ ), respectively. The difference is statistically significant for subsequent screening examinations.

The recall rate was significantly higher with FFDM, both in initial screening, where it increased from 2.3% to 4.4% ( $P < .001$ ), and in subsequent screenings, where it increased from 1.2% to 1.7% ( $P < .001$ ). Because of the increase in recalls, the PPV of recall decreased with digital mammography. For first screening examinations, the PPV decreased from 26.8% to 17.4%. For subsequent screening examinations, the PPV decreased from 43.1% to 30.4%.

The radiologic characteristics of lesions on the basis of which the women were recalled are shown in Table 2. It was found that with digital mammography, the fraction of cases recalled on the basis of clustered microcalcifications as only sign increased from 19.0% to 39.3% ( $P < .001$ ) in initial screenings and from 21.6% to 41.2% ( $P < .001$ ) in subsequent screenings. The majority of recalls remained related to the presence of masses, architectural distortion, and asymmetry. The PPV decreased for all lesion types. This decrease was most striking for recalls based on the presence of microcalcifications alone, for

Table 1

### Recall Rates and Screening Performance

Variable	SFM		FFDM		<i>P</i> Value ( $\chi^2$ Test)
	No.	Rate (%) <sup>*</sup>	No.	Rate (%) <sup>*</sup>	
Initial screening examinations <sup>†</sup>					
Screened women	38 754		10 307		
Recalled women	898	2.32	455	4.41	< .001
Breast cancers detected	241	.62	79	.77	.11
Invasive cancers	188	.49	56	.54	.46
DCIS	47	.12	23	.22	.015
Unknown disease	6	.02	0	.00	
Subsequent screening examinations <sup>‡</sup>					
Screened women	272 328		46 211		
Recalled women	3 173	1.17	784	1.70	< .001
Breast cancers detected	1 369	.49	238	.55	.12
Invasive cancers	1 122	.40	175	.40	.96
DCIS	216	.08	51	.12	.007
Unknown disease	31	.01	12	.03	

\* Detection rates for subsequent screening examinations are per 24 months.

<sup>†</sup> PPV of recall was 26.8% for SFM and 17.4% for FFDM.

<sup>‡</sup> PPV of recall was 43.1% for SFM and 30.4% for FFDM.

Table 2

### Radiologic Characteristics of Recall with PPV per Lesion Type

Variable	SFM				FFDM			
	No.	Fraction (%)	Rate (%)	PPV (%)	No.	Fraction (%)	Rate (%)	PPV (%)
Initial screening examinations								
Mass	650	72.4	1.68	24.0	249	54.2	2.42	17.3
Architectural distortion	45	5.0	.12	35.6	15	3.3	.15	26.7
Clustered microcalcifications	171	19.0	.44	31.0	179	39.3	1.73	15.6
Other	32	3.6	.08	34.4	12	2.6	.12	0
Subsequent screening examinations								
Mass	2250	70.9	.83	43.8	419	53.4	.91	34.1
Architectural distortion	118	3.7	.04	44.9	29	.37	.06	41.4
Clustered microcalcifications	685	21.6	.25	38.7	323	41.2	.70	23.8
Other	120	3.8	.04	27.5	13	1.7	.03	15.4



which the PPV decreased from 31.0% to 15.6% ( $P < .001$ ) for initial screenings and from 38.7% to 23.8% ( $P < .001$ ) for subsequent screenings.

Most of the recalled cases due only to microcalcifications that turned out to be malignant concerned DCIS. However, some women with invasive breast cancer presented with no other sign of abnormality than microcalcifications. Of the invasive cancers, 38 were reported with microcalcifications as the only sign of malignancy in the FFDM group and 106 in the group screened with SFM. The fraction of invasive breast cancer cases with microcalcifications alone increased significantly with FFDM, from 8.1% to 15.8% ( $P < .001$ ).

## Discussion

We found that the detection rates trended to be higher with FFDM than with SFM. Significantly more DCIS was found with FFDM. In the initial screening examinations, the detection rate of DCIS almost doubled. This finding confirms results reported in previous studies (4–7). Increased detection of DCIS with FFDM is related to better detection of microcalcifications. In our study, the fraction of recalls based solely on microcalcifications increased from 21.0% to 40.5% with FFDM. This suggests that microcalcifications are depicted better with FFDM. However, another factor is the use of CAD in our study, which most likely resulted in more sensitive detection. The fact that previous studies without CAD also reported increased detection of microcalcifications suggests that this result should not be attributed to CAD alone. It is noted that additional microcalcifications found with FFDM apparently are more difficult to interpret because the PPV associated with microcalcifications strongly decreased.

Results suggest that improvement of detection with FFDM may be more substantial at initial screenings, which include the youngest women in the screening population. At initial screening examinations, the mean age of participants was 51 years and the detection

rate was .77% with FFDM and .62% with SFM ( $P = .11$ ). At subsequent screening examinations, the mean age of the participants was 62 years and the detection rate was .54% with FFDM and .49% with SFM ( $P = .46$ ). These findings would be in accordance with the results of the Digital Mammographic Imaging Screening Trial, or DMIST (2), where a significant increase in performance with digital mammography was found only in younger women, while film-based mammography tended to perform nonsignificantly better for women aged 65 years or older and with fatty breasts (12). It should be noted, however, that the study period covered several screening intervals and many participants underwent more than one digital examination. A smaller effect of FFDM on the detection rate is expected after the first digital screening round, as earlier detection of screening-detected cancers does not have a permanent effect on the detection rate (11). Detection rate only increases because of a smaller proportion of interval cancers.

Recall rates were significantly higher with FFDM. This led to smaller PPVs of recall in initial and subsequent screening examinations. It is noted, however, that recall rates remain relatively low in comparison with other breast cancer screening programs. The largest decrease of PPV with FFDM was seen for microcalcifications. Although we did not have access to data to analyze detailed histopathologic characteristics of the cancers detected, we may assume detection of low-grade DCIS increased with FFDM. Early detection of these cancers may not have much effect on breast cancer mortality (13). However, with FFDM the fraction of invasive cancers visible only because of microcalcifications doubled from 8.1% to 16.4%. This shows that improved detection of microcalcifications with FFDM is beneficial for earlier detection of invasive cancers.

The higher detection found with FFDM may partly be explained by the higher recall rate. The effect of recall rate on screening performance has been studied by Otten et al (11). They estimated the effect of additional recalls in the screen-film setting to be much lower than

what we observed in this study. On the basis of data obtained within the context of the Dutch screening program, they estimated that in subsequent screening rounds increasing the recall rate from 0.9% to 2% would increase detection rate from .42% to .45% because of earlier detection of interval cancers. As in our study the increase in recall rate in subsequent screenings is much less, the expected effect of additional recalls is also smaller in our study. Therefore, it is unlikely that the higher detection with FFDM in this study was caused only by a higher recall rate.

In this study, potential biases due to changes in the screening center or the population, which hampered some of the previous studies, were avoided. By making a concurrent comparison in the same screening center, we ensured that variation in the group of readers or changes in their criteria for recall over time, other than those related to learning to work with FFDM and CAD, did not affect results. It is noted that the group of readers remained very stable during the study period. Since all readers were involved in both FFDM and SFM screenings, the risk of bias due to reading-skill differences was minimized.

Because of the study design, the fraction of women who underwent FFDM was larger in initial screening examinations than in subsequent screenings. Therefore, we analyzed results of initial and subsequent examinations separately. Because all women who underwent digital mammography automatically were assigned to digital mammography at subsequent screenings, we expected a slight difference in the mean age of the groups screened with both modalities. We found that the group screened with FFDM in subsequent screenings was 0.8 year younger on average. Since this difference is small, we do not believe this will have had a substantial effect on our results. It is noted that the effect of age difference on detection would be in favor of SFM, since the incidence is higher in older women.

The screening interval for FFDM was shorter than that for SFM. This was caused by the fact that scheduling was organized in a different way for mobile units, and most of the conventional units were mobile.

Therefore, we computed detection rates for subsequent screenings per standardized screening interval of 24 months. While this compensates for most of the bias caused by the different screening interval, this correction does not take into account that with a larger screening interval cancers can grow longer and will become easier to detect. Since the difference in screening interval was only 2 months, we believe the latter effect to be negligible. If it had some influence, it would be in favor of conventional mammography, where the interval was longer.

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. In particular, the high performance of CAD in detection of clustered microcalcifications is appreciated by radiologists, especially when soft-copy reading is practiced. 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 practiced, and findings of a recent study (14) suggest that single reading with CAD may yield results comparable to those of double reading. In our study, each mammogram was read by two readers who both could use CAD. Thus, results of this study show a combined effect of digital mammography and CAD.

It is noted that the screening approach in this study, and in Europe in general, is different from the practice in the United States and Canada: The screening interval was 2 years, studies were double read, part of the examinations had mediolateral oblique views only, and recall rates were much lower than typical in the United States. One should be aware of this when interpreting results. Detection rates are higher because the 2-year screening interval is higher than the typical interval of 1 year that is common in the United States. On the other hand, detection rates are lower due to interval cancers associated with the longer interval and to lower recall rates. Double reading minimizes perception errors and improves decisions, but most likely reduces the incremental benefit of CAD. Despite these differences, we do

not believe that our main findings strongly depend on the screening context, since these relate to a comparison of cancer detection by using the two techniques rather than to absolute values of performance indicators. It is unlikely that differences in detection strongly depend on the operating point chosen in a screening program, as long as this is the same in both modalities. On the other hand, the increase in recall we found with FFDM-CAD may be related to the low recall in the Netherlands and not translate to screening approaches where recall is much higher.

A limitation of our study design was that the contribution of FFDM and CAD could not be evaluated separately because they were introduced at the same time. Another important limitation was the unavailability of detailed pathology reports, which prohibited reliable analysis of the histologic grades of DCIS. In future research, we will address this issue. The study was not designed as a randomized controlled trial. Assignment of modality was determined according to availability, which was random, and also according to the previous screening, as women who once had undergone FFDM remained in the digital track. As in the initial phase, all FFDM screenings were initial screenings, this led to a slight bias toward younger women being assigned to FFDM. This was visible as a small bias in mean age in the two groups. The effect was judged to be small, and we did not correct for it. Bias would be in favor of SFM, since incidence increases with age. Finally, we mention the effect of multiple screening rounds on the expected screening outcome. When more early-stage cancers are found with FFDM using CAD, this will lead to less-invasive cancers in subsequent FFDM screenings and less interval cancers. Because of incomplete data on interval cancers, we could not investigate this issue.

To our knowledge, our study is the largest to date in comparisons of SFM with FFDM. Results indicate that with the FFDM-CAD combination and double reading, the detection is as good as that with SFM, and detection of clustered microcalcifications and DCIS is improved with FFDM using CAD.

## References

1. Lewin JM, Hendrick RE, D'Orsi CJ, et al. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4945 paired examinations. *Radiology* 2001;218(3):873-880.
2. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;353(17):1773-1783.
3. Skaane P, Young K, Skjennald A. Population-based mammography screening: comparison of screen-film and full-field digital mammography with soft-copy reading: the Oslo I study. *Radiology* 2003;229(3):877-884.
4. Heddson B, Ronnow K, Olsson M, Miller D. Digital versus screen-film mammography: a retrospective comparison in a population-based screening program. *Eur J Radiol* 2007;64(3):419-425.
5. Del Turco MR, Mantellini P, Ciatto S, et al. Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. *AJR Am J Roentgenol* 2007;189(4):860-866.
6. Skaane P, Hofvind S, Skjennald A. 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* 2007;244(3):708-717.
7. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P. 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* 2008;18(1):183-191.
8. Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol* 2009;50(1):3-14.
9. Roelofs AA, Karssemeijer N, Wedekind N, et al. Importance of comparison of current and prior mammograms in breast cancer screening. *Radiology* 2007;242(1):70-77.
10. Fracheboud J, de Koning HJ, Beemsterboer PM, et al. Nation-wide breast cancer screening in The Netherlands: results of initial and subsequent screening 1990-1995 National Evaluation Team for Breast Cancer Screening. *Int J Cancer* 1998;75(5):694-698.
11. Otten JD, Karssemeijer N, Hendriks JH, et al. Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. *J Natl Cancer Inst* 2005;97(10):748-754.
12. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008;246(2):376-383.
13. Duffy SW, Tabar L, Vitak B, et al. The relative contributions of screen-detected in situ and invasive breast carcinomas in reducing mortality from the disease. *Eur J Cancer* 2003;39(12):1755-1760.
14. Gilbert FJ, Astley SM, Gillan MG, et al. Single reading with computer-aided detection for screening mammography. *N Engl J Med* 2008;359(16):1675-1684.