Breast Cancer Screening Results 5 Years after Introduction of Digital Mammography in a Population-based Screening Program

1. Nico Karssemeijer, PhD,
2. Adriana M. Bluekens, MD,
3. David Beijerinck, MD,
4. Jan J. Deurenberg, MD,
5. Matthijs Beekman,
6. Roelant Visser, MS,
7. Ruben van Engen,
8. Annemieke Bartels-Kortland, MD and
9. Mireille J. Broeders, PhD

+ Author Affiliations

1. [1]*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.).*

1. **Address correspondence to** N.K. (e-mail: n.karssemeijer@rad.umcn.nl).

1. **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

Abstract

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

Screen-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 examinations 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. Nonconsultant 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 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^2. 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 χ^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.