Image Quality of Digital Direct Flat-Panel Mammography Versus an Analog Screen-Film Technique Using a Phantom Model

Kathrin Barbara Krug[1]
Hartmut Stützer[2]
Ralf Girnus[1]
Markus Zähringer[1]
Axel Goßmann[1]
Guido Winnekendonk[1]
Klaus Lackner[1]

OBJECTIVE. The objective of our study was to compare the detection and distinguishability of microcalcifications on mammograms obtained with a digital direct flat-panel detector versus an analog system using an anthropomorphic breast phantom.

MATERIALS AND METHODS. Studies were performed with a digital mammography system (Selenia) and an analog mammography system (Mammomat 3). Sixty-five transparent films were used as test specimens. Randomly distributed round and heterogeneous silicate particles (diameter, 100–1,400 μm) and an anthropomorphic scatter body were applied to the films. All radiographs were taken at identical settings and exposures. Six radiologists rated the films and monitor-displayed images independently of each other in random order on a standardized electronic questionnaire.

RESULTS. Interpretations based on monitor reading produced superior results over those based on digital image reading and analog film reading. In 41.1% (95% CI, 38.7–43.5%) of all the monitor readings, 20.2% (18.2–22.2%) of all digital images, and 19.6% (17.6–21.6%) of all analog films, the number of detectable microcalcifications agreed with the gold standard method. The diameter of visible microcalcifications was interpreted correctly in 35.6% (33.2–38.0%) of monitor readings, 19.0% (17.1–21.0%) of digital images, and 21.0% (18.9–23.0%) of analog films; and microcalcification shape was interpreted correctly in 53.8% (51.4–56.3%) of monitor readings, 28.2% (26.0–30.4%) of digital images, and 28.3% (26.0–30.5%) of analog films. Microcalcification number and size were underestimated more frequently than overestimated. Regardless of display medium, accuracy increased proportionately with the diameter of the simulated microcalcifications for all evaluation variables.

CONCLUSION. Digital flat-panel mammography is superior to the analog screen-film method for the detection and morphologic characterization of microcalcifications larger than 200 μm in diameter when the display medium is a monitor.

Keywords: breast cancer, digital imaging, screen-film mammography, mammography, women's imaging, X-ray technology

DOI:10.2214/AJR.05.2006

Received November 15, 2005; accepted after revision May 16, 2006.

[1]Department of Radiology, University of Cologne, Kerpenerstraße 62, Cologne, NRW 50924, Germany. Address correspondence to K. B. Krug (Barbara.Krug@uk-koeln.de).

[2]Department of Medical Statistics, University of Cologne, Cologne, NRW 50924, Germany.

AJR 2007; 188:399–407

0361–803X/07/1882–399

The technologic demands placed on mammographic imaging systems are particularly great. First, a higher local image resolution is required to detect and morphologically characterize microcalcifications in the breast than in other organs. Second, the dynamic range must be broad to enable simultaneous visualization of structures with strong X-ray absorption, such as calcifications, and structures with poor absorption, such as fatty tissue. Over time, various digital imaging systems have been developed: The most important of the currently available commercial technologies include computed radiography systems using photostimulable phosphor plates, digital flat-panel detector radiography with indirect conversion, digital flat-panel radiography with direct conversion, and linear quantum-counting digital X-ray detectors [1–8].

Digital flat-panel detector radiography with indirect conversion takes place in two steps. In the first step, a scintillator layer made of gadolinium oxysulfide or cesium iodide captures X-ray quanta and converts them to light. In the second step, the light is converted to electrical charges in photodiodes made of amorphous silicon. In digital flat-panel radiography using direct conversion, X-ray quanta are converted directly to electrical charges in a layer of amorphous selenium. Direct flat-panel detectors represent one of the more recent technologic approaches and, compared with the other methods, offer the advantages of a higher quantum efficiency because X-ray conversion takes place in just one step and, thanks to the chemical properties of amorphous selenium, larger-area imaging systems are possible.

The results of more recent experimental studies on digital mammography using stored-image and indirect flat-panel techniques have shown that advances in development have allowed the aforementioned technologies to be regarded as at least equivalent to analog screen-film mammography in terms of image quality and diagnostic accuracy [9–14]. A previous phantom study performed using direct flat-panel mammography suggested that direct flat-panel mammography was superior to the analog screen-film method, to digital stored-image radiography, and to indirect flat-panel detector radiography in detecting fibrous structures, although all methods detected microcalcifications and round densities equally well based on a comparably low number of interindividual observations [8]. For the study detailed in this article, we aimed to go beyond this by determining, on a larger sample and using an anthropomorphic model, the following: first, whether the detection rate and morphologic characterization of varyingly configured microcalcifications achievable with direct digital flat-panel mammography (test method) are comparable with the diagnostic profile of the analog screen-film technique (reference method); second, whether monitor interpretation is diagnostically equivalent or superior to film interpretation; and, third, whether within an interobserver comparison results are affected by the previous experience and expertise that the interpreting radiologists have in digital imaging techniques.

Materials and Methods

Imaging Systems

All equipment and machines were serviced before commencement of the study. The full-field digital mammography system (Selenia, Lorad/Hologic) had a double-focus molybdenum anode with a 25-μm filter made of molybdenum used for the test runs, or "takes"; a nominal focal spot size of 0.3 mm for survey images; and a honeycomblike grid. Thanks to its polygonal structure and because the interseptum material is air, the grid absorbs scatter radiation as usable radiation more intensely in all transverse directions. The focus–film distance was 65 cm. The flat-panel detector consisted of a semiconductor layer of amorphous selenium that was placed under a direct current before X-ray exposure. The absorption of X rays caused local charge equalization; the charges were captured in an array of electrodes, storage capacitors, and transistors located behind the selenium layer and were converted to electronic signals. After electronic enhancement and analog-to-digital signal conversion, the digital image data were sent to the imaging computer. The active field of view of the flat-panel detector was 24 × 29 cm, the matrix had an array of 3,328 × 4,096 pixels, and the pixel edge length was 70 × 70 μm, which is equivalent to a nominal local resolution of 7.2 lp/mm. Data were acquired in an 18 × 24 cm format, and 2,560 × 3,328 pixels of the active field of view were used for imaging. Before study initiation, the factory-set image processing variables were optimized by systematic variation and then were not changed during the studies.

The analog mammography unit (Mammomat 3, Siemens Medical Solutions) was equipped with a double-focus molybdenum–tungsten bimetal anode. The nominal focal spot size for survey images was 0.3 mm. The takes were done using a molybdenum filter (up to 40 kV, 0.03 mm). The focus–film distance was 60 cm. Min-R 2190 screens (Eastman Kodak) and single-coated MIN-R 2000 mammography films (Eastman Kodak) with a blue base support were used. The films had a format of 18 × 24 cm, equivalent to a nominal local resolution of 12 lp/mm, and were developed on a daylight load system (Miniloader 2 Plus, Eastman Kodak).

Phantom and Simulated Microcalcifications

On 65 transparent universal laser printer films with a format of 21 × 30 cm (type A P/N 003R96019, Xerox), a circle of 10 cm in diameter was drawn and the circle was divided into four quadrants (I–IV). Each universal laser printer film was then labeled for radiologic visualization by gluing a wire to the film (Fig. 1). After this step, round or heterogeneous silicate particles (composition: SiO_2 [65.0%], Al_2O_3 [0.5–2.0%], Fe_2O_2 [0.15%], MgO [2.5%], CaO [8.0%], Na_2O [14.0%]; diameter classes: 1, 100–199 μm; 2, 200–399 μm; 3, 400–599 μm; 4, 600–800 μm) were scattered within the quadrants by random distribution and were fixed with adhesive tape (Fig. 2). Single quadrants were left out of this particle assignment. The shape and size of the particles on the films were recorded for each quadrant. The number of simulated microcalcifications located in each quadrant was counted under a surgical microscope (Wild M680, Leica) (Tables 1 and 2).

The 65 films (phantoms) with four quadrants each (phantom quadrants) were imaged with a digital mammography system (Selenia) and an analog mammography system (Mammomat 3). A 1.5-cm-thick piece of acrylic plastic sheet (Plexiglas, Degussa) made of polymethylmethacrylate (PMMA) and a similarly thick layer of ground meat placed on the film before X-ray exposures were taken served as scattering bodies. Because the two X-ray systems were located in one spatial unit, the digital and analog images could be produced in direct succession without having to change the scatter-body configuration.

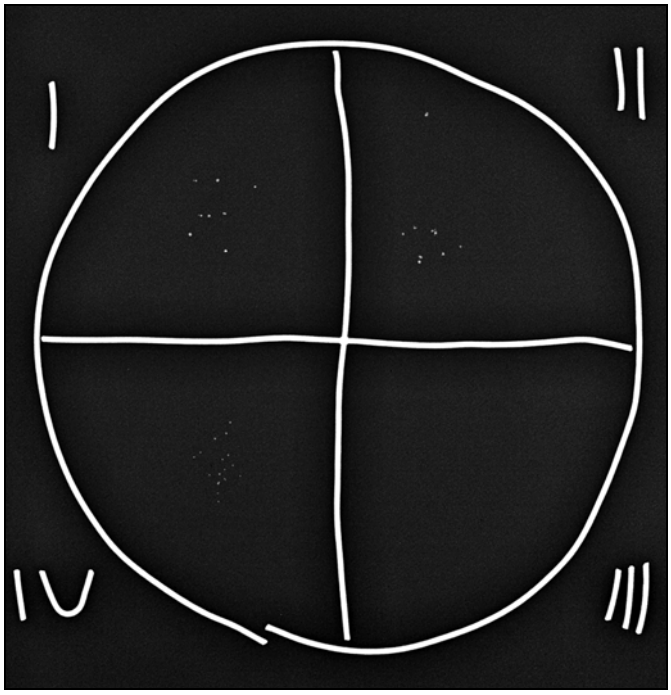
Imaging Technique

All images were obtained in the craniocaudal plane. The same compression setting was used for both methods. Preliminary tests for defining the optimal exposure settings were conducted in which the kilovoltage settings were systematically changed between 20 and 32 kV using the automatic exposure control and a molybdenum filter on the analog mammography system. After these tests, all analog and digital scans were taken at a tube voltage of 29 kV and an exposure tube current of 18 mAs. The half-value layer (HVL), entrance skin exposure, and average glandular dose were 0.40, 1.925 mGy, and 0.593 mGy for the digital system; and 0.41, 1.875 mGy, and 0.588 mGy for the analog system, respectively. The entrance skin exposure was 1.875 mGy for the Mammomat 3 system and 1.925 mGy for the Selenia system, and the average glandular dose was 0.588 mGy for the Mammomat 3 system and 0.593 mGy for the Selenia system. The doses were measured with a calibrated dosimeter (Solidose 300, RTI Electronics AB). The manufacturer states that the measuring error of the dosimeter is ± 5%. This test proved that any minor differences observed between the analog and digital methods with regard to the exposure doses are within the range of error. Thus, the two machines are comparable in terms of overall exposure levels.

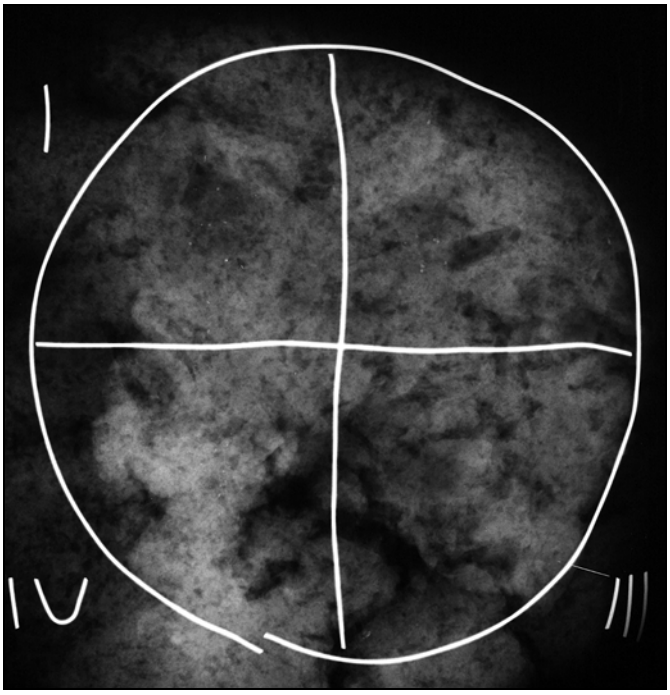
Image Display

The digital imaging data sets were printed on a dry laser printer (DryView 8610, Eastman Kodak) that operates according to the direct thermography principle. The geometric print resolution was 655 dots per inch (DPI), the effective pixel edge length was 36 × 39 μm in a 5,025 × 6,200 pixel matrix, and the contrast resolution had a 12-bit gray-scale depth (4,096 gray levels). An infrared-sensitive photothermographic film with a blue film base coating on one side was used in a 18 × 24 cm format (DryView laser imaging film, Eastman Kodak).

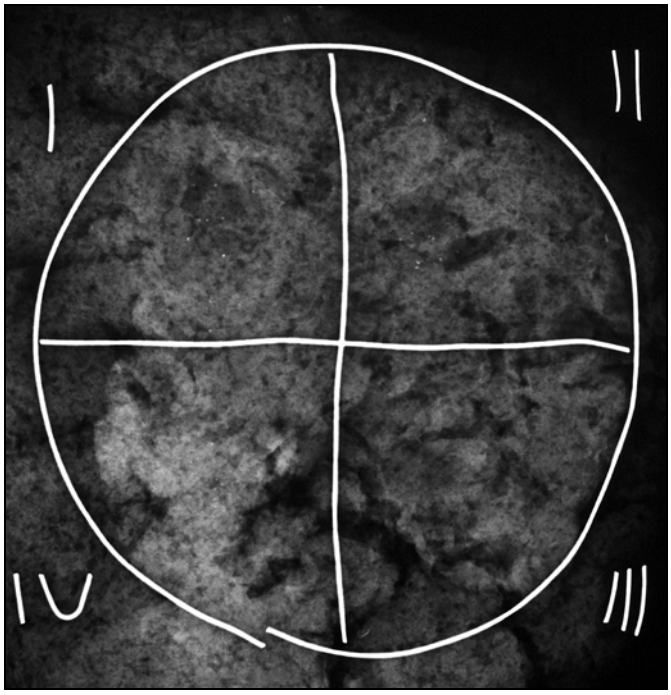
The digital system was connected to a viewing workstation (Selenia Softcopy Workstation, Lorad/Hologic) that was equipped with soft-copy reading software (MeVis BreastCare, MeVis Bremen). The graphics controllers (BarcoMed 5MP1H, Barco NV) operated with an internal 12-bit gray-scale depth equivalent to 4,096 levels of gray and a 10-bit digital-to-analog converter (DAC) that converts the signal for a display depth equivalent to 1,024 levels of gray. The two monitors (Barco, Barco NV) had an image field size of 30 × 40 cm in diameter with a line resolution of 2,048 × 2,560 pixels (effective pixel edge length, 147 × 156 μm). Whenever a scan with a format of 18 × 24 cm was completely displayed on the monitor, a reduction in the geometric resolution of approximately a factor of 0.8 resulted, which is equivalent to 7 lp/mm. For a 1:1 reproduc-



A



B



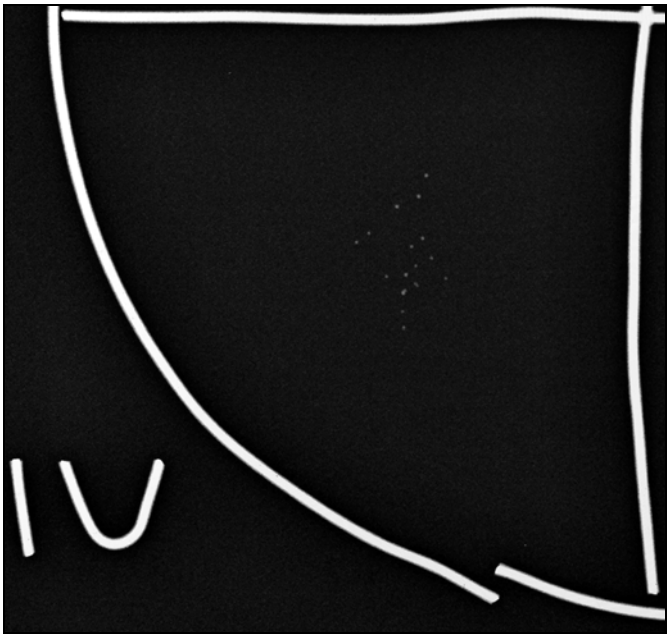
C

Fig. 1—Radiographic study of universal laser printer film. Four quadrants were labelled by wire fixed to film for better radiologic visualization. Quadrant I contains round (class of number, 3; class of size, 3), quadrant II contains round (class of number, 3; class of size, 3), and quadrant IV contains round (class of number, 3; class of size, 2) particles; quadrant III contains no silicate particles. Number of microcalcifications was characterized as follows: class 0, no microcalcifications; 1, 1–4 microcalcifications; 2, 5–9 microcalcifications; 3, 10–19 microcalcifications; 4, 20–39 microcalcifications; or 5, > 39 microcalcifications. Diameter classes were class 1, 100–199 μm; 2, 200–399 μm; 3, 400–599 μm; and 4, 600–800 μm.
A, Digital image of film taken without scatter body for demonstration purposes.
B, Analog film image.
C, Digital film image.
(Fig. 1 continues on next page)

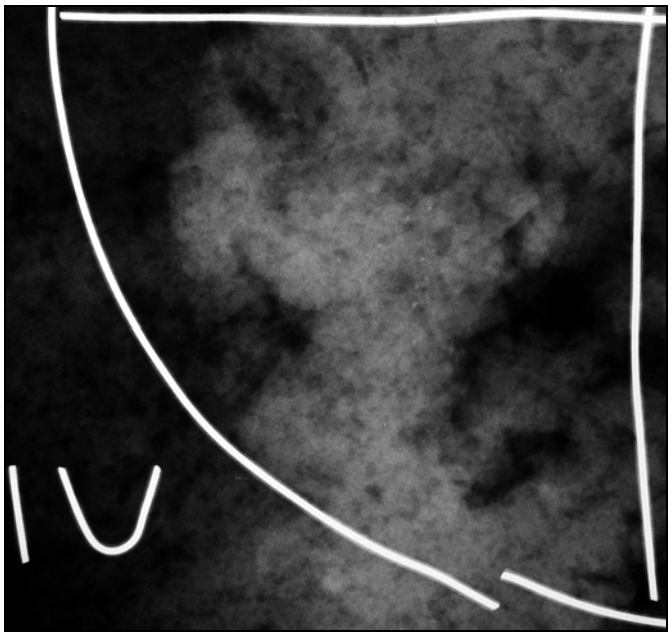
tion of the digital image data set, the image was cut off slightly on one edge, producing an effective display resolution of approximately 8 lp/mm. Through bicubic interpolation, the optical magnifying glass option allowed factor-2 zooming of the digital image data set.

Image Evaluation

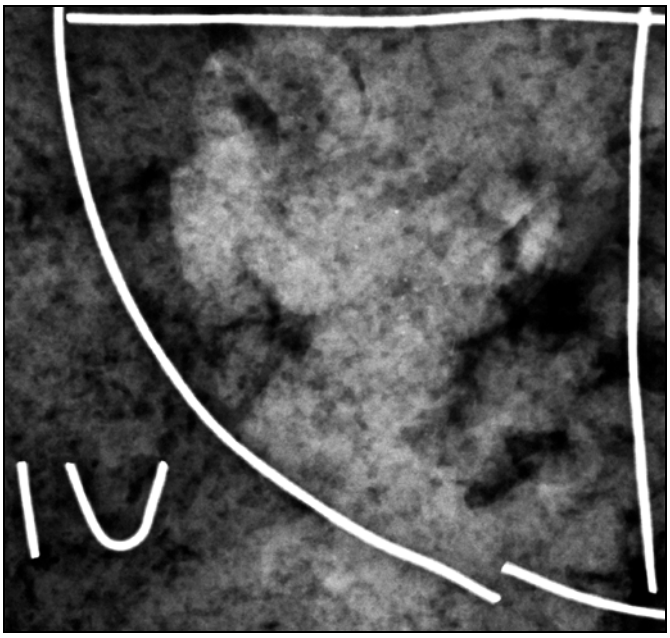
A team of six board-certified radiologists rated the analog films, digital scans, and monitor-displayed digital images in randomized order and independently of each other. Each observer specializes in diagnostic radiology and at the time of the study had 3–15 years' experience in interpreting routine diagnostic analog mammograms and 1.5 years' experience interpreting digital mammograms. Observers 1 and 2 each had completed many years of scientific research in digital radiography. Observers 1 and 3 had the most experience with routine diagnostic mammography.



D



E



F

Fig. 1 (continued)—Radiographic study of universal laser printer film. Four quadrants were labelled by wire fixed to film for better radiologic visualization. Quadrant I contains round (class of number, 3; class of size, 3), quadrant II contains round (class of number, 3; class of size, 3), and quadrant IV contains round (class of number, 3; class of size, 2) particles; quadrant III contains no silicate particles. Number of microcalcifications was characterized as follows: class 0, no microcalcifications; 1, 1–4 microcalcifications; 2, 5–9 microcalcifications; 3, 10–19 microcalcifications; 4, 20–39 microcalcifications; or 5, > 39 microcalcifications. Diameter classes were class 1, 100–199 μm; 2, 200–399 μm; 3, 400–599 μm; and 4, 600–800 μm.
D, Magnified cutout of quadrant IV shown in **A**.
E, Magnified cutout of quadrant IV shown in **B**.
F, Magnified cutout of quadrant IV shown in **C**.

The observers' notations were entered on structured questionnaires. The observers were required to rank the number of simulated microcalcifications visible as follows: class 0, no microcalcifications; 1, 1–4 microcalcifications; 2, 5–9 microcalcifications; 3, 10–19 microcalcifications; 4, 20–39 microcalcifications; or 5, > 39 microcalcifications. Size was ranked according to class as follows: 0, not applicable (i.e., no microcalcifications in the phantom quadrant); 1, 100–199 μm; 2, 200–399 μm; 3, 400–599 μm; or 4, ≥ 600 μm. The size rankings were based on the maximum size of simulated microcalcifications observed for each cluster. Shape was defined as round, heterogeneous, or not applicable (NA) if no microcalcifications were visible in the phantom quadrant. Data regarding diameter and shape were related to the respectively largest silicate particle visible in one quadrant. For orientation, each observer was given templates produced by the analog and digital techniques and showing a perfect depiction of all shapes and sizes without the scatter layer of ground meat (Fig. 2). The images were evaluated in darkened rooms. While reviewing images on viewboxes, the observers were allowed to use illumination lamps and magnifying glasses.

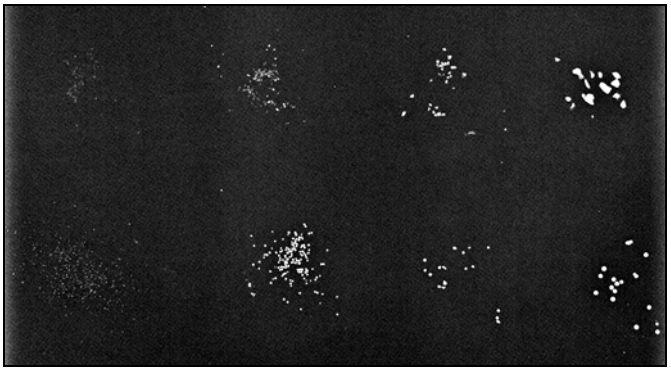


Fig. 2—Digital image shows all heterogeneous (*top row*) and round (*bottom row*) silicate particles ranging from size class 1 (left) to 4 (right), respectively, without scatter body. Diameter classes were class 1, 100–199 μm; 2, 200–399 μm; 3, 400–599 μm; and 4, 600–800 μm.

TABLE 1: Number of Round Simulated Microcalcifications as a Function of Size

No. of Round Simulated Microcalcifications	Maximum Diameter (μm)				Total
	100–199	200–399	400–599	≥ 600	
1–4	1	3	1	0	5
5–9	2	8	16	5	31
10–19	9	11	16	2	38
20–39	8	2	2	0	12
≥ 40	20	1	0	0	21
Total	40	25	35	7	107

Note—Gold standard = total coverage in all 260 quadrants, 53 of which were not covered with silicate particles.

TABLE 2: Number of Heterogeneous Simulated Microcalcifications as a Function of Size

No. of Heterogeneous Simulated Microcalcifications	Maximum Diameter (μm)				Total
	100–199	200–399	400–599	≥ 600	
1–4	0	0	1	1	2
5–9	1	1	9	9	20
10–19	3	13	18	15	49
20–39	5	6	4	0	15
≥ 40	11	2	1	0	14
Total	20	22	33	25	100

Note—Gold standard = total coverage in all 260 quadrants, 53 of which were not covered with silicate particles.

Statistical Analysis

The notations made by the six observers for the ranked variables' number and size and for the qualitative variable shape of detected microcalcifications, separated according to the three display media, were compared with the experimentally predefined constellation in the phantom quadrants. In a first step, the correct interpretations of the six observers were cross-tabulated in tables to summarize the results. Kappa statistics were used to indicate the degree of agreement between the gold standard and the ratings. Second, the differences between the ranks of the observers' interpretations and those of the predefined experimental constellations of microcalcifications were used to describe the direction and magnitude of the error (deviation from reference) for each quadrant for the first two variables. Assuming that each of the four quadrants of a film counted as an independent rating, each observer produced 4×65 interpretations, or a total of 260 interpretations, per display medium that were available for deviation analysis. With regard to the variables' number and size, the sign of the error allowed us to distinguish between observers' underestimations (negative sign) and overestimations (positive sign) with regard to the predefined phantom constellation on the scale of the respective rating entity.

For better graphic illustration of interobserver variability, the deviation distributions were presented in box plots (number and size) or contingency tables (shape) separated according to observers and display medium and independently of the simulated microcalcification size. To describe the effect that simulated microcalcification size had on the quality and accuracy of microcalcification detectability, the analyses of the variables' number and shape were divided into classes for size of the actual experimentally predefined constellations and were then repeated.

Using global comparisons of all three display media (Kruskal-Wallis test), the distribution of the differences in the observers' notations from the experimentally predefined actual constellation (gold standard) with respect to the display medium were analyzed on the basis of the aforementioned errors in direction and size and were supplemented by explorative paired comparisons between the different types of scans (Mann-Whitney tests). The p values were not corrected for multiple test scenarios. Therefore, the interpretation of any differences observed was only of an explorative nature.

Results

Monitor interpretation proved superior to both the interpretation of digital images and the review of analog films (Table 3 and Figs. 1 and 2). In this respect, the number of microcalcifications counted irrespective of their size was in agreement with the gold standard count in 41.1% (95% CI, 38.7–43.5%) of all quadrants when rated on monitor interpretation, in 20.2% (18.2–22.2%) when rated on digital images, and in 19.6% (17.6–21.6%) when rated on analog films. As measured by the kappa value (κ), the difference between perfect agreement and agreement expected by chance for monitor interpretation ratings was approximately 25.3% and has been accounted for in the agreement between observers' interpretations and the gold standard; for digital images and analog films, kappa was less than 1%. When all observers and quadrants were considered, the diameter of the microcalcifications was correctly identified in 35.6% (33.2–38.0%) using monitor interpretation ($\kappa = 0.193$), in 19.0% (17.1–21.0%) on digital images ($\kappa = -0.010$), and in 21.0% (18.9–23.0%) on analog films ($\kappa = 0.013$) (Table 4). Microcalcification shape was classified correctly using monitor interpretation in 53.8% (51.4–56.3%), digital images in 28.2% (26.0–30.4%), and analog films in 28.3% (26.0–30.5%); the measurement of agreement calculated to kappa values of 0.338, –0.028, and –0.014, respectively.

Figures 3A and 3B present the absolute magnitudes of the observer classification deviations (error) from the gold standard. Observers 1 and 2, the two radiologists most experienced in digital techniques, produced better results for monitor interpretations and digital

TABLE 3: Correct Classifications as a Function of the Number of Silicate Particles

Gold Standard: No. of Silicate Particles (Class)	No. (%) of Correct Classifications			No. (%) of Quadrants Evaluated
	Monitor Interpretation	Digital Images	Analog Films	
0	234 (73.6)	128 (40.3)	142 (44.7)	318 (100.0)
1–4	12 (28.6)	3 (7.1)	5 (11.9)	42 (100.0)
5–9	148 (48.4)	66 (21.6)	62 (20.3)	306 (100.0)
10–19	207 (39.7)	113 (21.6)	91 (17.4)	522 (100.0)
20–39	21 (13.0)	4 (2.5)	5 (3.1)	162 (100.0)
≥ 40	16 (7.8)	0 (0)	0 (0)	204 (100.0)
Total	638 (41.1)	314 (20.2)	305 (19.6)	1,554[a] (100.0)
κ	0.253	–0.002	–0.007	
p	< 0.0001	0.844	0.550	

Note—Correct readings of the six observers summarized for all quadrants of the 65 laser printer films, with 1,560 observations per display medium.

[a]One quadrant was not assessable for all six observers.

TABLE 4: Correct Classifications as a Function of Size

Gold Standard Diameter of Silicate Particles (μm)	No. (%) of Correct Classifications			No. (%) of Quadrants Evaluated
	Monitor Interpretation	Digital Images	Analog Films	
No particles	234 (73.6)	128 (40.3)	142 (44.7)	318 (100.0)
100–199	77 (21.4)	62 (17.2)	40 (11.1)	360 (100.0)
200–399	59 (20.9)	72 (25.5)	77 (27.3)	282 (100.0)
400–599	73 (17.9)	33 (8.1)	33 (8.1)	408 (100.0)
≥ 600	112 (58.3)	3 (1.6)	35 (18.2)	192 (100.0)
Total	555 (35.6)	298 (19.1)	327 (21.0)	1,560 (100.0)
κ	0.193	–0.010	0.013	
p	< 0.0001	0.395	0.258	

Note—Correct readings of the six observers summarized for all quadrants of the 65 laser-printer films, with 1,560 observations per display medium.

image interpretations than the other four observers. The number of microcalcifications was underestimated more frequently than overestimated—a finding that particularly applied to the comparatively inexperienced observers 4–6. Regardless of the display medium, less experienced observers tended to underestimate microcalcification size, whereas the more experienced observers 1, 2, and 3 tended to overestimate microcalcification size.

The accuracy of classifications increased in correlation with the diameter of the silicate particles. Figures 4A and 4B present the results separately for the variables' number (Fig. 4A) and size (Fig. 4B) according to size classes 1–4 of the gold standard as summarized across all quadrants and observers. Although errors in the diagnostic classification of microcalcifications with a diameter of 100–199 μm (class 1) deviated the greatest from the predefined phantom constellation, the classifications mainly agreed with the gold standard for microcalcifications of size class 4 (diameter, ≥ 600 μm). An advantage of monitor interpretation was observed for the variable number of detectable microcalcifications starting from a real diameter of 200 μm and for the variable size starting from a real diameter 400 μm.

An explorative analysis of observer classification error with respect to the number of visible microcalcifications showed a global difference among the three display media pooled across all observers for both the variable number (Table 5) and the variable size (Table 6) of detectable microcalcifications (Kruskal-Wallis test, $p < 0.05$). When pooled across all observers, the Mann-Whitney tests produced a difference ($p < 0.05$) only for the paired comparisons analog films versus monitor interpretation and digital images versus monitor interpretation, which in most cases was likewise confirmed in the analyses conducted separately for the individual observers.

Discussion

Experimental studies are able to give only a simplified, more or less one-sided reproduction of a complex medical situation. They cannot make predictions about how a diagnostic test method will perform in routine clinical practice that will account for all technologic, biologic, physiologic, and medical factors and influences. Experimental designs, however, do allow an estimation of the diagnostic accuracy to be expected of a test method when it is used in clinical routine. In this context, it is not to be expected that every lesion visible on a phantom image will also be detectable under clinical conditions when superimposing anatomic structures are present. Conversely, an object that could not be visualized under the ideal circumstances of an experimental study will presumably not be detectable under clinical conditions either. This statement is particularly true for microcalcifications that proved to have a minimum detectable diameter of 130–200 μm in experimental setups [15, 16].

The quality of a mammography method is determined by the capability of the imaging system to convert the attenuation of the X-ray beam in the breast as dimensionally accurately as possible into a perceivable image. The most important variables affecting image quality are sharpness, contrast, and noise. Local resolution and contrast resolution of an imaging system are interdependent: An improvement in local resolution causes the lower number of X-ray quanta received per pixel to lead to a poorer contrast-to-noise ratio and thus to poorer contrast resolution and vice versa. The mutual dependency of these two factors is expressed in terms of digital quantum efficiency (DQE). The DQE, which is 50–80% for digital mammography systems, indicates how many photons hitting the detection plane are used for imaging and is determined by comparing the local dose measured behind different filters with the exposure dose and includes every step in the process—from radiation exposure to when the detection system receives the radiation pattern. What are not accounted for are process steps downline from the signal acquisition, such as the analog-to-digital conversion of the signals, technical noise, image processing, and the performance of the image reproduction system (film, monitor).

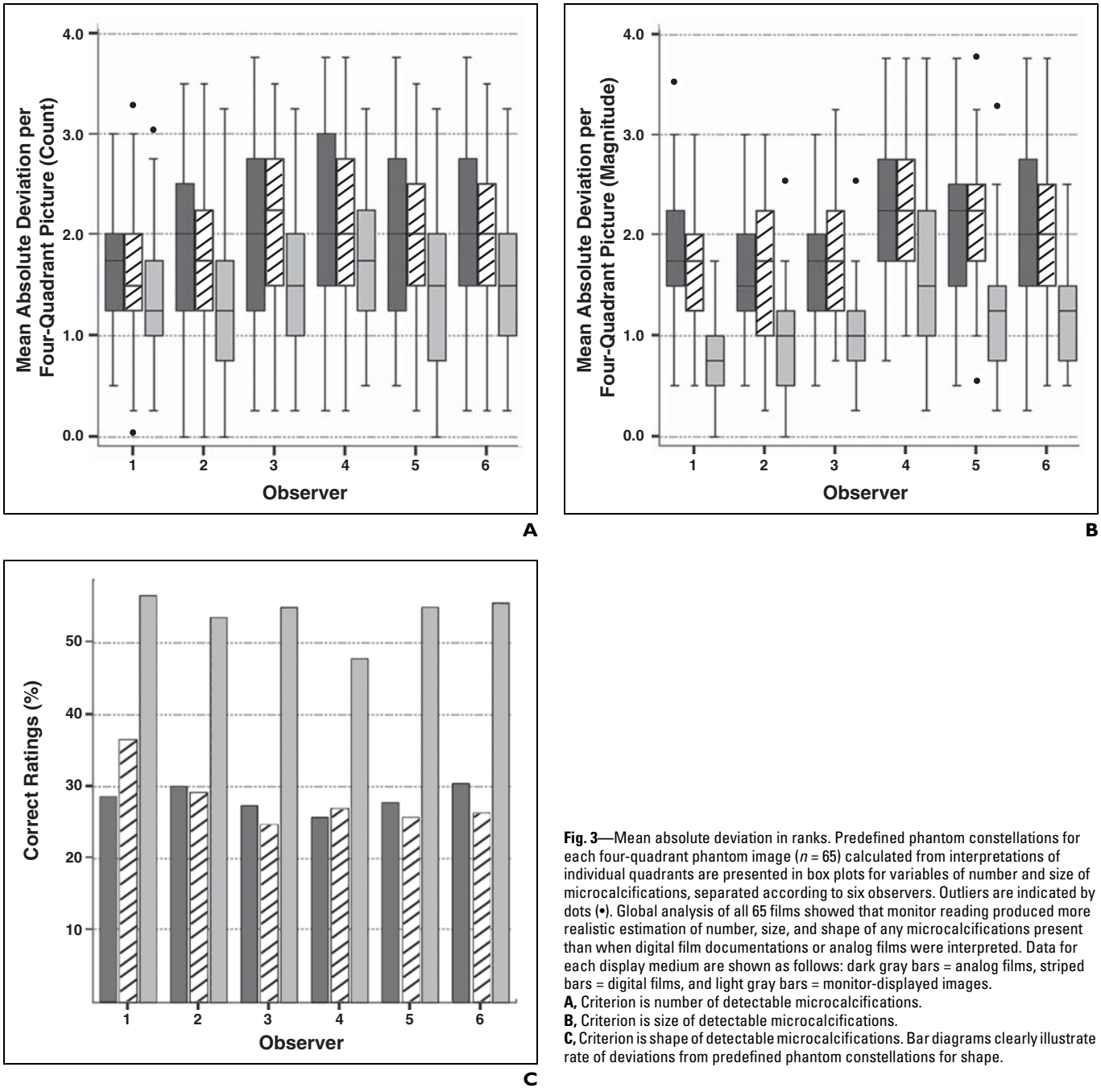


Fig. 3—Mean absolute deviation in ranks. Predefined phantom constellations for each four-quadrant phantom image (n = 65) calculated from interpretations of individual quadrants are presented in box plots for variables of number and size of microcalcifications, separated according to six observers. Outliers are indicated by dots (•). Global analysis of all 65 films showed that monitor reading produced more realistic estimation of number, size, and shape of any microcalcifications present than when digital film documentations or analog films were interpreted. Data for each display medium are shown as follows: dark gray bars = analog films, striped bars = digital films, and light gray bars = monitor-displayed images.
A, Criterion is number of detectable microcalcifications.
B, Criterion is size of detectable microcalcifications.
C, Criterion is shape of detectable microcalcifications. Bar diagrams clearly illustrate rate of deviations from predefined phantom constellations for shape.

Digital mammography must be measured against the gold standard of analog screen-film mammography, which has proved its merits in clinical routine over decades. Its advantages include low investment and operating costs, a high local resolution of up to 20 lp/mm for high-contrast objects, and simple presentation options offered by the viewbox [2, 17]. Disadvantages involve the fact that a trade-off has to be made between image contrast and radiation exposure. As a result of the sigmoidal gradation curve, under- or overexposure can produce inadequate image contrast. The DQE of screen-film systems is comparably low, between 20% and 40%, as a result of the low absorption of the X rays on intensifier films [2].

Experimental studies offer a feasible method of estimating the diagnostic accuracy of digital mammography, given that methodologically high-quality clinical studies on the qualitative and quantitative detection of microcalcifications are difficult to perform owing to the problems associated with the gold standard and, not least, for reasons of radiation safety.

The present study showed that direct flat-panel mammography (film interpretation) is equivalent to analog screen-film mammogra-

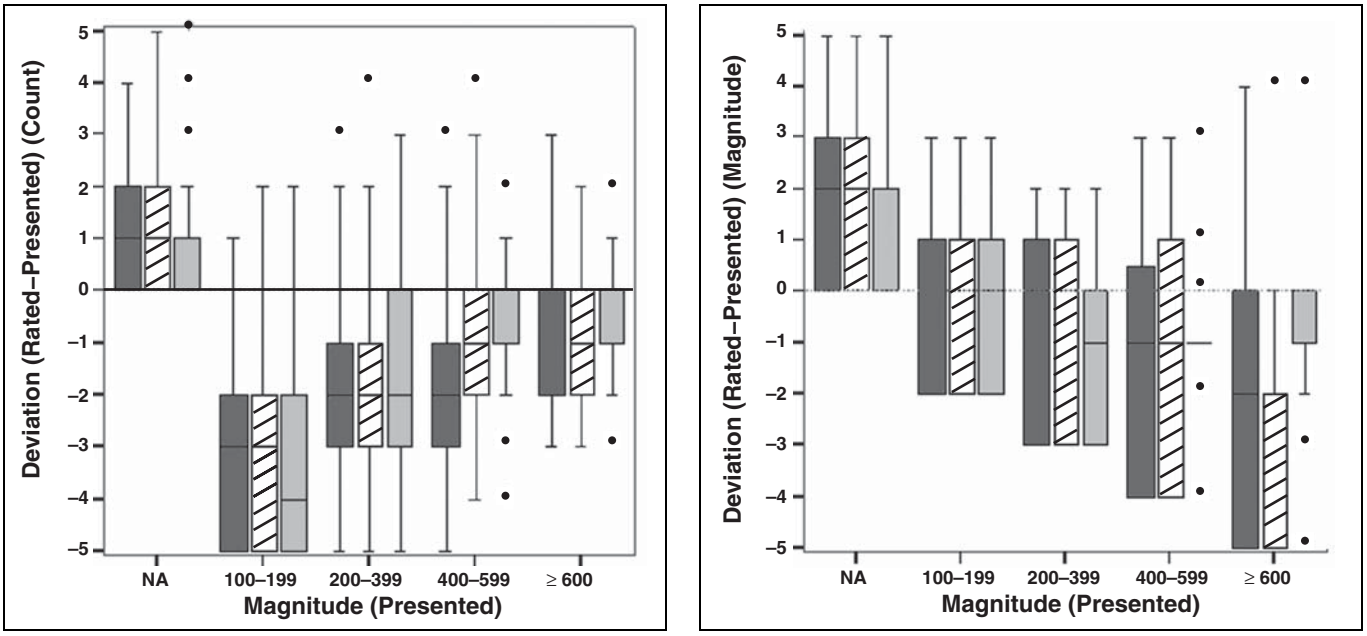


Fig. 4—Box plots of deviations in six observers' classifications from actual value of gold standard as function of presence and size of artificial microcalcifications applied to individual quadrants on 65 films. Outliers are indicated by dots (•). NA = not applicable (i.e., no microcalcifications in phantom quadrant). Data for each display medium are shown as follows: dark gray bars = analog films, striped bars = digital films, and light gray bars = monitor-displayed images.
A, Criterion is number of detectable microcalcifications as function of their presented number.
B, Criterion is size of detectable microcalcifications as function of their presented size.

phy in the detection and morphologic characterization of microcalcifications, thus confirming the results of more recent phantom studies on digital storage-plate and indirect flat-panel mammography [9–14, 16, 17]. In addition to this finding, the following were shown for direct flat-panel mammography: First, monitor interpretation of digital image data sets was diagnostically superior to the analog and digital film interpretations for microcalcifications greater than 200 μm in diameter. Second, not only was the number of simulated microcalcifications as accurately estimated as with analog and digital methods of film interpretation, but the description of their shape and size was also comparably reliable. Third, a positive correlation existed between the diagnostic accuracy of imaging systems and size of simulated microcalcifications. Finally, methodologic expertise in digital radiography was more crucial for the correct perception of microcalcifications than experience in routine clinical diagnostics. The latter unquestionably plays a role in differential diagnostics with regard to the interpretation of once-detected microcalcifications.

Given the technologic characteristics discussed earlier in this article, the superiority of monitor interpretation versus analog film interpretation and digital film interpretation of large microcalcifications might be attributable to the type of digital image data processing and the light intensity of the monitors. The decline in accuracy of monitor-viewed interpretations when microcalcifications smaller than 200 μm in diameter were involved can be explained by the pixel size of the monitors, an observation also described by Rong et al. [16] in their experimental study of simulated microcalcifications with diameters between 112 and 160 μm imaged with digital screen-film, charge-coupled device (CCD), and indirect flat-panel systems. The superior diagnostic results for microcalcifications with a diameter greater than 200 μm achieved by monitor interpretation are clinically relevant insofar as this finding proves that radiologic malignancy criteria—that is, grouping trends and polymorphisms—can be interpreted more reliably when interpretations are performed using a high-resolution monitor rather than analog or digital films.

The power of the present study is partially limited by the lack of data on detector dose, considering that image quality and radiation dose are interdependent [18]. Even if the same anode filter combination is used, the same tube voltage and same milliampere-seconds (mAs) settings will not yield a matching detector dose, particularly given the difference in focus–film distance of 5 cm, the different construction of the tubes, and the different grids.

The present research investigated the diagnostic performance profiles of only mammography systems used in the high-contrast range. A similarly designed study on the low-contrast range is currently in progress. More recently, standardized test phantoms have become available for digital mammography that allow quantification of the interplay of the characteristic physical variables of contrast, sharpness, and noise in a single test sequence. Test specimens, such as the Mammo-152 phantom (Wellhöfer) or the PAS1054 phantom (Artimis), provide robust tools for the testing and medicophysical expert evaluation of digital mammography systems. Nevertheless, they appear unsuitable for interobserver comparisons by virtue of the uniformity of their investigative items.

In conclusion, an anthropomorphic model showed that the detection rate and morphologic characterization of microcalcifications achieved with direct digital flat-panel mammography are comparably as effective as the diagnostic profile of the analog screen-film technique, that monitor interpretation is diagnostically superior to film interpretation

TABLE 5: Results of Explorative Analysis of the Deviations in Observers' Classifications of the Number of Microcalcifications from the Experimentally Predefined Actual Value

	p					
Statistical Test	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5	Observer 6
Kruskal-Wallis test	**0.001**	**0.004**	0.322	0.535	**0.023**	0.443
Mann-Whitney tests						
Analog vs digital film	**0.001**	0.290	0.842	0.819	0.707	0.746
Analog film vs monitor interpretation	**0.001**	**0.001**	0.133	0.260	**0.008**	0.224
Digital film vs monitor interpretation	0.588	**0.034**	0.285	0.453	**0.037**	0.346

Note—These results show a global difference between the three display media for three of the six radiologists (Kruskal-Wallis test). The Mann-Whitney tests produced remarkable differences ($p < 0.05$) in the comparison of analog and digital film interpretations for observers 1 and 2, the two who were most experienced in digital radiography; in the comparison of analog film and monitor interpretations for three observers; and in the comparison of digital film and monitor interpretations for two of the six observers. The *p* values were not corrected for multiple test scenarios. Values shown in **boldface** indicate $p < 0.05$.

TABLE 6: Results of Explorative Analysis of the Deviations in Observers' Classifications of Size of the Microcalcifications from the Experimentally Predefined Actual Value

	p					
Statistical Test	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5	Observer 6
Kruskal-Wallis test	**0.031**	**< 0.0005**	**< 0.0005**	0.349	0.785	0.792
Mann-Whitney tests						
Analog vs digital film	0.141	0.172	0.720	0.933	0.812	0.592
Analog film vs monitor interpretation	**0.008**	**< 0.0005**	**< 0.0005**	0.263	0.823	0.895
Digital film vs monitor interpretation	0.235	**< 0.0005**	**< 0.0005**	0.165	0.419	0.521

Note—These results show a global difference between the three display media for three of the six radiologists (Kruskal-Wallis test). The Mann-Whitney tests produced remarkable differences in the comparison of analog with digital film interpretations, the comparison of analog film and monitor interpretations for three observers, and the comparison of digital film and monitor interpretations for two of the six observers. The *p* values were not corrected for multiple test scenarios. Values shown in **boldface** indicate $p < 0.05$.

for detecting large microcalcifications, and that the evaluating radiologists' experience with digital radiography affects diagnostic accuracy. Consequently, digital mammography in conjunction with monitor interpretation should be the preferred method for clinical diagnostics and breast cancer screening in the future.

Acknowledgments

We thank Claudia Morgenroth and Daniele Steinhaus-Wittig for their evaluations of the analog films, digital scans, and monitor-displayed digital images. We also thank Gerald Haupt (formerly of the Clinic and Policlinic for Urology of the University of Cologne and currently of the Department of Urology, St. Vincentius-Krankenhaus, Speyer) for counting the number of simulated microcalcifications located in each quadrant.

References

1. Haus AG, Yaffe MJ. Screen-film and digital mammography. *Radiol Clin North Am* 2000; 38:871–898
2. Hermann K-P, Funke M, Grabbe E. Physical and technical aspects of digital mammography [in German]. *Radiologe* 2002; 42:256–260
3. James JJ. The current status of digital mammography. *Clin Radiol* 2004; 59:1–10
4. Pisano ED, Kuzmiak C, Koomen M. Perspective on digital mammography. *Semin Roentgenol* 2001; 36:195–200
5. Pisano ED, Yaffee MJ. Digital mammography. *Radiology* 2005; 234:353–362
6. Pisano ED, Gatsonis C, Hendrick E, et al. Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group: diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005; 353:1773–1783
7. Schulz-Wendtland R, Wenkel E, Aichinger U, et al. Film-screen mammography versus digital storage plate mammography: hard copy and monitor display of microcalcifications and focal findings—a retrospective clinical and histologic analysis [in German]. *Röfo* 2003; 175:1220–1224
8. Schulz-Wendtland R, Hermann K-P, Böhner C, et al. Phantom study for the detection of simulated lesions in five different digital and one conventional mammography system [in German]. *Röfo* 2004; 176:1127–1132
9. Diekmann F, Diekmann S, Bick U, et al. Comparing the visualization of microcalcifications with direct magnification in digital full-field mammography vs. film-screen mammography [in German]. *Röfo* 2002; 174:297–300
10. Diekmann S, Bick U, von Heyden H, Diekmann F. Visualization of microcalcifications on mammographies obtained by digital full-field mammography in comparison to conventional film-screen mammography [in German]. *Röfo* 2003; 175:775–779
11. Obenauer S, Hermann K-P, Schorn C, Funke M, Fischer U, Grabbe E. Full-field digital mammography: a phantom study for detection of microcalcifications [in German]. *Röfo* 2000; 172:646–650
12. Obenauer S, Hermann K-P, Schorn C, Fischer U, Grabbe E. Full-field digital mammography: dose-dependent detectability of breast lesions and microcalcinosis [in German]. *Röfo* 2000; 172:1052–1056
13. Schulz-Wendtland R, Aichinger U, Säbel M, Böhner C, Dobritz M, Bautz W. Experimental studies on image quality in conventional film screen system, digital phosphor storage plate mammography in magnification technique and digital mammography in CCD-technique [in German]. *Röfo* 2000; 172:965–968
14. Schulz-Wendtland R, Aichinger U, Lell M, Kuchar I, Bautz W. Experiences with phantom measurements in different mammographic systems [in German]. *Röfo* 2002; 174:1243–1246
15. Cowen AR, Launders JH, Jadav M, Brettle DS. Visibility of microcalcifications in computed and screen-film mammography. *Phys Med Biol* 1997; 42:1533–1548
16. Rong XJ, Shaw CC, Johnston DA, et al. Microcalcification detectability for four mammographic detectors: flat-panel, CCD, CR, and screen-film. *Med Phys* 2002; 29:2052–2061
17. Yip WM, Pang SY, Yim WS, Kwok CS. ROC analysis of lesion detectability on phantoms: comparison of digital spot film mammography with conventional spot film mammography. *Br J Radiol* 2001; 74:621–628
18. Hermann K-P, Obenauer S, Grabbe E. Radiation exposure in full-field digital mammography with a flat-panel x-ray detector based on amorphous silicon in comparison with conventional screen-film-mammography [in German]. *Röfo* 2000; 172:940–945