

FULL-FIELD DIGITAL MAMMOGRAPHY

A Technology Assessment

INTRODUCTION

The California Technology Assessment Forum has been asked to conduct a review of the scientific literature on the safety and efficacy of full-field digital mammography for the diagnosis of breast cancer.

BACKGROUND

Breast Cancer

In 2005, there were an estimated 211,240 new cases of invasive breast cancer in the U.S. and an estimated 40,410 deaths from this cancer. This represents approximately 31% of all new cancer cases in women and 15% of all cancer deaths in women.¹ In addition to invasive breast cancer, 58,490 new cases of *in situ* breast cancer will be diagnosed in women in 2005. Cancer of the breast is the most common form of cancer in women. Every American woman is estimated to have a one in eight chance of developing breast cancer at some time during her life.

Film Mammography (aka Screen-Film Mammography)

Mammography is one of the most commonly performed radiological procedures. Approximately 55% of women over the age of 40 have had a mammogram within the past year. Thus, more than 32 million mammograms are performed annually in the U.S.

There has been a tremendous amount of research on the efficacy of screen-film mammography to reduce mortality from breast cancer. Eight large clinical trials have randomized over 260,000 women and followed them for more than 15 years.²⁻¹⁶ While the interpretation of these results remain hotly debated¹⁷⁻²⁴, the consensus in recent systematic reviews and practice guidelines is that screen film mammography reduces breast cancer mortality for women 50-75 years of age and probably for women ages 40 years and older.²⁵⁻²⁹ A series of complex models based on U.S. data all support the hypothesis that mammography has contributed to a reduction in breast cancer mortality in the U.S.³⁰

However, there are a number of known limitations to mammography. For women under the age of 50 years, the benefits of mammography are smaller and take longer to become apparent. Many factors contribute to this including the lower incidence of breast cancer among younger women and the lower sensitivity of mammography in this group.^{31, 32} The primary explanation for the lower sensitivity of mammography appears to be that younger women have a higher proportion of breast tissue that is mammographically dense.^{31, 33-35}

In the U.S., the standard for reporting mammography results is the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) which uses six categories to describe the likelihood of cancer (1 = negative; 2 = benign finding; 3 = probably benign finding; 4 = suspicious abnormality; 5 = highly suggestive of malignancy; 0 = need additional imaging evaluation). BI-RADS also defines four categories to describe breast density (1 = almost entirely fat; 2 = scattered fibroglandular densities; 3 = heterogeneously dense; 4 = extremely dense).

Digital Mammography (aka Full-Field Digital Mammography)

In contrast to traditional mammography, full-field digital mammography captures the x-ray image of the breast digitally. The images can either be printed on film for review (hard copy) or read on computer monitors (soft copy). Most other radiographic studies are either inherently digital (computed tomography, magnetic resonance imaging) or have transitioned to digital imaging because of the ease of storage, retrieval and transmission of digital images. Mammography has continued to use film because of the need to examine fine details of the image in order to identify breast cancers when they are still small. Until recently, digital acquired images and the workstations to view them, have not had sufficiently high resolution to satisfy the demands of radiologists interpreting mammograms.

The potential advantages of digital mammography are based on the separation of the image acquisition, image presentation and storage of the images. Separating these three tasks allows each to be optimized. Digital image acquisition may improve the signal to noise ratio of x-ray detection over a wider range of intensities, compared to film.³⁶⁻³⁸ Computer aided enhancement of the images at the computer workstations may also improve the accuracy of mammographic interpretation.³⁹ Each manufacturer includes proprietary image enhancement algorithms that may have unique advantages and disadvantages. Digital enhancement, with increased contrast resolution, has particular promise in improving detection of low contrast lesions in radiographically dense breasts. Furthermore, implementation of computer aided detection methods would be simplified, as there would no longer be the separate step of digitizing the film mammograms.⁴⁰ Digital mammography has the potential to improve workflow by allowing for electronic transmission, storage and retrieval of the images. This may also allow for improved interpretation of mammograms as the images can be obtained locally, but sent to a central location for interpretation by experts at centers specializing in

mammographic interpretation. Finally, digital storage and retrieval may increase the likelihood that comparison images from prior mammograms are available to aid the radiologist in interpreting the mammogram.

There are at least four different digital mammography systems that are commercially available. All four systems were included in the pivotal Digital Mammographic Imaging Screening Trial (DMIST) described below. The Trex Digital Mammography System, which was also used in the DMIST, has been withdrawn from the market. Each system has potential advantages and disadvantages.

The system used in the majority of published studies of digital mammography is the Senographe 2000D (General Electric Medical Systems). It uses a flat panel phosphor system with thin film transistor switches used to identify signal from each pixel. The individual detector element measures 100 μm in a matrix of 1,920 by 2,304 pixels.

The Senoscan system (Fischer Medical Imaging) also uses a phosphor system, but uses millions of optical fibers to carry light from the phosphor detector to a charge-coupled device. The detector is long and narrow—to acquire the image, both the x-ray beam and detector scan across the breast together. This approach takes longer than the Senographe system to acquire each image, but the device is simpler and may require less radiation to capture a comparable image. The individual detector element measures 54 μm in a matrix of 4,096 by 5,625 pixels.

The Computed Radiography for Mammography System (Fuji Medical Systems) uses a phosphor screen with photostimulable luminescence similar to that used in many other radiographic applications. X-ray absorption causes electrons to be trapped in a crystal lattice. The image is then read -using a red laser beam, which frees the electrons to drop back to their resting state. The energy released corresponds to the x-ray absorption at the pixel location. The individual detector element measures 50 μm in a matrix of approximately 4,700 by 6,000 pixels. This system has not yet received FDA approval.

Finally, the Selenia Digital Mammography System (Lorad/Hologic) uses selenium instead of phosphor to absorb x-rays. Electrodes on either side of the selenium surface record the charge generated by x-ray exposure. The individual detector element measures 70 μm in a matrix of approximately 3,000 by 4,000 pixels.

Technology Assessment (TA)

TA Criterion 1: The technology must have the appropriate regulatory approval.

There are several FDA approved FFDM systems:

Senographe 2000D (General Electric Medical Systems) received FDA approval on 1/28/2000. SenoScan FFDM System (Fischer Imaging Corporation) received FDA approval on 9/25/2001. The Lorad Digital Breast Imager (Hologic, Inc.) received FDA approval on 3/15/2002. The Siemens Mammomat Novation DR FFDM System (Siemens Medical Solutions USA, Inc) received FDA approval on 8/20/2004.

TA Criterion 1 is met.

TA Criterion 2: The scientific evidence must permit conclusions concerning the effectiveness of the technology regarding health outcomes.

The Medline database, Cochrane clinical trials database, Cochrane reviews database and the Database of Abstracts of Reviews of Effects (DARE) were searched using the keywords mammography and digital. These were cross-referenced with the keyword human. The search was performed for the period from 1966 through October 2005 and identified 786 articles. The bibliographies of systematic reviews and key articles were manually searched for additional references. The abstracts of citations were reviewed for relevance and all potentially relevant articles were reviewed in full. In order to be included in this systematic review, articles had to compare the results of digital mammography with film mammography performed on at least 250 patients for screening studies and 50 patients for diagnostic tests. Studies had to report test characteristics based on histologically confirmed cancer diagnoses, ideally with at least one-year follow-up from the mammogram to ensure that negative results represent true negatives. Multiple publications described five studies⁴¹⁻⁴⁷ comparing the two technologies for screening mammography and four studies⁴⁸⁻⁵¹ focusing primarily on diagnostic mammography. A number of other clinical studies were reviewed⁵²⁻⁵⁹ but not included because they lacked controls, did not include sufficient data to evaluate test characteristics, compared image quality rather than clinical outcomes or compared different image processing algorithms. The studies of screening mammography are summarized in Table 1 and diagnostic studies are summarized in Table 2.

Ideally, randomized clinical trials would compare digital to film mammography using breast cancer mortality as the primary outcome. However, such studies would require hundreds of thousands of women to be randomized and followed for at least five to ten years. Such studies are unlikely to be performed. A

reasonable surrogate is to evaluate the relative sensitivity and specificity of digital versus film mammography performed independently in the same population. Digital mammography offers the potential for both improved sensitivity and specificity. However, if it slightly improves sensitivity while sacrificing specificity, the overall harms from increased false positive results and increased biopsies are likely to outweigh the increased detection rate as the vast majority of women presenting for screening mammography do not have breast cancer (approximately 1 in 200). Other potential benefits from digital mammography include lower radiation exposure and more convenient storage and retrieval of mammograms.

Level of evidence: 1 - 5

TA Criterion 2 is met.

TA Criterion 3: The technology must improve the net health outcomes.

The primary outcomes of interest are the relative sensitivity and specificity (measured by call back rates and biopsy rates) of digital mammography compared with film mammography. Ideally, these measures are summarized using receiver operating characteristics (ROC) curve analysis. The area under the ROC curve (AUC) summarizes the overall characteristics of a diagnostic test.

Screening Mammography

Lewin et al. evaluated 6,736 sets of mammograms from 4,489 women who presented for screening mammography at two university medical centers in the U.S.^{41, 42} Women were eligible for the study if they were at least 40 years old and did not have breast implants. The average age of the participants was 55.6 years. All digital mammograms were obtained within three days of the film mammograms (91% on the same day by the same technologist). Digital mammography was performed using a prototype of the Senographe 2000D (General Electric Medical Systems). Digital mammograms were interpreted on a prototype workstation that had lower resolution and fewer software tools than the FDA approved commercial system. Board certified radiologists experienced in mammography interpreted the film and digital images independently. The radiologists interpreted the mammograms using standard BI-RADS categories and a probability of malignancy (0-100) for use in ROC curve analyses. BI-RADS 0, 4 and 5 results were considered positive for the calculation of the sensitivity and specificity. Each subject was followed for one year to determine whether she developed breast cancer. A total of 42 cancers were diagnosed by mammography and eight additional cancers were diagnosed during follow-up. Digital mammography was

less sensitive than film mammography (54% vs. 66%, $p>0.1$), but was more specific (89% vs. 85%) and had a lower recall rate (12% vs. 15%). The AUCs were not significantly different, although digital mammography was lower than film mammography (0.74 vs. 0.80, $p=0.18$).

Strengths of this study include the independent reading of all mammograms by trained radiologists, a standard protocol, appropriate one year follow-up for outcomes and the use of appropriate statistical analyses (sensitivity, specificity and ROC curves). The major weakness of the study is the lack of power to detect a difference in sensitivity. An absolute difference of 12 points in sensitivity is clinically significant. In this study, film mammography detected 22% more cancers than digital mammography. The delay in diagnosis in these women could result in worse clinical outcomes from treatment of their breast cancers. The higher specificity of digital mammography resulted in a lower recall rate and lower biopsy rate, which would be clinically important if sensitivity was not sacrificed. Another weakness of the study is the use of a prototype system for the digital mammography. The improved resolution and software used in the commercial system will likely yield different test characteristics. The results of this study may not generalize to the commercial system.

Skaane et al. in the Oslo I Study compared digital to film mammography in 3,683 women screened as part of the Norwegian Breast Cancer Screening Program.⁴⁵ They also collected data on 6,249 women evaluated using screen-film mammography alone, during the same screening round. Women aged 50-69 years were eligible (mean age of participants: 58.2 years). Digital and film mammography were performed on the same day. Digital mammography was performed using the commercial version of the Senographe 2000D. Eight radiologists with at least four years of experience in screening mammography interpreted the mammograms. Two radiologists interpreted both film and digital images independently. A five point rating scale was used to estimate the probability of cancer (1=normal or benign, 2=probably benign, 3=indeterminate, 4=probably malignant, 5=malignant). If one of the two readers categorized the mammogram as 2 or higher, it was reviewed in a consensus conference. Mammographic interpretations requiring further imaging or biopsy at the consensus conference were considered positive. All others were considered negative. The diagnosis of cancer ($n=31$) was only made on the basis of evaluation of positive results. For participants with negative mammograms, breast cancers developing in the 12 months following mammography were not considered in these analyses. Thus, the specificities calculated in this study are likely to be somewhat higher than those calculated on the basis of 12 months of follow-up. In this study, digital mammography was less sensitive than film mammography (74% vs. 90%, $p=0.23$), less specific (96% vs. 97%) and had a higher recall rate (4.6% vs. 3.5%). No ROC analysis was performed. It is useful to compare these results with those from the non-study population screened by film mammography alone in the same screening round in Norway. The non-study population had a lower recall rate (2.6% vs. 3.5% for

film mammography in the study and 4.6% for digital mammography) and a lower cancer detection rate (0.40% vs. 0.71% for film mammography in the study and 0.54% for digital mammography), compared with the study population. This suggests that the extra scrutiny from the paired reading, plus consensus conference, led to improved sensitivity for breast cancer at the expense of a higher recall rate.

As with the prior study,⁴¹ the authors did not have sufficient power to demonstrate a large and clinically important difference (16%) in sensitivity between the two technologies. The lower sensitivity and specificity of the digital technology, while not statistically significant, is concerning. Their methods for interpretation (double read plus consensus conference) does not reflect usual clinical practice, even in Norway, and this was reflected in the higher detection rate and call back rate in the study, compared with usual practice in a similar population. The authors note that the low level of experience of radiologists in the study with soft-copy reading and the lack of a dedicated room with minimal extraneous light may have biased the study against digital mammography.

Skaane and Skjennald⁴⁴ also used a different study design to address the same question in a second study in Norway (Oslo II Study). They randomized 25,263 to either film or digital mammography using a two-to-one randomization stratified by age group (45-49 years and 50-69 years). The authors excluded 352 women who were not screened as randomized. They note that three cancers were diagnosed in this group. Digital mammography was performed using the Senographe 2000D. As in the prior study, two readers interpreted each mammogram using the same five-point scale and a consensus conference was held for all mammograms when at least one reader categorized the mammogram as not normal or benign. There was no follow-up for interval cancers and no imaging with a second modality, so sensitivity and specificity could not be calculated. The authors reported recall rate, cancer detection rate, positive predictive value and median tumor size as the primary outcomes. In the 45-49 year age group, 3,012 women had digital mammography and 7,607 had film mammography. The recall rate was higher in the digital group (3.7% vs. 3.0%, p NS) as was the cancer detection rate (0.27% vs. 0.22%, p = 0.69). The positive predictive value (7.1% vs. 7.4%, p NS) and tumor size (10 mm. vs. 11 mm, p not reported) were similar in the two groups. In the 50-69 year age group, 3,985 women had digital mammography and 10,304 had film mammography. Again, the recall rate was higher in the digital group (3.8% vs. 2.5%, p<0.05) as was the cancer detection rate (0.83% vs. 0.54%, p = 0.053). The positive predictive value (21.6% vs. 22.1%, p NS) and tumor size (15 mm. vs. 13 mm, p not reported) were similar in the two groups.

This study corrected some of the issues that were of concern in the Oslo I Study. Seven of the eight radiologists had experience with soft copy reading of digital films and the digital mammograms were interpreted in a dedicated room. In this study the cancer detection rate was higher in the digital group,

suggesting that the problems with sensitivity seen in the Oslo I study had been overcome with reader experience and an improved work environment. However, the study designs are quite different. In the Oslo I study, each person serves as her own control. However, in the Oslo II study, there are two independent samples. Randomization should guarantee the equal distribution of risk factors and cancers in the two groups; however there is clear evidence that randomization was violated in the study. For example, women with breast implants who were randomized to film mammography were imaged with digital mammography because the authors felt that digital mammography was better than film. Ideally, all women with breast implants would have been excluded from the study prior to randomization. Instead, those randomized to film were excluded and those randomized to digital mammography were included. Apparently, no attempt was made to conceal the allocation to groups or to blind investigators, staff or the participants. Table 1 is presented comparing the distribution of breast cancer risk factors by randomization group. Finally, the clear violation of the intent-to-treat principle calls the validity of their results into question. After randomization, they excluded participants for a variety of reasons resulting in at least six cancers being dropped from the analyses. At a minimum, the intention to treat results should have been presented along with their per protocol results. Thus, from a methodological perspective, this is a poor quality study.

A small study⁴⁶ by Yamada et al. compared digital to film mammography in 480 women in Japan using the Senographe 2000D system. Only two cancers were detected and they were found by both digital and film mammography. The recall rate was higher in the digital group (4.2% vs. 2.9%) as seen in both of the Oslo studies. The small number of participants and the lack of follow-up limit the interpretation of this study. Furthermore, the digital mammograms were read as hard copy (laser printed films) rather than as soft copy at dedicated workstations. Thus, the potential advantage arising from computer-aided enhancement of the images was lost.

Finally, DMIST, a large multi-center trial funded by the National Cancer Institute was designed to address the deficiencies of the prior studies. The study was designed to be large enough to detect clinically important but small differences in sensitivity and specificity between digital and film systems. The primary aim was to compare the diagnostic accuracy of digital mammography with that of film mammography when used for screening asymptomatic women. There were many secondary aims based on hypothesized reductions in false-positive mammography results, higher throughput cost savings and subgroups with expected greater benefits from the digital technology, as well as comparing the diagnostic accuracy of the available digital systems. Pisano et al. recently published the primary results with some subgroup analyses.^{43, 47} Investigators enrolled 49,528 women from 33 academic and community practice sites in Canada and the U.S. All women presenting for screening mammograms were eligible unless they complained of a breast mass, had nipple discharge, had breast implants, might be pregnant, had a history of

breast cancer treated with lumpectomy and radiation, had undergone mammography within the prior 11 months or would not be available for follow-up. Women were randomized to have either digital or film mammography performed first. Both were performed by the same technologist with positioning, dose and compression held as constant as possible. One reader interpreted each mammogram independent and blinded; from the radiologist interpreting the mammogram obtained using the other modality. A total of 153 radiologists interpreted the study mammograms. All radiologists had at least eight hours of training in digital mammography including the use of the soft-copy display, if applicable, and met national standards for mammographic interpretation, but were not required to have special training or expertise in mammography. All mammograms were interpreted using four different scales: 0-100% probability of malignancy, the BI-RADS assessment categories, a seven-point scale for probability of malignancy and a five-point scale on the need for the patient to return for diagnostic work-up.

From the 49,528 women enrolled in the trial, 42,760 (87%) had data available for the primary analysis. The primary reasons for exclusion from the analysis were lack of follow-up information (4,339), protocol violations at one institution (1,489) and indeterminate cancer status (474). The average age of women in the analysis data set was 54.9 years and 84% were white. A total of 254 breast cancers were diagnosed within 365 days of study entry and 335 cancers were diagnosed within 455 days. The primary outcome was the comparison between the AUC for digital (AUC 0.78) and film mammography (AUC 0.74). The difference (0.03, 95% confidence interval -0.02 to 0.08) favored digital mammography, but was not statistically significant ($p=0.18$). The performance of digital mammography was significantly better than film mammography for women under the age of 50 years, women with heterogeneously dense or extremely dense breasts and for women who were not yet post-menopausal. For example, the AUC for digital mammography in women under the age of 50 was 0.84, compared with an AUC of 0.69 for film mammography ($p=0.002$). The most commonly used criteria for the sensitivity and specificity of mammographic screening is based on a positive test defined by a BI-RADS assessment score of 4 or 5 and cancer outcomes defined over one year (365 days) of follow-up. Using this definition, digital mammography had non-significantly higher sensitivity (70% vs. 66%, $p=0.37$) with equal specificity (92% vs. 92%, $p=0.74$). For women under the age of 50 years, the sensitivity was significantly higher for digital mammography (78% vs. 51%, $p=0.002$) with identical specificity (90% vs. 90%, $p=0.89$). Recall rates and biopsy rates were identical in the two groups (Table 1). The authors reported that the AUC for digital mammography did not differ significantly from film mammography according to race, risk of breast cancer and type of digital machine, although no data were presented.

DMIST appears to definitively demonstrate that digital mammography is equivalent to film mammography when used to screen asymptomatic women for breast cancer. The large number of women dropped from

the analysis due to protocol violations and incomplete follow-up are concerning, but are unlikely to substantially alter the findings of the study. As hypothesized, digital mammography performed best, relative to film mammography, in younger women with denser breasts. Unfortunately, the authors did not present data in the corresponding older patients with less dense breasts. Film mammography may offer some advantages in that population, though it is difficult to assess without the data.

Placed in the context of the four earlier studies,^{41, 44-46} it is apparent that without sufficient attention to training and the work environment, digital mammography may be both less sensitive and less specific than film mammography. In the two published studies that led to FDA approval of digital mammography systems,^{41, 49} the AUC was lower for the digital systems, although the differences were not statistically significant. However, with careful training and attention to quality control, results with digital mammography should be equivalent to film mammography.

TA Criterion 3 is met for screening mammography.

Diagnostic Mammography

The four published studies comparing digital to film mammography in patients referred for diagnostic mammography are summarized in Table 2. The primary difference from screening mammography is the population undergoing the test. Most patients referred for diagnostic mammography may have a palpable lump, nipple discharge or have had an abnormal screening mammogram and are being referred for further imaging.

Venta et al.⁵¹ published the first cohort comparing the accuracy of digital mammography using the Senographe 2000D to that of film in a diagnostic population. They recruited all women scheduled for diagnostic mammography who were at least 40 years old and did not believe they were pregnant. Of the 991 patients invited to participate, 692 (70%) enrolled in the study. All patients were imaged with the same three views by the same technologist with both digital and film mammography. A radiologist interpreted the film images during the visit in order to determine further management of the patient. To limit the radiation dose to the patient, any additional images were performed using film mammography alone. A second radiologist independently reviewed the digital images blinded to the film mammography interpretation. The interpretations were classified using BI-RADS categories. The digital and film interpretations were considered to be in agreement if both radiologists assigned the same category or one assigned BI-RADS 1 and the other assigned BI-RADS 2 or one assigned BI-RADS 4 and the other assigned BI-RADS 5. Partial agreement occurred when one assigned BI-RADS 3 and the other assigned BI-RADS 1 or 2. Disagreement occurred when one assigned BI-RADS 1, 2 or 3 and the other BI-RADS 4 or 5. A total of 18 cancers were

diagnosed. The sensitivity of digital mammography was lower than that of film mammography (72% vs. 89%, p NS). The authors highlight the low rate of disagreement between the digital and film interpretations: 4% (50/1,147) analyzed on a per breast basis. However, this is not the appropriate statistic to focus on. The better measure of agreement between the two interpretations is the kappa statistic (agreement beyond what would be expected by chance). Kappa was only 0.20. While there is no general rule, many statisticians consider kappa less than 0.40 as poor, from 0.40-0.59 as fair, from 0.60-0.74 as good as and greater than 0.74 as excellent. Thus, the two interpretations did not agree well beyond what might be expected by chance. Another way to understand how the disagreement rate can be misleading is to assume that screening mammography was correct for all examinations (BI-RADS 4 or 5 for the 18 cancers and 1 or 2 for the rest). If the digital mammograms were all read as BI-RADS 1 or 2, the disagreement rate would have been 18/1,147 or 1.6%, even though the sensitivity for cancer was 0%. The primary strength of this study was the enrollment of a consecutive sample of all patients presenting for diagnostic mammography. Unfortunately, a substantial number of women declined participation in the study, limiting somewhat the generalizability of the results. Additionally, the small number of cancers in the study gives it low power for evaluating differences in the sensitivity for cancer between the two technologies. The substantially lower sensitivity of digital mammography for cancer and the poor agreement in interpretation between the two technologies (kappa=0.20) raises concerns about the equivalence of digital mammography to film mammography in a diagnostic population.

The study by Becker et al. had a somewhat different design.⁴⁸ The authors present data on the yield of stereotactic biopsy for microcalcifications guided by either digital or film mammography. The study is a case series with historical controls. They reviewed all cases of breast biopsies performed for lesions with microcalcifications as the only abnormality. Between 1993 and 1997, a film mammography unit guided the biopsies. Between 1999 and 2000, a digital unit guided the biopsies. The primary outcome was the percentage of biopsy specimens with microcalcifications seen on x-ray of the specimen or on microscopic review. Among the 111 cases performed with digital guidance, microcalcifications were obtained in 96% of the specimens. Among the 121 cases performed with film guidance, microcalcifications were obtained in 93% of the specimens. Open biopsy was avoided in 74% of the digital cases and in 70% of the film cases. There were also no significant differences in the number of needle passes per lesion, although, again, the trend favored the digitally guided procedures. The improvements seen with the digital unit were not statistically significant and they may simply have represented general improvement in core biopsy technique with time as they were performed an average of four to five years later than the biopsies guided by film. On the other hand, the equivalent yield supports the use of the digital unit for this indication.

The study published by Fischer et al.⁵⁰ compared digital (Senographe 2000D) to film mammography for the classification of microcalcifications initially identified by digital mammography. In 35 patients, 37 clusters of microcalcifications were initially classified as BI-RADS 2 (n=7), BI-RADS 4 (n=7) or BI-RADS 5 (n=23). Biopsy of these 37 clusters revealed 21 cancers. The remaining 20 patients had clusters that were stable over two years of follow-up and were considered to not represent cancers. The authors reported the average sensitivity and specificity from four radiologists who read both the digital and film mammograms. When BI-RADS 1-2 were considered negative and BI-RADS 3-5 considered positive, digital mammography was slightly more sensitive (95% vs. 92%) and specific (41% vs. 39%) than film mammography. If the cut point was set at the more conventional BI-RADS 1-3 considered negative and BI-RADS 4-5 considered positive, they were equally sensitive (68% vs. 68%), but digital mammography was slightly more specific (81% vs. 76%). Even though these results were encouraging, the study was too small to provide definitive results on the relative sensitivity and specificity of the two forms of mammography. For example, I estimated the 95% confidence intervals for the sensitivity of 68% to range from 44% to 86% and the 95% confidence interval for the difference in sensitivity between digital and film mammography to range from -39% to +39%, even though the absolute difference was 0. Additionally, it would have been helpful if the authors had presented the kappa statistic comparing the interpretations and an analysis comparing the AUCs.

Finally, Cole et al.⁴⁹ published an analysis of the data submitted in support of the FDA application for the Fischer Sensoscan Digital Mammography System. They reported data from a complicated sample of patients done in three phases. All women presenting for “problem solving” mammography who were older than 21 years and did not think they were pregnant were eligible for participation. Phase 1 included 560 women who were either recommended for biopsy, had abnormal film mammograms or had symptoms leading to diagnostic mammography. After one year of follow-up, only 25 of these patients were diagnosed with breast cancer. In Phase 2, an unknown number of patients scheduled for biopsy identified an additional 101 cancers. In Phase 3, mammography case sets from an unknown number of women recruited to other clinical trials who met the entry criteria for this study yielded an additional 15 cases of cancer. At the time of the diagnostic visit, women underwent film mammography as clinically indicated and then underwent digital mammography. The original films were sent to the University of North Carolina for the study. The digital data was sent to the Fischer Imaging Corporation where a technologist with five years experience in digital mammography manually adjusted the contrast and brightness of the digital images, printed them and sent them to the University of North Carolina. Of the 676 cases, 247 were selected for this study: 120 from a screening visit and 126 from a diagnostic visit.

The authors report that there were 136 benign diagnoses and 111 cancers identified. The benign diagnoses were randomly selected from the set of benign mammograms. In the abstract, the authors stated that all

cancers were included, although the total number of cancers in Phases 1 to 3 appears to add up to 141. Eight radiologists interpreted the images according to a five-point scale: 1 definitely not malignant; 2 probably not malignant; 3 possibly malignant; 4 probably malignant; and 5 definitely malignant. BI-RADS assessment was not used. The readers all read different sets of films as the original film mammograms because the original site requested their return. The primary analysis compared the average AUC for digital mammography (AUC 0.715) to that of film (AUC 0.765). The difference between the two did not achieve statistical significance, but was close: -0.05, 95% CI -0.101 to 0.002. Indeed, the authors' last sentence of the article states: "As most of the confidence interval is negative, it is possible that a study with more power would show the Fischer SenoScan as inferior to screen-film mammography." The poorer performance of digital mammography compared with film mammography was also reflected in lower average sensitivity (66% vs. 74%, p not reported) although the specificity was higher for digital mammography (67% vs. 60%, p not reported).

There are a number of limitations of this study. Only half of the films chosen for the study represent diagnostic mammograms and the results for the screening and diagnostic films were not presented separately. The phased recruitment of subjects was unusual, inconsistent and could have been subject to bias. Furthermore, the manual adjustment of the digital images by an experienced technologist who may not have been blinded to the case status, does not correspond to clinical practice with the commercial unit, which uses an automatic image-processing algorithm prior to printing the films. Additionally, enrollment was partially based on findings identified on film mammography, which may have biased the results in favor of film mammography. Radiologists interpreting the mammograms used printed films from the digital mammography (hard copy) rather than evaluating them on a high-resolution digital monitors (soft copy). Soft copy display allows the reader to adjust image contrast, magnification and use other image processing algorithms that may improve the accuracy of the interpretations.

In summary, the best quality study⁵¹ comparing digital to film mammography reported poor agreement in the interpretations across technologies ($\kappa = 0.20$) and a large absolute difference in sensitivity for cancer that favored film mammography, although the number of cancers was very low. The only study evaluating a large number of cancers⁴⁹ also found that the sensitivity of digital mammography was lower than that of film mammography by a wide margin, although statistical comparisons were not reported. A small study reported equivalent outcomes when evaluating microcalcifications and a second reported good yield when stereotactic biopsy was guided by digital imaging. However, because of their limited populations and small sizes, neither of these studies was sufficient to demonstrate that overall outcomes with digital imaging are equivalent to film, when used for diagnostic mammography.

TA Criterion 3 is not met for diagnostic mammography.

TA Criterion 4: The technology must be as beneficial as any established alternatives.

The established alternative to digital mammography is film mammography. All of the studies discussed under TA Criterion 3 compare digital to film mammography. For screening mammography, the early results were mixed with some studies reporting higher specificity for digital mammography, but lower sensitivity and others reporting higher sensitivity, but lower specificity. All of the early studies used the Senographe 2000D (General Electric). The two studies from Norway suggested that there is an important learning curve with digital mammography, with improved diagnostic accuracy after more experience with digital mammography and the use of dedicated space designed for optimal use of high-resolution workstations. It is difficult to translate the results of the Norwegian studies to the U.S. because the approach to reading the mammograms (two independent readers followed by consensus conference for any abnormalities) was quite different than the usual practice in the U.S. A further remarkable difference is the low recall rate achieved in the Norwegian studies, compared with the U.S. studies. The methods used in the DMIST trial more faithfully represent usual care in the U.S. A large number of radiologists at 33 centers across the U.S. and Canada evaluated screening mammograms individually. More patients were examined in DMIST than in all of the other trials combined. Digital mammography was performed using five different technology platforms and according to the report, there were no significant differences between them. However, no data were presented on the number of mammograms assessed using the different platforms, nor were any specific data presented on the relative accuracy of the five platforms in the study. Appropriate statistical comparisons were made and the overall results indicated that digital mammography was equivalent to film mammography with the trend favoring digital mammography. Subgroup analyses supported the *a priori* hypothesis that digital mammography should perform better than film when evaluating younger women with denser breast tissue. It is unclear why the results of DMIST were better than those of the earlier studies. There was exceptional attention paid to quality control with a study physicist visiting each site twice yearly to reinforce calibration and other quality control procedures. There was also an attempt to keep the radiation dose equivalent between the digital and film images. Prior studies may have attempted to use a lower dose of radiation to obtain the digital images because one of the proposed advantages of digital imaging was the ability to obtain quality images with lower radiation exposure. Additionally, better imaging processing algorithms and more efficient training of radiologists in the interpretation of digital images may also have influenced the results.

Unfortunately, the same cannot be said for diagnostic mammography. The available studies are relatively small, generally of poor quality and the sensitivity of digital mammography for breast cancer detection was generally lower than that of film mammography in this setting. Low sensitivity is particularly troubling in a diagnostic setting as these patients have a higher pre-test probability for cancer. In the diagnostic setting, if a trade-off has to be made, it is usually better to err on the side of improved sensitivity at the expense of specificity. As the technology matures and radiologists gain greater experience with digital mammography, the accuracy of digital mammography for diagnostic imaging is likely to improve. However, the studies to date do not support the equivalence of digital imaging to film when used for diagnostic mammography.

TA Criterion 4 is met for screening mammography.

TA Criterion 4 is not met for diagnostic mammography.

TA Criterion 5: The improvement must be attainable outside the investigational setting.

The large DMIST trial reported data from over 30 centers including community-based sites. The study trained over 150 radiologists to interpret digital images and most of the radiologists were not specialists in mammography. They clearly demonstrated the equivalence of digital mammography to film mammography with an overall trend towards better results for digital mammography. With appropriate training and quality control, results comparable to those seen in the DMIST trial should be attainable at all sites. Digital mammography also offers the opportunity to transmit the data electronically, without loss of image quality, to sites skilled at interpreting digital images.

From the studies published to date, it is not clear why studies of diagnostic mammography have not been able to demonstrate the benefits seen with screening mammography. Until better data are available, digital mammography should focus on screening mammography, particularly for women under that age of 50 with dense breasts.

TA Criterion 5 is met for screening mammography.

TA Criterion 5 is not met for diagnostic mammography.

Table 1: Studies comparing full-field digital mammography to screen-film mammography in a screening population

Study	Population	Mammogram	N	Detection rate	Sensitivity	Specificity	Recall rate	Biopsy rate	PPV	Comments
FFDM device										
Lewin 2002	U.S.: Women ≥ 40 years old, some with multiple mammograms during study period. Mean age 56.	Digital - GE	4489	0.40%	54% (27/50)	89%	12%	1.4%	3.4%	Two institutions; high recall rate, low specificity for both modalities. Eight incident cancers during one year of follow-up.
Senographe 2000D prototype		Film	4489	0.49%	66% (33/50)	85%	15%	2.1%	3.3%	
Skaane 2003	Norway: Screening mammograms women aged ≥ 50 years. Mean age 58.	Digital - GE	3683	0.62%	74% (23/31)	96%	4.6%	NR	12%	9/31 cancers detected were palpable at diagnosis.
Senographe 2000D		Film	3683	0.76%	90% (28/31)	97%	3.5%		20%	
Skaane 2004	Norway: Screening mammograms women aged 45-69 years. Subgroups by age.	Digital - GE	6,997	0.59%	NA	NA	3.8%	NR	15%	2:1 randomization to FM or DM stratified by age group. ITT violated. No follow-up for missed cancers; known missed cancers excluded from analysis – can't calculate sensitivity or specificity.
Senographe 2000D		Film	17,911	0.41%			2.7%		15%	
		<u>Age 45-49</u>	3,012	0.27%			3.7%		7.1%	
		Digital - GE	7,607	0.22%			3.0%		7.4%	
		Film	3,985	0.83%			3.8%		22%	
		<u>Age 50-69</u> Digital - GE	10,304	0.54%			2.5%		22%	
	Film									
Yamada 2004	Japan: Screening mammograms women aged 50-69 years. Mean age NR.	Digital - GE	480	0.42%	100% (2/2)	96%	4.2%	NR	10%	DM read on printed copy, not soft copy at workstation. No follow-up for missed cancers.
Senographe 2000D		Film	480	0.42%	100% (2/2)	97%	2.9%		14%	

Pisano 2005	U.S. and Canada: All women presenting for screening mammograms. Mean age 55 years.	Digital	42,760	0.41%	70% (177/254)	92%	8.4%	1.6%	5%	Largest and methodologically best study. Higher recall rate than European and Japanese studies.	
<u>5 Digital Systems:</u>		Film	42,760	0.39%	66% (167/254)	92%	8.4%	1.6%	5%		
Senographe 2000D		<u>Age 40-49</u>			NR	78%	90%				3%
Senoscan		Digital Film				51%	90%				2%
CRSM		<u>Age 50-69</u>				NR	NR				NR
DMS		Digital Film									
Selenia FFDM											

* Test characteristics are based, when possible, on a positive test being defined by BI-RADS 0, 4 or 5 and false negatives by interval cancers diagnosed within 365 days of the mammogram.

DM: Digital mammography
GE: General Electric

FM: Film mammography
DMS: Digital Mammography System

FFDM: Full field digital mammography
CRSM: Computed Radiography System for Mammography
NR: Not reported

ITT: Intention to treat

PPV: Positive predictive value

NA: Not applicable

Table 2: Studies comparing full-field digital mammography to screen-film mammography in a diagnostic testing population

Study	Population	N cancers	Mammogram	N	Detection rate	Sensitivity	Specificity	Comments
FFDM device								
Venta 2001 Senographe 2000D	Women undergoing diagnostic mammography	18	Digital - GE Film	692 692	1.9% 2.3%	72% (13/18) 89% (16/18)	NR	Kappa = 0.20. Underpowered: the difference in sensitivity was NS: p=0.20.
Becker 2001 SenoVision digital	Patients undergoing core biopsy for microcalcifications.	43	Digital – GE Film	111 121	96% 93%	NA	NA	Historical controls: digital unit replaced film unit. No significant differences, trend towards higher yield and fewer cores per lesion with digital unit.
Fischer 2002 Senographe 2000D	Women with microcalcifications on digital mammography.	20	Digital - GE Film	55 55	25% 25%	68% (13/19) 68% (13/19)	81% 76%	Hard copy digital. Better sensitivity (95% and 92%) but lower specificity (41% and 39%) if consider BI-RADS 3-5 as positive.
Cole 2004 Senoscan	Primarily women with abnormal film mammograms, cancer cases over sampled.	111	Digital - Fischer Film	247 247	NA	66% (73/111) 74% (82/111)	67% 60%	Hard copy digital. Average area under ROC curve for multiple readers lower for DM (0.715) than for FM (0.765), p NR.

* Test characteristics are based, when possible, on a positive test being defined by BI-RADS 4 or 5 and false negatives by interval cancers diagnosed within 365 days of the mammogram.

The detection rate in the microcalcification study refers to the percentage of lesions in which microcalcifications were identified in the core biopsy.

DM: Digital mammography
GE: General Electric
PPV: Positive predictive value

FM: Film mammography
DMS: Digital Mammography System
NA: Not applicable

FFDM: Full field digital mammography
CRSM: Computed Radiography System for Mammography
NR: Not reported

CONCLUSION

It is very difficult to interpret mammograms. As many of the studies comparing digital to film mammography demonstrated, up to a third of breast cancers that are present are missed by radiologists interpreting images obtained using either technology, even though as many as 20 women are called back for “positive” mammograms for every one cancer identified. Furthermore, the long-term benefits from mammography screening are relatively small and continue to be debated. The delicate balance of risks and benefits could be affected by small differences in the accuracy of mammography. Thus, it is important to carefully evaluate the test characteristics of any technology that may replace film mammography.

Five studies directly compared digital to film mammography in populations of women undergoing screening for breast cancer. The sensitivity (or detection rate) of digital mammography for breast cancer was lower than film in two of the studies, higher in two and equal in the small Japanese trial. The specificity of digital mammography was only better in the earliest trial. The DMIST trial included more women (n=42, 760) than all of the other trials combined and included five different technologies for digital mammography. In DMIST, digital mammography had equal recall rates and biopsy rates compared with film mammography and the specificities were identical. However, digital mammography was more sensitive than film, particularly for younger women with denser breasts. The AUC was significantly greater for digital compared with film mammography among women under the age of 50 years, women who were not postmenopausal and women with heterogeneously or extremely dense breasts. The study was performed at a large number of both academic and community-based centers with over 150 radiologists interpreting the images using methods similar to those widely used in the U.S. and Canada today. Thus, the results of DMIST should be obtainable outside the research setting.

Unfortunately, the same cannot be said for digital mammography used for diagnostic mammography. Sensitivity becomes even more important in the diagnostic setting as patients are being evaluated for signs or symptoms suggestive of breast cancer and a higher proportion will have cancer. None of the four relatively small studies demonstrated a greater sensitivity of digital mammography for cancer. In the study with the largest number of cancers (n=111 cancers), the average sensitivity of multiple readers interpreting the images was 66% for digital mammography and 74% for film mammography. This is a large and clinically important difference. The AUC was also lower for digital mammography (0.715 vs. 0.765). There is no fundamental reason why digital mammography should perform differently in diagnostic compared with screening populations. However, at this time, the evidence does not support the equivalence of digital to film mammography in a diagnostic setting.

RECOMMENDATION

- It is recommended that the use of digital mammography meets Technology Assessment Criteria 1 through 5 for safety, effectiveness and improvement in health outcomes when used for screening mammography.
- It is recommended that the use of digital mammography does not meet Technology Assessment Criteria 3 through 5 for safety, effectiveness and improvement in health outcomes when used for diagnostic mammography.

The California Technology Assessment Forum panel voted to approve the recommendation.

February 15, 2006

RECOMMENDATIONS OF OTHERS

Blue Cross Blue Shield Association (BCBSA)

The BCBSA Technology Evaluation Center Medical Advisory Panel reviewed this technology at their October 2005 meeting and determined that TEC criteria were met for screening and diagnosis of breast cancer.

Centers for Medicare and Medicaid Services (CMS)

CMS does not distinguish between film and digital mammography in their National Coverage policy.

California Radiological Society (CRS)

A CRS representative attended the meeting and provided testimony in support of the use of this technology for screening and diagnosis.

American Cancer Society (ACS)

ACS has not taken a formal position on the use of digital mammography as more or less appropriate for any specific subgroup. A representative was not able to participate at the meeting. The most current ACS statement

American College of Obstetrics and Gynecology (ACOG)

ACOG was not able to participate at the meeting and did not provide an opinion regarding the use of this technology.

American Society of Breast Surgeons (ASBS)

ASBS did not participate at the meeting regarding this topic and did not provide a position/opinion statement specific to the use of this technology

ABBREVIATIONS USED IN THIS ASSESSMENT:

AUC: Area Under the ROC Curve
BI-RADS: Breast Imaging Reporting and Data System
DMIST: Digital Mammographic Imaging Screening Trial
ROC: Receiver Operating Characteristic

REFERENCES

1. American Cancer Society. *Breast cancer facts & figures 2005-2006*. Atlanta: American Cancer Society, Inc.; 2005.
2. Alexander FE. The Edinburgh Randomized Trial of Breast Cancer Screening. *J Natl Cancer Inst Monogr*. 1997(22):31-35.
3. Alexander FE, Anderson TJ, Brown HK, et al. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet*. Jun 5 1999;353(9168):1903-1908.
4. Andersson I, Aspegren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *Bmj*. Oct 15 1988;297(6654):943-948.
5. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. *J Natl Cancer Inst Monogr*. 1997(22):63-67.
6. Bjurstam N, Bjorneld L, Duffy SW, et al. The Gothenburg Breast Cancer Screening Trial: preliminary results on breast cancer mortality for women aged 39-49. *J Natl Cancer Inst Monogr*. 1997(22):53-55.
7. Bjurstam N, Bjorneld L, Duffy SW, et al. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. *Cancer*. Dec 1 1997;80(11):2091-2099.
8. Larsson LG, Andersson I, Bjurstam N, et al. Updated overview of the Swedish Randomized Trials on Breast Cancer Screening with Mammography: age group 40-49 at randomization. *J Natl Cancer Inst Monogr*. 1997(22):57-61.
9. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study: update on breast cancer mortality. *J Natl Cancer Inst Monogr*. 1997(22):37-41.
10. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. *J Natl Cancer Inst Monogr*. 1997(22):27-30.
11. Tabar L, Fagerberg G, Chen HH, Duffy SW, Gad A. Screening for breast cancer in women aged under 50: mode of detection, incidence, fatality, and histology. *J Med Screen*. 1995;2(2):94-98.

12. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years--update of mortality results in the Stockholm mammographic screening trial. *Breast Cancer Res Treat.* Sep 1997;45(3):263-270.
13. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *Cmaj.* Nov 15 1992;147(10):1477-1488.
14. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst.* Sep 20 2000;92(18):1490-1499.
15. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med.* Sep 3 2002;137(5 Part 1):305-312.
16. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet.* Mar 16 2002;359(9310):909-919.
17. Dean PB. Final comment. The articles by Gotzsche and Olsen are not official Cochrane reviews and lack scientific merit. *Lakartidningen.* Jun 21 2000;97(25):3106.
18. Duffy S, Tabar L, Smith RA. The mammographic screening trials: commentary on the recent work by Olsen and Gotzsche. *J Surg Oncol.* Dec 2002;81(4):159-162; discussion 162-156.
19. Duffy SW, Tabar L, Smith RA. The mammographic screening trials: commentary on the recent work by Olsen and Gotzsche. *CA Cancer J Clin.* Mar-Apr 2002;52(2):68-71.
20. Gotzsche PC. Debate on screening for breast cancer is not over. *Bmj.* Sep 22 2001;323(7314):693.
21. Gotzsche PC. Mammographic screening: no reliable supporting evidence? *Lancet.* Feb 23 2002;359(9307):706.
22. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet.* Jan 8 2000;355(9198):129-134.

23. Miller AB. Is mammography screening for breast cancer really not justifiable? *Recent Results Cancer Res.* 2003;163:115-128; discussion 264-116.
24. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet.* Oct 20 2001;358(9290):1340-1342.
25. ACOG practice bulletin. Breast cancer screening. Number 42, April 2003. *Int J Gynaecol Obstet.* Jun 2003;81(3):313-323.
26. Smith RA, Saslow D, Sawyer KA, et al. American Cancer Society guidelines for breast cancer screening: update 2003. *CA Cancer J Clin.* May-Jun 2003;53(3):141-169.
27. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med.* Sep 3 2002;137(5 Part 1):347-360.
28. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *Jama.* Mar 9 2005;293(10):1245-1256.
29. Fletcher SW, Elmore JG. Clinical practice. Mammographic screening for breast cancer. *N Engl J Med.* Apr 24 2003;348(17):1672-1680.
30. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med.* Oct 27 2005;353(17):1784-1792.
31. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med.* Feb 4 2003;138(3):168-175.
32. Porter PL, El-Bastawissi AY, Mandelson MT, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst.* Dec 1 1999;91(23):2020-2028.
33. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *Jama.* Jul 3 1996;276(1):33-38.

34. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst.* Jul 5 2000;92(13):1081-1087.
35. Rosenberg RD, Hunt WC, Williamson MR, et al. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. *Radiology.* Nov 1998;209(2):511-518.
36. Feig SA, Yaffe MJ. Digital mammography. *Radiographics.* Jul-Aug 1998;18(4):893-901.
37. Pisano ED, Yaffe MJ, Hemminger BM, et al. Current status of full-field digital mammography. *Acad Radiol.* Apr 2000;7(4):266-280.
38. Pisano ED, Zong S, Hemminger BM, et al. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *J Digit Imaging.* Nov 1998;11(4):193-200.
39. Pisano ED, Cole EB, Hemminger BM, et al. Image processing algorithms for digital mammography: a pictorial essay. *Radiographics.* Sep-Oct 2000;20(5):1479-1491.
40. Astley SM. Computer-aided detection for screening mammography. *Acad Radiol.* Oct 2004;11(10):1139-1143.
41. Lewin JM, D'Orsi CJ, Hendrick RE, et al. Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. *AJR Am J Roentgenol.* Sep 2002;179(3):671-677.
42. Lewin JM, Hendrick RE, D'Orsi CJ, et al. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. *Radiology.* Mar 2001;218(3):873-880.
43. Pisano ED, Gatsonis CA, Yaffe MJ, et al. American College of Radiology Imaging Network digital mammographic imaging screening trial: objectives and methodology. *Radiology.* Aug 2005;236(2):404-412.

44. Skaane P, Skjennald A. Screen-film mammography versus full-field digital mammography with soft-copy reading: randomized trial in a population-based screening program--the Oslo II Study. *Radiology*. Jul 2004;232(1):197-204.
45. Skaane P, Young K, Skjennald A. Population-based mammography screening: comparison of screen-film and full-field digital mammography with soft-copy reading--Oslo I study. *Radiology*. Dec 2003;229(3):877-884.
46. Yamada T, Saito M, Ishibashi T, et al. Comparison of screen-film and full-field digital mammography in Japanese population-based screening. *Radiat Med*. Nov-Dec 2004;22(6):408-412.
47. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med*. Oct 27 2005;353(17):1773-1783.
48. Becker L, Taves D, McCurdy L, Muscedere G, Karlik S, Ward S. Stereotactic core biopsy of breast microcalcifications: comparison of film versus digital mammography, both using an add-on unit. *AJR Am J Roentgenol*. Dec 2001;177(6):1451-1457.
49. Cole E, Pisano ED, Brown M, et al. Diagnostic accuracy of Fischer Senoscan Digital Mammography versus screen-film mammography in a diagnostic mammography population. *Acad Radiol*. Aug 2004;11(8):879-886.
50. Fischer U, Baum F, Obenauer S, et al. Comparative study in patients with microcalcifications: full-field digital mammography vs. screen-film mammography. *Eur Radiol*. Nov 2002;12(11):2679-2683.
51. Venta LA, Hendrick RE, Adler YT, et al. Rates and causes of disagreement in interpretation of full-field digital mammography and film-screen mammography in a diagnostic setting. *AJR Am J Roentgenol*. May 2001;176(5):1241-1248.
52. Brettell DS, Ward SC, Parkin GJ, Cowen AR, Sumsion HJ. A clinical comparison between conventional and digital mammography utilizing computed radiography. *Br J Radiol*. May 1994;67(797):464-468.

53. Cole EB, Pisano ED, Kistner EO, et al. Diagnostic accuracy of digital mammography in patients with dense breasts who underwent problem-solving mammography: effects of image processing and lesion type. *Radiology*. Jan 2003;226(1):153-160.
54. Cole EB, Pisano ED, Zeng D, et al. The effects of gray scale image processing on digital mammography interpretation performance. *Acad Radiol*. May 2005;12(5):585-595.
55. Jarlman O, Borg A, Braw M, Kehler M, Lyttkens K, Samuelsson L. Breast imaging: a comparison of digital luminescence radiographs displayed on TV-monitor and film-screen mammography. *Cancer Detect Prev*. 1994;18(5):375-381.
56. Jarlman O, Samuelsson L, Braw M. Digital luminescence mammography. Early clinical experience. *Acta Radiol*. Mar 1991;32(2):110-113.
57. Kimme-Smith C, Bassett LW, Gold RH, Gormley L. Digital mammography. A comparison of two digitization methods. *Invest Radiol*. Nov 1989;24(11):869-875.
58. Yamada T, Ishibashi T, Sato A, et al. Comparison of screen-film and full-field digital mammography: image contrast and lesion characterization. *Radiat Med*. Jul-Aug 2003;21(4):166-171.
59. Fischmann A, Siegmann KC, Wersebe A, Claussen CD, Muller-Schimpfle M. Comparison of full-field digital mammography and film-screen mammography: image quality and lesion detection. *Br J Radiol*. Apr 2005;78(928):312-315.