Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved

Improvement in Sensitivity of Screening Mammography with Computer-Aided Detection: A Multiinstitutional Trial

Rachel F. Brem[1]
Janet Baum[2]
Mary Lechner[3]
Stuart Kaplan[4]
Stuart Souders[5]
L. Gill Naul[6]
Jeff Hoffmeister[7]

OBJECTIVE. Our study evaluated radiologist detection of breast cancer using a computer-aided detection system.

MATERIALS AND METHODS. Three radiologists reviewed 377 screening mammograms interpreted as showing normal or benign findings 9–24 months before cancer diagnosis from 17 of the 18 participating centers. In 313 cases, study radiologists recommended additional mammographic evaluation. In 177 cases, the area warranting additional workup precisely correlated with the subsequently diagnosed cancer. These 177 missed cancers were evaluated with computer-aided detection. The proportion of radiologists identifying the missed cancers was used to determine radiologist sensitivity without computer-aided detection.

RESULTS. The study radiologists determined that 123 of the 377 missed cancer cases warranted workup. Therefore, 123 additional cancers cases could have been found. The calculated radiologist sensitivity without computer-aided detection was therefore 75.4% (377 / [377 + 123]). Similarly, using the performance of the system on the missed cancers, we estimated that 80 (65.0%) of these 123 missed cancer cases would have been identified with the use of computer-aided detection. Consequently, the estimated sensitivity of radiologists using computer-aided detection was 91.4% ([377 + 80] / [377 + 123])—resulting in a 21.2% ([91.4% / 75.4%] – 1) increase in radiologist sensitivity with computer-aided detection.

CONCLUSION. Use of the computer-aided detection system significantly improved the detection of breast cancer by increasing radiologist sensitivity by 21.2%. Therefore, for every 100,000 women with breast cancer identified without the use of computer-aided detection, an estimated additional 21,200 cancers would be found with the use of computer-aided detection.

Received December 17, 2002; accepted after revision March 4, 2003.

R. F. Brem is a paid consultant to CADx Systems. J. Hoffmeister is the medical director for CADx Systems and Qualia Computing and a stockholder in Qualia Computing.

Supported by CADx Medical Systems, QC, Canada.

[1]Department of Breast Imaging and Intervention, George Washington University Medical Center, 2150 Pennsylvania Ave. N.W., Washington, DC 20037. Address correspondence to R. F. Brem.

[2]Department of Radiology, Beth Israel Deaconess Medical Center, 330 Brookline Ave., CC0492, Boston, MA 02215.

[3]Jane Brattain Breast Center, Park-Nicollett Clinic, St. Louis Park, MN 55416.

[4]Strax Institute, 4300 N. University Dr., Lauderhill, FL 33332.

[5]Department of Radiology, Kaiser Permanente Medical Center, 9961 Sierra Ave., Fontana, CA 92335.

[6]Department of Radiology, Scott and White Clinic, 2401 S. 31st St., Temple, TX 76508.

[7]CADx Systems, Ste. 100, 2689 Commons Blvd., Beavercreek, OH 45431.

AJR 2003;181:687–693

0361–803X/03/1813–687

© American Roentgen Ray Society

According to the American Cancer Society, invasive breast cancer was diagnosed in 192,000 American women in 2001, and an additional 47,100 women were diagnosed with in situ breast cancer. Breast cancer was the cause of death in 40,200 in 2001 [1], making breast cancer the second most common cause of cancer death in women. The lifetime risk of a woman developing this disease has been established as one in eight [2].

Although a recent analysis of eight randomized breast cancer screening trials questions the reduction in mortality from screening with mammography [3] and the Physician Data Query Panel sponsored by the National Cancer Institute concurs with this uncertainty [4], other United States and international organizations still affirm that screening with mammography, clinical breast examination, or both is effective in reducing breast cancer mortality. The United States Preventive Services Task Force, sponsored by the United States Department of Health and Human Services, commissioned an evaluation of eight mammography screening trials and concluded that mammography reduces breast cancer mortality by 16% [5]. The International Agency for Research on Cancer, which is part of the World Health Organization, concluded that these randomized trials showed that mammography reduces breast cancer mortality by 25–35% in women who are 50–69 years old [6].

The sensitivity of mammography ranges from approximately 70% to 90% [7–14]. Thus, for a woman with breast cancer, the probability that her cancer will be detected on mammography is 70–90% and the probability that it will not be detected is 10–30%. Therefore, even though mammography is an effective tool to use to detect breast cancer, thereby reducing mortality, further improvement in mammographic sensitivity is needed.

Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved

Studies have shown that the accuracy of mammographic interpretation increases when a mammogram is evaluated by two radiologists; mammographic interpretation by two radiologists improves breast cancer detection by 5–15% [15–26]. However, this approach is not currently advocated as a standard of care and requires substantial additional resources that are often not available [26]. A cancer can be missed because it is not mammographically visible, because it is not detected by the interpreting radiologist (i.e., an oversight), or because the radiologist misinterprets the finding. Clinical studies have shown that 30–70% of breast cancers diagnosed at screening mammography are visible in retrospect on prior examinations and that detection errors are responsible for approximately half of missed breast cancers, with interpretation errors accounting for the other half [10, 27–33].

In view of the frequency of missed breast cancers, computer-aided detection of breast lesions on mammograms has been evaluated as a method to improve mammographic sensitivity. In a study of the potential benefit of adding computer-aided detection to a single radiologist review [34], an analysis method was developed to estimate the increase in sensitivity a radiologist could expect with computer-aided detection. The results indicated that review by a single radiologist combined with computer-aided detection represents a possible alternative to mammographic interpretation by a single radiologist only or by two radiologists for reducing the rate of detection errors that lead to missed cancers. In fact, a prior study reported breast cancer detection is improved with the use of computer-aided detection—even when used by experienced mammographers [35].

The purpose of our study was to evaluate radiologist detection of breast cancer using a computer-aided detection system.

Materials and Methods

The mammograms used in this study were obtained from 18 participating clinical centers that consist of academic and community hospitals and mammography clinics; the specific locations are listed in the Acknowledgments. The protocol of the trial was approved by the institutional review board at each center. Informed consent was not required by any of the participating institutional review boards because this study is a retrospective study that used images from the archives of the institutions.

Study Design

Case characteristics.—This multicomponent study was conducted between April 1999 and May 2001. None of the participating centers enrolled more than 25% of the total study population for any case category. All cases of breast cancer were in women 29 years old or older (age range, 29–93 years). Patients with normal findings on mammograms were approximately 10 years younger (mean age, 52.3 years) than those who had either potentially or actually missed cancers (mean age, 62.3 and 64.4 years, respectively). These findings reflect the fact that the risk of developing breast cancer increases with patient age. All mammograms in this trial dated from less than 7 years before the start of the study so that only images of current quality were included. All mammograms consisted of craniocaudal and mediolateral oblique views for both the left and right breasts.

Missed cancers.—This study was designed to evaluate the efficacy of a computer-aided detection system in the detection of missed cancers. From a total of 930 screening-detected breast cancers accrued from 17 of the 18 sites, 377 potentially missed cancers were identified (Fig. 1). Potentially missed cancer cases were defined as mammographically detected breast cancer cases diagnosed 9–24 months after a mammogram with "normal" or "benign" findings (i.e., not requiring any additional workup). The interval of 9–24 months was chosen for efficiency of data collection because the most recent prior mammograms for women who follow typical breast cancer screening guidelines were generally expected to be within this interval. Mammograms known to have been reviewed by two radiologists were excluded.

Six radiologists were designated as the study radiologists. They reviewed the 377 potentially missed cancer cases to determine whether areas warranting additional workup were present on the prior mammograms in order to identify the (actually) missed cancer cases. Each case was reviewed by three of the six study radiologists. All study radiologists were board-certified mammographers, met the qualification standards of the Mammography Quality Standards Act [36], and had interpreted a minimum of 800 mammograms per year.

The study radiologists evaluated the potentially missed cancer cases without clinical information and without other images or additional prior mammograms. As stated in the protocol, in order to enhance blinding, the study radiologists were told that normal cases might be included as distracters among the cases they reviewed. In fact, no normal cases were included in the review of the prior screening mammograms. None of the study radiologists could review a case from his or her participating center. During the review of the potentially missed cancer cases, each study radiologist identified regions on the prior mammograms that warranted workup and characterized each region as

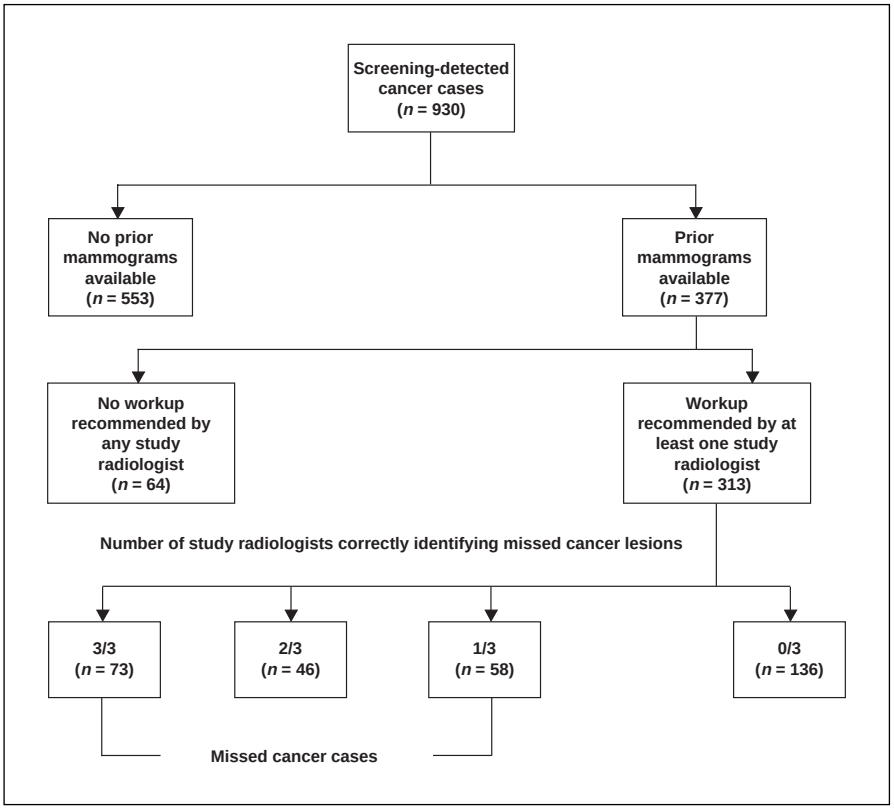

Fig. 1.—Flowchart of study cases shows cases of potentially missed breast cancer, workup recommended, and missed cancers.

Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved

mass, microcalcifications, or mixed (i.e., a mass lesion with microcalcifications), where "mass" lesions included architectural distortions and focal asymmetric densities. The study radiologists assigned an assessment recommendation on the basis of the Breast Imaging Reporting and Data System (BI-RADS) categories of the American College of Radiology [37] and precisely identified the location of the region that warranted workup using coordinates from a scaled grid overlay. The coordinates from the scaled grid overlay were used to confirm that the mammographic abnormality correlated precisely with the location of recommended workup. Coordinates of the overlying grid had to be within 1 grid or 5 mm to be considered in the same location.

Any case from the 377 potentially missed cancer cases for which additional workup was recommended by at least one of the three study radiologists was forwarded for a "truthing" review to assess whether the region for which workup had been recommended corresponded precisely to the location of the pathologically proven cancer, again with the use of a scaled grid overlay. The characterization of the lesion recommended for workup was also compared with the mammographic appearance of the cancer. Different radiologists constituted the truthing radiologists than the study radiologists. All radiologists participating in this study had interpreted more than 800 mammograms per year, as per the protocol, and all truthing radiologists were board-certified and met the qualification standards of the Mammography Quality Standards Act.

The prior screening mammogram and radiology report for each case were independently reviewed by two truthing radiologists with the use of supporting clinical and pathology data. Additionally, the mammogram that showed the cancer was also used to identify the precise location and mammographic appearance of the cancer. The truthing radiologists confirmed that the mammographic abnormality was in precisely the same location as the cancer by using a scaled grid overlay. Lesions visible in only one view were eligible to be included in this study. Any discrepancies in interpretation between the two truthing radiologists were resolved by a consensus review during which both radiologists reviewed the case together.

If one or more of the three study radiologists recommended workup with the appropriate lesion characterization in the location that was confirmed to correlate precisely with the location of the cancer, the case was classified as a missed cancer and the prior screening mammogram was processed by the computer-aided detection system (Second Look [version 3.4], CADx Systems, Beavercreek, OH). The precise location of the region recommended for workup was correlated with the location of the cancer by comparing the grid coordinates of the region the study radiologist recommended for workup with the grid coordinates of the location of the cancer as determined by the truthing radiologists. For the study radiologists to correctly identify the missed cancer lesion, they also had to characterize the lesion appropriately, where a mixed characterization was also considered correct for lesions determined to be only a mass or only microcalcifications by the truthing radiologists.

Of the 377 potentially missed cancer cases, there were 177 cases in which the region recommended for workup by at least one of the three study radiologists was confirmed to correlate to the precise location of the cancer and the lesion was characterized correctly. These 177 cases are the missed cancer cases that were then evaluated by the computer-aided detection system. The areas of mammographic abnormalities marked by the computer-aided detection system were then assessed to determine whether the location and characterization of the computer-aided detection mark corresponded precisely to the mammographically detected cancer. This assessment was performed by two scoring radiologists with scaled grid overlays used on both the computer-aided detection printout and the prior mammogram. Any discrepancies in interpretation between the scoring radiologists were resolved by consensus.

Each case was determined to be a true-positive or a false-negative by comparing the grid coordinates of the marks from the computer-aided detection system with the grid coordinates of the confirmed mammographic location of the cancer (the "truth lesion"). If a case had a mass or microcalcification mark at the location of a retrospectively visible lesion of the same characterization (or mixed) on at least one mammogram, it was a true-positive. Otherwise, it was a false-negative.

Normal cases.—One hundred fifty-five mammograms with normal findings were obtained from eight of the 18 participating centers. Although all participating centers could accrue cases with normal findings, only eight of the 18 sites chose to do so. One of the centers submitted only cases with normal findings. A normal case was defined as a screening mammogram that was interpreted as showing normal findings, not requiring any further workup (BI-RADS category 1), and for which there were at least 3 subsequent years of screening mammography with normal findings (BI-RADS category 1 or 2). Mammograms known to have been reviewed by two reviewers were excluded. Because the normal cases contained no cancer lesions, all computer-aided detection marks in these cases were considered false-positives. All normal cases were accrued sequentially.

Analysis of Mammograms by the Computer-Aided Detection System

The 177 missed cancer cases and 155 mammograms with normal findings used in this study were analyzed by the computer-aided detection system. The report generated by the computer-aided detection system consists of the digitized images with ellipses and rectangles highlighting potential areas of concern. The ellipses mark potential masses (nonspiculated masses, spiculated masses, architectural distortions, and asymmetric densities), and the rectangles mark potential microcalcifications. Figure 2A is a computer-aided detection printout that shows potentially suspicious areas highlighted, and Figure 2B is a computer-aided detection printout that a radiologist used while reviewing the mammogram.

Statistical Methods

The sensitivity of the original interpreting radiologists at the study centers was estimated as follows:

$$(N_{found} / [N_{found} + N_{missed}]).$$

The sensitivity of these radiologists using computer-aided detection (CAD) was estimated as follows:

$$([N_{found} + N_{CAD}] / [N_{found} + N_{missed}]),$$

where N_{found} is the number of potentially missed cancer cases with subsequent screening mammograms that led to the diagnosis of cancer obtained 9–24 months after the potentially missed cancer mammogram [34]. Constancy in patient demographics and in numbers of patients screened annually at the participating centers was assumed; therefore, this same number of cancer cases would have been obtained by the original reviewing radiologists during the prior screening interval.

N_{missed} is the estimated number of missed cancer cases that could have been found by the original reviewing radiologists at the study centers at the time of the missed cancer mammogram. N_{missed} was calculated using the proportion of study radiologists who correctly identified the truth lesion or lesions as a likelihood multiplier. The appropriate proportion of study radiologists (i.e., 0.3333, 0.6667, or 1.0) was multiplied by the number of missed cancers identified by one of three, two of three, and three of three study radiologists, respectively. The sum of these three quantities was calculated.

N_{CAD} is the estimated number of cancer cases that a radiologist using computer-aided detection would identify in addition to N_{found}. N_{CAD} was calculated by multiplying the proportion of study radiologists (i.e., 0.3333, 0.6667, or 1.0) by the number of cases correctly marked by the computer-aided detection system in each of these three categories and then calculating the sum of these three quantities.

A 95% confidence interval (CI) for the relative risk of the estimated sensitivities was calculated to assess the statistical significance of the relative increase in sensitivity with the use of computer-aided detection as compared with the sensitivity without computer-aided detection. This interval was then expressed as a relative change by subtracting 1 from both the lower and upper bounds and multiplying by 100.

Results

General Findings

The screening mammograms from each of the 377 potentially missed cancer cases were reviewed by three study radiologists. Of these 377 cases, 64 were not recommended for further

Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved

workup by any study radiologist. The remaining 313 cases were recommended for workup by at least one study radiologist. These cases underwent a truthing review to determine whether the mammographic area for which workup was recommended corresponded precisely to the mammographic location and characterization (i.e., mass, mixed, or microcalcifications) of the cancer. Of those 313 cases recommended for workup, 96 were recommended by one radiologist, 85 by two, and 132 by three of the three study radiologists.

Malignant lesions at the locations that had been recommended for workup by the study radiologists with the correct characterization were confirmed in 177 of the 313 cases reviewed by the truthing radiologists. These 177 cases were the missed cancer cases that were processed by the computer-aided detection system and then forwarded to the scoring radiologists. Of these 177 cases, the lesions were correctly identified by one of the three reviewing radiologists in 58 cases, by two in 46 cases, and by all three radiologists in 73 cases.

In 71 of the 313 cases, the truthing radiologists found no malignant lesion. In 65 cases, the truthing radiologists found at least one malignant lesion but at a different location or with a different characterization from any region recommended for workup by the groups of three study radiologists. These 136 cases were not analyzed by the computer-aided detection system.

Tumor Stage

Information about tumor size from pathology and histology results was collected if available at the participating site. Although there were 377 potentially missed cancer cases included in this study, pathologic cancer size was available in 249. Similarly, of the 177 missed cancer cases, pathologic size was available in 114. Table 1 refers to those cancers in which pathologic cancer size was available.

Tumor staging data (TNM stage for tumor size, lymph nodes, and metastasis) were obtained from sites with ready access to these data. These results are summarized in Table 2. The mean tumor size was 1.20 cm, and 76.9% of the tumors were invasive or microinvasive cancers (Table 1). In 23 of the 177 missed cancer cases, primary tumor staging was unknown (all were invasive or microinvasive tumors with size unavailable or unknown) (Table 2). Among the 154 missed cancer cases with known primary tumor staging, 102 (66.2%) were stage T1 (tumors ≤ 2 cm in greatest diameter), and 41 (26.6%) were classified as Tis (carcinoma in situ). Other stages (T2, T3, and T4) accounted for 11 (7.1%). In 90 of the 177 missed cancer cases, lymph node status could not be assessed (NX) or was unknown. Among the 87 cases with known lymph node status, 75 (86.2%) had no regional lymph node metastasis (N0). None of the cases had metastasis to the ipsilateral internal mammary lymph node (N3) or known distant metastasis (M1) at the time of diagnosis.

Truthing Results

Missed cancer cases with mammographically visible malignant lesions were classified as primarily a mass in 116 (65.5%) of the 177 cases and as primarily microcalcifications in 61 (34.5%) of the 177 cases. In 170 (96.0%) of the 177 cases, there was one malignant lesion; the remaining seven cases (4.0%) consisted of two malignant lesions each. Lesion size ranged from 0.3 to 9.0 cm in largest diameter on mammograms, with an average diameter of 1.26 cm. Approximately equal numbers of cancers were seen in the left and right breasts. One patient had bilateral cancers.

Radiologist Sensitivity With and Without Computer-Aided Detection

There were 177 missed cancer cases in this study. The number of study radiologists who correctly identified the location of a lesion on the prior mammogram with the appropriate characterization and recommended workup was used to estimate the number of missed cancer cases that could have been found by the original reviewing radiologists at the study centers. This number of potentially detectable missed cancers was calculated by multiplying the corresponding proportion of study radiologists (i.e., 33%, 67%, or 100%) by the number of missed cancers identified by one of three, two of three, and three of three study radiologists, respectively. The sum of these three quantities was calculated (Table 3). The total result for N_{missed} was 123

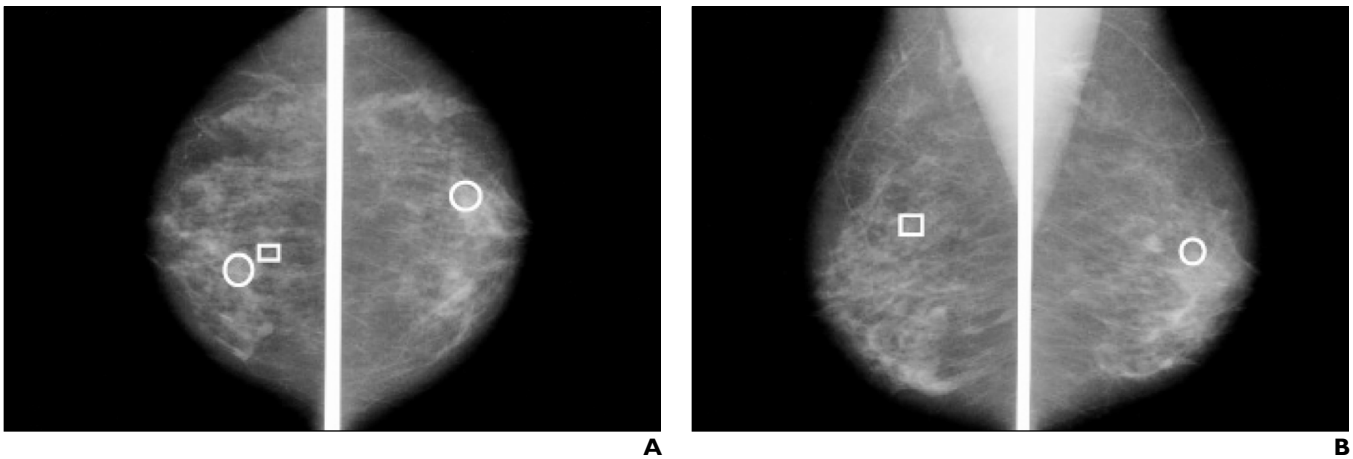

Fig. 2.—Printouts of digitized mammograms in 56-year-old woman who underwent screening mammography for breast cancer. Printouts were generated by computer-aided detection software (Second Look [version 3.4], CADx Systems, Beavercreek, OH).
A and **B,** Printouts of craniocaudal (**A**) and mediolateral oblique (**B**) views highlight areas with suspicious findings. Ellipses mark potential masses (nonspiculated masses, spiculated masses, architectural distortions, and asymmetric densities), and rectangles mark potential microcalcifications.

Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved

cases; this value represents the estimated number of mammograms with retrospectively visible cancers recommended for workup by the study radiologists conducting the blinded review that were missed by the original interpreting radiologists.

Because there were 377 (N_{found}) potentially missed cancer cases with subsequent screening mammograms that led to a cancer diagnosis, the estimated sensitivity of a radiologist was 75.4% (377 / [377 + 123]). This calculated number of mammograms warranting workup (n = 123) was compared with the number of mammograms with retrospectively visible cancers that were correctly identified by the majority (two or three of three) of the study radiologists. The truth lesion or lesions were correctly identified by two (67%) of three study radiologists in 46 cases and by three (100%) of three study radiologists in 73 cases (Table 3). Thus, the majority of the study radiologists correctly identified and recommended 119 (46 + 73) cases for workup, which compares well to the calculated 123 potentially detectable missed cancer cases.

To estimate the number of additional cancers that would have been detected with the use of computer-aided detection, we multiplied the proportion of study radiologists recommending a workup by the number of cases that the computer-aided detection system correctly marked (Table 3). The computer-aided detection system detected 63% of the missed cancer cases that had a subsequent screening mammogram that led to a cancer diagnosis (111/177). Similar to the calculation of the 123 potentially detectable missed cancers, the performance of the system was used to estimate that 80 (N_{CAD}) of these 123 missed cancer cases would have been identified by the original interpreting radiologists if they had used computer-aided detection. The estimated sensitivity of a radiologist using computer-aided detection was therefore 91.4% ([377 + 80] / [377 + 123]). This sensitivity of a radiologist using computer-aided detection is a 21.2% ([91.4% / 75.4%] – 1) increase (95% confidence limits = 14.5%, 28.3%) compared with the sensitivity of a single radiologist.

Normal Cases

In the 155 normal cases, the average number of mass and microcalcification marks per mammogram was 1.00 and 0.25, respectively.

Discussion

This study is one of the largest retrospective reviews of the screening mammograms of patients whose subsequent mammograms led to the diagnosis of breast cancer. Our results indicate that mammograms obtained 9–24 months before cancer diagnosis had evidence of an abnormality in 47% of the breast cancer cases (177/377) and that the earlier mammograms warranted additional workup as determined by the study radiologists in 32% of the breast cancer cases (123/377).

A prior study found that 36% of the most recent mammograms obtained 6–39 months (mean, 14 months) before the mammogram on which cancer was diagnosed were considered to have evidence of that cancer [30]. A recent study by Warren Burhenne et al. [34] indicated that 27% of prior mammograms (i.e., 115/427 mammograms obtained 9–24 months earlier) showed findings that warranted workup. Similarly, we found that mammograms obtained 9–24 months before the cancer diagnosis showed findings that warranted follow-up in 32% of the cancer cases in our study.

Researchers have inferred that detection or interpretation errors could be largely responsible for the incidence of missed breast cancers [10, 27, 33]. This may explain why breast cancer detection rates are higher with double interpretation [15–26]. Computer-aided detection is another way to reduce detection errors because it prompts the radiologist to reexamine the mammogram at the location of the computer-aided detection marks to determine whether further workup is necessary.

This study evaluated the benefit of using a computer-aided detection system by analyzing the mammograms obtained before a cancer diagnosis with computer-aided detection and estimating the increased sensitivity of breast cancer detection. The benefit derived from the computer-aided detection system was calculated as proportional to the number of radiologists who correctly identified the lesions at blinded review, so a more conservative estimate of computer-aided detection benefits could be obtained. For instance, if only one of three study radiologists identified a lesion, then only 33% of the cases correctly marked by computer-aided detection were considered in the calculation of computer-aided detection benefit (i.e., to determine that the original radiologist would have recommended workup for the cancer marked by computer-aided detection). It was only when all three study radiologists identified a lesion that 100% of the cases correctly marked by computer-aided detection were considered to calculate its benefit.

The estimated number of cases that would have benefited from the use of the computer-aided detection system (i.e., the cancer would have been detected on the prior mammogram if the computer-aided detection system had been used) was 80 (65%) of the 123 cancer cases found to warrant workup by study radiologists (i.e., cases that could have been detected by the radiologists who conducted the original mammography review). This value indicates that computer-aided detection could have helped reduce the rate of missed cancers by 65%.

To assess the impact of using computer-aided detection on radiologist sensitivity, we estimated that the sensitivity of a radiologist using computer-aided detection was 91.4%, which represented a 21.2% increase in radiologist sensitivity. Given that a 5–15% increase in breast cancer detection—as observed with double interpretation—is considered clinically significant [15–26], this 21.2% increase in radiologist sensitivity with the use of computer-aided detection is a clinically meaningful improvement in the detection of breast cancer and is similar to the improvement in sensitivity reported by Warren Burhenne et al. [34]. These estimates of sensitivity are based on a population of patients who had screening mammograms 9–24 months before cancer diagnosis. If it can be assumed that those cases without screening mammograms 9–24 months earlier are not appreciably different, similar results could be

TABLE 1 Pathology and Histology Findings for Potentially Missed and Missed Breast Cancers

Findings	Cancers	
	Potentially Missed	Missed
Pathology		
No. of tumors evaluated	249	114
Tumor diameter (cm)		
Mean	1.24	1.20
SD	0.85	0.64
Median	1.0	1.0
Minimum	0.1	0.2
Maximum	7.0	3.4
Histology		
No. of tumors evaluated	377	177
No. (%) invasive or microinvasive	290 (76.9)	136 (76.8)
No. (%) noninvasive	87 (23.1)	41 (23.2)

Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved

TABLE 2 Analysis of Tumor Staging Using TNM Classification System

TNM Stage	Definition of TNM Stage	No. (%) of Cancers	
		Potentially Missed (*n* = 377)	Missed (*n* = 177)
Primary tumor (T)			
Tis	Noninvasive in situ carcinoma	87 (23.1)	41 (23.2)
T1	Tumor ≤ 2 cm	213 (56.5)	102 (57.6)
T2	Tumor > 2–5 cm	25 (6.6)	10 (5.6)
T3	Tumor > 5 cm	1 (0.3)	0
T4	Tumor of any size with direct extension to chest wall or skin	1 (0.3)	1 (0.6)
Unknown		50 (13.3)	23 (13.0)
Regional lymph nodes (N)			
NX	Regional lymph nodes cannot be assessed	30 (8.0)	16 (9.0)
N0	No regional lymph node metastases	154 (40.8)	75 (42.4)
N1	Metastasis to movable ipsilateral axillary lymph nodes	29 (7.7)	11 (6.2)
N2	Metastasis to ipsilateral axillary lymph nodes that are fixed	5 (1.3)	1 (0.6)
N3	Metastasis to ipsilateral internal mammary lymph nodes	0	0
Unknown		159 (42.2)	74 (41.8)
Distant metastasis (M)			
MX	Presence of distant metastasis cannot be assessed	69 (18.3)	39 (22.0)
M0	No distant metastases	36 (9.5)	12 (6.8)
M1	Distant metastases	0	0
Unknown		272 (72.1)	126 (71.2)

expected for the entire population. Thus, for every 100,000 women with breast cancer currently identified by a radiologist without using computer-aided detection, an estimated additional 21,200 cancers would be found with the use of computer-aided detection.

In conclusion, this study has shown that a computer-aided detection system can significantly improve radiologist sensitivity in detecting breast cancer during screening mammography. The computer-aided detection system could also be helpful in screening mammography for the earlier detection of breast cancer—a particularly important benefit because women with breast cancer that is detected earlier have a better prognosis. The results of this multiinstitutional trial indicate that computer-aided detection may be viewed as a potential tool to help reduce the number of missed breast cancers, assist in the earlier detection of breast cancers, and serve as an alternative to double interpretation.

TABLE 3 Assessment of Mammograms Obtained 9–24 Months Before Cancer Diagnosis

No. of Study Radiologists Who Correctly Identified Lesion or Lesions	No. (%) of Cases	No. of Missed Cancer Cases That Radiologist Might Have Found	No. of Lesions Correctly Marked with Computer-Aided Detection System[a]	Estimated No. of Cases That Would Have Benefited from Analysis with Computer-Aided System[b]
One of three	58 (32.8)	19.3	33 (56.9)	11.0
Two of three	46 (26.0)	30.7	26 (56.5)	17.3
Three of three	73 (41.2)	73.0	52 (71.2)	52.0
Total	177 (100)	123	111 (62.7)	80.3

Note.—Estimated sensitivity of radiologists without computer-aided detection: [377 / (377 +123)] = 75.4%. Estimated sensitivity of radiologists with computer-aided detection: [(377 +80) / (377 + 123)] = 91.4%. Estimated relative increase of radiologist sensitivity with computer-aided detection: 21.2%. 95% confidence limits = 14.5%, 28.3%.

[a]Calculated by multiplying study radiologist score by respective number of missed cancer cases.

[b]Calculated by multiplying study radiologist score by number of cases correctly marked by computer-aided detection.

Acknowledgments

We acknowledge all of the additional sites that contributed cases and all the additional investigators of the study: Nagi Khouri, Johns Hopkins University Hospital, Baltimore, MD; Lawrence Bassett, Iris Cantor Center for Breast Imaging, UCLA School of Medicine, Los Angeles, CA; Wendie Berg, Department of Radiology, University of Maryland, Baltimore, MD; Richard Bird, Presbyterian Breast Center, Charlotte, NC; Steven Edell, Delaware Spect Imaging Center, Newark, DE; G. Richard Locke, Department of Clinical Research, Decatur Memorial Hospital, Decatur, IL; Jim Lynde, Mammography Department, Virginia Baptist Hospital, Lynchberg, VA; Thomas Poulton, Radiology Department, Aultman Hospital, Canton, OH; Robert Rosenberg, Department of Radiology, University of New Mexico, Albuquerque, NM; Barbara Monsees, Washington University, Mallinckrodt Institute of Radiology, St. Louis, MO; Darrell Smith, Department of Radiology, Brigham and Women's Hospital, Boston, MA; and Don Young, Department of Radiology, University of Iowa, Breast Imaging and Diagnostic Center, Iowa City, IA.

We thank Nadege Chery for assistance in the preparation of the manuscript. We thank Glenn Pixton for assistance with statistical analysis.

References

1. Thomas AG, Jemal A, Thun MJ. *Breast cancer facts & figures 2001–2002*. Atlanta: American Cancer Society, 2001:3
2. Feuer EJ, Wun LM, Boring CC, et al. The lifetime risk of developing breast cancer. *J Natl Cancer Inst* 1993;85:892–897
3. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129–134
4. Breast Cancer (PDQ) Web Site. Available at: http://cancer.gov/cancerinfo/pdq/screening/breast/healthprofessional/. Accessed October 16, 2002
5. Humphrey LL, Helfand M, Chan BK, et al. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137:347–360

6. Vainio H, Bianchini F. *Breast cancer screening*, vol. 7. Lyon, France: IARC Press, 2002:248
7. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of the factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002;225:165–175
8. Yankaskas BC, Cleveland RJ, Schell MJ, Kozar R. Association of recall rates with sensitivity and positive predictive values of screening mammography. *AJR* 2001;177:543–549
9. Kerlikowske K, Carney PA, Geller B, et al. Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Ann Intern Med* 2000;133:855–863
10. Bird RE, Wallace TW, Yankaskas BC. Analysis of cancers missed at screening mammography. *Radiology* 1992;184:613–617
11. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. *J Natl Cancer Inst* 1993;85:1644–1656
12. Goergen SK, Evans J, Cohen GP, MacMillan JH. Characteristics of breast carcinomas missed by screening radiologists. *Radiology* 1997;204:131–135
13. Robertson CL. A private breast imaging practice: medical audit of 25,788 screening and 1,077 diagnostic examinations. *Radiology* 1993;187:75–79
14. Sickles EA. Auditing your practice. In: Kopans DB, Mendelson EB, eds. *RSNA categorical course in breast imaging*. Oak Park, IL: Radiological Society of North America, 1995:81–91
15. Anderson ED, Muir BB, Walsh JS, Kirkpatrick AE. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994;49:248–251
16. Anttinen I, Pamilo M, Soiva M, Roiha M. Double reading of mammography screening films: one radiologist or two? *Clin Radiol* 1993;48:414–421
17. Beam C, Sullivan D, Layde P. Effect of human variability on independent double reading in screening mammography. *Acad Radiol* 1996;3:891–897
18. Ciatto S, Del Turco MR, Morrone D, et al. Independent double reading of screening mammograms. *J Med Screen* 1995;2:99–101
19. Warren RM, Duffy SW. Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. *Br J Radiol* 1995;68:958–962
20. Brown J, Bryan S, Warren R. Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms. *BMJ* 1996;312:809–812
21. Bird RE. Professional quality assurance for mammographic screening programs. *Radiology* 1990;177:587
22. Kopans DB. *Breast imaging*, 2nd ed. Philadelphia: Lippincott-Raven, 1998:214
23. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187–210
24. Thurjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. *Radiology* 1994;191:241–244
25. Hendee WR, Beam C, Hendrick E. Proposition: all mammograms should be double-read. *Med Phys* 1999;26:115–118
26. Kopans DB. Double reading. *Radiol Clin North Am* 2000;38:719–724
27. Birdwell RL, Ikeda DM, O'Shaughnessy KF, Sickles EA. Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology* 2001;219:192–202
28. Ikeda DM, Andersson I, Wattsgard C, Janzon L, Linell F. Interval carcinomas in the Malmo Mammographic Screening Trial: radiographic appearance and prognostic considerations. *AJR* 1992;159:287–294
29. Baines CJ, McFarlane DV, Miller AB. The role of the reference radiologist: estimates of inter-observer agreement and potential delay in cancer detection in the National Breast Screening Study. *Invest Radiol* 1990;25:971–976
30. Harvey JA, Fajardo LL, Innis CA. Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation. *AJR* 1993;161:1167–1172
31. Kopans DB. *Breast imaging*, 2nd ed. Philadelphia: Lippincott-Raven, 1998:793
32. Martin JE, Moskowitz M, Milbrath JR. Breast cancer missed by mammography. *AJR* 1979;132:737–739
33. van Dijck JA, Verbeek AL, Hendriks JH, Holland R. The current detectability of breast cancer in a mammographic screening program: a review of the previous mammograms of interval and screen-detected cancers. *Cancer* 1993;72:1933–1938
34. Warren Burhenne LJ, Wood SA, D'Orsi CJ, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology* 2000;215:554–562
35. Brem RF, Schoonjans JM. Radiologist detection of microcalcifications with and without computer-aided detection: a comparative study. *Clin Radiol* 2001;56:150–154
36. Mammography Quality Standards Act of 1992. 42 USC §263b
37. American College of Radiology. *Breast imaging reporting and data system (BI-RADS)*, 3rd ed. Reston, VA: American College of Radiology, 1998

Downloaded from www.ajronline.org by Utrecht University Library on 01/10/20 from IP address 145.136.140.249. Copyright ARRS. For personal use only; all rights reserved