Screening Mammograms: Interpretation with Computer-aided Detection—Prospective Evaluation[1]

Marilyn J. Morton, DO
Dana H. Whaley, MD
Kathleen R. Brandt, MD
Kimberly K. Amrami, MD

Purpose: To prospectively determine the effect of a commercially available computer-aided detection (CAD) system on interpretations of screening mammograms.

Materials and Methods: Institutional review board approval was granted; informed consent and HIPAA compliance were waived. A total of 21 349 screening mammograms obtained in 18 096 women were interpreted first without and then with review of CAD images to determine the effect of CAD analysis on the screening breast cancer detection rate, recall rate, and positive predictive value (PPV) for biopsy. The percentage of total cancers detected by the radiologists independent of CAD and the percentage correctly marked by the CAD system were determined.

Results: On the basis of pre-CAD interpretations, 2101 patients were recalled for diagnostic evaluation, 256 biopsies were performed, and 105 breast cancers were diagnosed. The breast cancer detection rate per 1000 screening mammograms was 4.92 (105 of 21 349 mammograms), the recall rate was 9.84% (2101 of 21 349 mammograms), and the PPV for biopsy was 41.0% (105 of 256 biopsies). After CAD image review, 199 additional patients were recalled, 21 additional biopsies were performed, and eight additional cancers were detected. The effect was a 7.62% (eight of 105) increase in the number of breast cancers detected, an increase in the recall rate to 10.77% (2300 of 21 349 mammograms), and a slight decrease in the PPV to 40.8% (113 of 277 biopsies). Radiologists detected 92.9% (105 of 113 cancers) of the total cancers, and CAD correctly marked 76.1% (86 of 113 cancers).

Conclusion: The use of CAD improved the detection of breast cancer, with an acceptable increase in the recall rate and a minimal increase in the number of biopsies with benign results.

[1] From the Divisions of Breast Imaging and Intervention (M.J.M., D.H.W., K.R.B., K.K.A.) and Biostatistics (J.N.M.), Mayo Clinic, 200 First St SW, Rochester, MN 55905. Received December 14, 2004; revision requested February 4, 2005; revision received May 12; accepted June 13; final version accepted July 1. **Address correspondence to** M.J.M. (e-mail: *morton.marilyn@mayo.edu*).

Screening mammography is recognized as the single most important tool for the detection of early-stage, clinically occult breast cancer, and numerous studies have demonstrated that screening mammography reduces breast cancer mortality (1–5). Mammography is not a perfect screening tool, however, and some breast cancers are not detectable at screening mammography. The reported false-negative rate for screening mammography is 10%–30%, with higher rates reported in women with mammographically dense breasts (6–11).

Several studies have shown that nonpalpable breast cancers detected at screening are often visible in retrospect on prior mammograms. In studies in which a nonblinded retrospective review was performed (ie, the current mammogram that led to a diagnosis of breast cancer was compared with the prior mammogram), up to 77% of breast cancers were visible on the prior mammogram (6,10,12–14). By comparison, studies that involved blinded review of prior mammograms found that 25%–41% of screening-detected breast cancers were visible retrospectively (7,10,14). Although the specific reason that a given breast cancer is overlooked at mammographic interpretation is not always readily apparent, many undetected breast cancers are overlooked because of perceptual errors on the part of the interpreting radiologist (8,10,15,16).

To reduce perceptual errors, independent double reading of screening mammograms by two or more radiologists has been investigated. Double reading increased the number of screening-detected breast cancers by 5%–15% (17–20). Although commonly performed in European countries, this method of double interpretation is not widely practiced in the United States, owing to the increased resources required for implementation.

To date, most studies in which an evaluation of the potential usefulness of computer-aided detection (CAD) as an aid to radiologists in their interpretation of mammograms has been performed have been retrospective. Several investigators have evaluated the potential of CAD to reduce the false-negative interpretations by performing CAD analysis on prior mammograms of patients with screening-detected breast cancers. On the basis of the ability of the CAD system to mark retrospectively visible cancers, investigators in two similarly designed studies (14,21) estimated the potential benefit of CAD to be an increase in the breast cancer detection rate of 21%. Although these results indicate that CAD systems do have the potential to increase the mammographic detection of early-stage breast cancer, the researchers in these studies did not directly measure the effect of CAD as an interpretation aid in actual clinical practice. To our knowledge, the investigators in only one published study (22) have evaluated the effects of CAD on the interpretation of screening mammograms in a clinical practice setting. In this prospective study, CAD analysis resulted in a 19.5% increase in the breast cancer detection rate.

The purpose of our study, thus, was to determine prospectively the effect of a commercially available CAD system on interpretations of screening mammograms.

Advances in Knowledge

- Computer-aided detection increased breast cancer detection by 7.62% with an absolute increase in the recall rate of 0.93%.
- The absence of a computer-aided detection mark on any potential mammographic finding should not dissuade the radiologist from recalling a patient for additional imaging.

Materials and Methods

Patients and Study Period

Institutional review board approval was granted for review of screening mammograms in patients who were recalled for additional imaging evaluation; informed consent and Health Insurance Portability and Accountability Act compliance were waived. Patients who did not give written authorization for the use of their medical records for research purposes were excluded from the study. Between April 1, 2001, and October 1, 2002, 21 349 screening mammograms were obtained consecutively in 18 096 asymptomatic women who were 23–98 years old (mean age, 60 years).

Image Acquisition and CAD Analysis

Each screen-film mammogram was acquired with one of four identical screen-film mammographic systems (Lorad MIV; Hologic Systems, Bedford, Mass) by a certified mammographic technologist in accordance with the rules and regulations set forth by the Mammography Quality Standards Act of 1992 (23). The CAD analysis was performed with a commercially available CAD system (ImageChecker M1000, version 2.2; R2 Technology, Sunnyvale, Calif) composed of a film digitizer to convert the analog films to digital images, a processing computer for application of the CAD algorithms, and an image display system. CAD images were displayed on low-resolution video monitors built into a motorized film viewer (RADX 614A; S&S Technology, Houston, Tex). The CAD system algorithms use a triangle to mark clusters of bright spots (suggestive of clustered microcalcifications) and an asterisk to mark areas of central density and radiating lines (suggestive of a mass or architectural distortion). With use of a continuous-loading digitizer, a routine four-view mammogram

Published online before print
10.1148/radiol.2392042121

Radiology 2006; 239:375–383

Abbreviations:
CAD = computer-aided detection
PPV = positive predictive value

Author contributions:
Guarantor of integrity of entire study, M.J.M.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, M.J.M., D.H.W.; clinical studies, all authors; statistical analysis, M.J.M., D.H.W., K.R.B.; and manuscript editing, all authors

Authors stated no financial relationship to disclose.

was digitized and processed in approximately 4 minutes.

Interpretation of Mammograms

Each screening mammogram was independently interpreted by one of 12 board-certified radiologists (including M.J.M., D.H.W., K.R.B., and K.K.A.), each of whom was aware of his or her participation in the study. The range of experience in mammographic interpretation was 3–30 years, with a mean of 12 years. In each case, the radiologist made an initial interpretation and assessment without viewing the CAD images. The initial assessment was either negative or benign (Breast Imaging Reporting and Data System 1 or 2) (24) or incomplete (Breast Imaging Reporting and Data System 0) (24). An incomplete assessment generated a recall of the patient for diagnostic imaging or a request to obtain and compare prior mammograms from other facilities, or both. The CAD images were then viewed on video display monitors, and the regions of interest marked by the CAD system were reexamined on the hard-copy film images. An abnormal mammographic finding detected by the radiologist was considered to be similarly detected with CAD if the appropriate CAD mark was present on at least one of the two views of each breast. If the potential abnormality had features of both a mass and microcalcifications, then the presence of either CAD mark, asterisk, or triangle was considered correct. The absence of a CAD mark on a potential abnormality did not alter the radiologist's initial decision to recall the patient on the basis of the pre-CAD interpretation. A final assessment and a recommendation were made after the completion of the diagnostic evaluation.

Data Collection

For each mammographic finding that resulted in the recall of a patient for additional diagnostic imaging (Breast Imaging Reporting and Data System 0), the following data were recorded by the interpreting radiologist: (*a*) for the method of detection, findings detected at the pre-CAD interpretation and also marked by the CAD system were recorded, findings not detected at the pre-CAD interpretation but marked by the CAD system with subsequent recall of the patient were recorded, and findings detected at the pre-CAD interpretation but not marked by the CAD system were recorded; (*b*) the predominant mammographic finding (mass or density, architectural distortion, or clustered microcalcifications); (*c*) the location within the breast (quadrant or clock position); and (*d*) the distance of the finding from the nipple recorded in centimeters.

The final interpretation assessment and recommendation for each screening mammogram in a patient recalled for diagnostic imaging were retrieved from our mammography database by one of the authors (D.H.W.). All percutaneous and surgical breast biopsy results and breast cancer staging data are maintained within this database. The histopathologic diagnosis of all lesions on which a biopsy was performed and the maximum dimension and stage of all screening-detected breast cancers were recorded by two of the authors (M.J.M. and D.H.W.).

Statistical Analysis

The recorded data were collectively analyzed by three individuals (including M.J.M., and D.H.W.) to determine the effect of CAD on the screening breast cancer detection rate (the number of occult cancers detected per 1000 examinations), the screening recall rate (the number of patients in whom screening mammograms were obtained who were recalled for diagnostic imaging divided by the total number of screening examinations), and the positive predictive value (PPV) for biopsy (the number of cancers diagnosed divided by the total number of biopsies performed).

The number of breast cancers detected by the radiologists independent of CAD analysis and the number correctly marked by the CAD system were calculated as percentages of the total breast cancers detected. For each breast cancer diagnosed that was not initially perceived by the radiologist and for each breast cancer not marked by the CAD system but detected by the radiologist, the specific mammographic finding (mass or density, architectural distortion, or microcalcifications), tumor histologic findings, and maximum tumor dimension were determined. In the absence of a paired control group, 95% confidence intervals were calculated from the recorded data. All analyses were performed with statistical software (SAS, version 8, 2001; SAS Institute, Cary, NC).

Results

The pre-CAD mammographic interpretations resulted in 2101 patients being recalled, 256 breast biopsies, and 105 breast cancer diagnoses (Table 1). Use of the CAD system to supplement the radiologists' interpretations resulted in an additional 199 patients being recalled, 21 breast biopsies, and eight breast cancer diagnoses (Table 1).

Screening Recall Rate

From the pre-CAD interpretations, findings on 2101 mammograms necessitated the recall of patients for diagnostic breast imaging evaluation, resulting in a screening recall rate of 9.84% (2101 of 21 349 mammograms) (Table 2). After CAD image review by the radiologists, mammographic interpretations generated an additional recalling of 199 patients for diagnostic evaluation. The recall rate increased to 10.77% (2300 of 21 349 mammograms), an absolute increase of 0.93% and a relative increase of 9.47% (199 of 2101).

Breast Cancer Detection Rate and PPV

All diagnostic evaluations, recommended biopsies, and follow-up evaluations generated by the recalling of 2300 patients were completed (Table 2). As an outcome of the recall of 2101 patients generated by the pre-CAD mammographic interpretations, 256 breast biopsies were performed, and 105 clinically occult breast cancers were diagnosed. This resulted in a screening breast cancer detection rate per 1000 screening mammograms of 4.92 (105 of 21 349 mammograms) and a PPV for biopsy of 41.0% (105 of 256 biopsies).

After the CAD images were reviewed

and the computer-marked regions of interest were reexamined by the radiologist, the an additional 199 patients were recalled. In these patients who were recalled, an additional 21 breast biopsies were performed, and this number resulted in an 8.20% (21 of 256) relative increase in the total number of breast biopsies performed. From these additional biopsies, eight breast cancers were diagnosed that had not been detected at the pre-CAD interpretations. Detection of these eight additional breast cancers increased the breast cancer detection rate from 4.92 to 5.29 (113 of 21 349 mammograms), a relative increase of 7.62% (eight of 105) in the number of breast cancers detected.

The proportion of benign and malignant breast lesions from which tissue samples were removed at biopsy was only slightly affected by the use of CAD. The PPV for the CAD-generated breast biopsies was 38% (eight of 21 biopsies) compared with the PPV of 41.0% (105 of 256 biopsies) for biopsies generated by the radiologists on the basis of their pre-CAD interpretations. The combined total PPV for biopsy was 40.8% (113 of 277 biopsies), a minimal decrease from the pre-CAD PPV of 41.0%.

Mammographic Findings

Among the total 113 breast cancers diagnosed from the 21 349 screening mammographic interpretations, a visible mass or density was the most common finding (72 [63.7%] of 113), followed by clustered microcalcifications (31 [27.4%]) and architectural distortion (10 [8.85%]). The radiologists detected 105 (92.9%) of 113 malignancies: 70 (97%) of 72 malignant masses, nine (90%) of 10 malignant architectural distortions, and 26 (84%) of 31 malignant clustered microcalcifications. The CAD system correctly marked 86 (76.1%) of 113 malignancies: 47 (65%) of 72 malignant masses, eight (80%) of 10 malignant architectural distortions, and all 31 (100%) malignant clustered microcalcifications.

Cancers Not Marked by CAD System

Of the 105 breast cancers detected by the radiologists, 27 (25.7%) were not marked by the CAD system (Table 3). In 25 (93%) of the 27 unmarked cancers, a mass or density was the predominant mammographic finding (Fig 1). These lesions ranged in size from 0.5 to 2.2 cm in greatest diameter. Architectural distortion was the predominant mammographic finding in the remaining two cancers not marked by CAD. Both were invasive ductal carcinomas, not otherwise specified, measuring 0.7 and 1.7 cm in greatest diameter. In none of the cancers overlooked by the CAD system were microcalcifications the predominant mammographic finding.

Cancers Not Perceived by Radiologists

The use of CAD as a supplement to the radiologists' interpretations resulted in the detection of eight breast cancers not perceived by the radiologists at pre-CAD interpretations (Table 4). In five (62%) of eight breast cancers, clustered microcalcifications were the predominant mammographic finding (Fig 2). All five of these lesions were ductal carcinoma in situ; three were low grade and two were intermediate grade. In two (25%) of eight cancers, the predominant mammographic finding was a noncalcified mass. One mass was a 1.1-cm invasive lobular carcinoma, and the other was a 1.2-cm invasive ductal carcinoma, not otherwise specified. Architectural distortion was the predominant mammographic finding in one (12%) cancer, a 1.6-cm invasive lobular carcinoma.

Table 1

Effect of CAD on Numbers of Patients Recalled, Breast Biopsies, and Detected Breast Cancers from 21 349 Screening Mammographic Interpretations

Interpretation	No. of Patients Recalled	No. of Biopsies	No. of Detected Cancers
Radiologist alone	2101	256	105
CAD alone	199	21	8
Radiologist and CAD	2300	277	113

Note.—Relative increase after CAD analysis was 9.47%, 8.20%, and 7.62% for numbers of patients recalled, of biopsies, and of detected cancers, respectively.

Table 2

Effect of CAD on Recall Rate, PPV for Biopsy, and Number of Screening-detected Breast Cancers

Interpretation	Screening Recall Rate		PPV for Biopsy		Cancers Detected per 1000 Screening Mammograms	
	Percentage*	95% CI	Percentage*	95% CI	Percentage*	95% CI
Radiologist alone	9.84 (2101/21 349)	9.44, 10.25	41.02 (105/256)	34.93, 47.31	4.92 (105/21 349)	4.00, 6.00
Radiologist and CAD	10.77 (2300/21 349)	10.36, 11.20	40.79 (113/277)	34.95, 46.84	5.29 (113/21 349)	4.40, 6.40
Relative increase after CAD analysis	9.47 (199/2101)	8.25, 10.80	NA	NA	7.62 (8/105)	3.35, 14.46

Note.—CI = confidence interval, NA = not applicable.

* Numbers in parentheses were used to calculate the percentages.

Discussion

The first report of the use of computer analysis to assist radiologists in the detection of mammographic abnormalities was published in 1967 (25). Since then, extensive scientific research has led to marked improvements in CAD signal-processing algorithms, and there has been growing interest in the use of this technology to assist radiologists in the interpretion of screening and diagnostic mammograms. In 1998, the U.S. Food and Drug Administration approved the first CAD system for clinical use in screening mammography. Today, three commercially available CAD systems have been approved for clinical use by the Food and Drug Administration: ImageChecker, Second Look (CADx Medical Systems, Laval, Quebec, Canada), and MammoReader (Intelligent Systems Software, Clearwater, Fla).

In a retrospective study by Warren Burhenne et al (14), 1083 clinically occult breast cancers were detected from current screening mammograms. In 427 of these cases, the most recent prior mammogram was available for review. In 67% (286 of 427) of these prior mammograms, the breast cancers were found to be visible in retrospect. At independent blinded review, 27% (115 of 427) of the prior mammograms were interpreted as warranting a recall of the patient for diagnostic evaluation. The 1083 current mammograms and each of the 427 prior mammograms were then analyzed with the ImageChecker CAD system. The CAD system correctly marked 84% (906 of 1083) of the current cancers, including 99% (400 of 406) of malignant calcifications and 75% (506 of 677) of cancers manifesting as a mass, asymmetric density, or architectural distortion. The CAD system correctly marked 89 (77%) of 115 retrospectively visible breast cancers considered to warrant a recall of the patient. Warren Burhenne et al (14) determined that the original radiologists had a false-negative rate of 21% and estimated that the use of CAD analysis could have reduced this rate by 77%.

In a more recent retrospective study, Brem et al (21) evaluated the ability of Second Look to correctly mark breast cancers overlooked on prior mammograms. Screening mammograms from which 930 occult breast cancers were detected were identified, and a total of 377 prior mammograms were available for review. In 47% (177 of 377) of the prior mammograms, the breast cancers were considered to be visible in retrospect by a panel of radiologists. These 177 prior screening mammograms were subsequently analyzed with Second Look, and 63% (111 of 177) of the overlooked cancers were correctly marked by the CAD system. In these two retrospective studies, the investigators determined that, if the initial mammographic interpretations had been established with CAD assistance, there could have been a 20% (14) and 21% (21) increase in the breast cancer detection rate.

To our knowledge, the only published prospective study of the use of CAD to supplement screening mammographic interpretations in clinical practice was reported by Freer and Ulissey (22). These authors reported their results with ImageChecker to interpret 12 860 screening mammograms. The radiologists detected 41 cancers (detection rate, 3.2 cancers per 1000 screening mammograms) without CAD assistance. Eight additional cancers, seven of which were clustered microcalcifications, were detected only after the radiologists reviewed the CAD-marked regions of interest. Detection of these additional cancers resulted in an increase in the breast cancer detection rate of 19.5% (eight of 41). This finding is similar to the potential estimated increases in breast cancer detection of 20% and 21% reported in the retrospective studies by Warren Burhenne et al (14) and Brem et al (21).

Freer and Ulissey (22) found that the radiologists detected 84% (41 of 49) of the total number of breast cancers at the pre-CAD interpretations, and the CAD system correctly marked 82% (40 of 49 cancers). In their experience, the effect of CAD analysis on the recall of patients was an absolute increase in the screening recall rate of 1.2% (from 6.5% to 7.7%) and a relative increase of 18.46% (1.2 of 6.5). The PPV for biopsy before and after CAD was 38%. All eight additional cancers detected only after CAD image review were stage 0 and 1 lesions.

To our knowledge, our study is the largest prospective evaluation of the effect of CAD assistance on the interpretation of screening mammograms in clinical practice. From the 21 349 interpretations of screening mammograms, 105 breast cancers were detected (4.9 per 1000 screening mammograms) by the radiologists independent of CAD image review. CAD analysis resulted in the detection of eight additional breast cancers that the radiologists did not detect prospectively. Thus, a total of 113 clinically occult breast cancers (5.3 per 1000 screening mammograms) were detected by the radiologists with CAD assistance. These additional eight CAD-detected cancers represent a relative increase in the breast cancer detection rate of 7.62% (eight of 105). This result differs from the estimated increase in breast cancer detection of 20%–21% found in two retrospective studies

Table 3

Mammographic Findings, Histologic Types, and Tumor Size in 27 Cancers Not Marked by CAD System

Mammographic Finding	No. of Mammograms*	Histologic Types†	Size (cm)
Mass or density	25 (93)	20 IDC	0.5–1.9
		1 mixed IDC and ILC	2.2
		4 DCIS	0.5–0.9
Architectural distortion	2 (7)	2 IDC	0.7, 1.7
Microcalcifications	0	NA	NA

* Numbers in parentheses are percentages.

† DCIS = ductal carcinoma in situ, IDC = invasive ductal carcinoma, ILC = invasive lobular carcinoma, NA = not applicable.

(14,21), as well as from the 19.5% increase reported in the prospective study by Freer and Ulissey (22).

We found that the radiologists detected 92.9% (105 of 113 cancers) of the total breast cancers without CAD assistance and that the CAD system correctly marked 76.1% (86 of 113) of the total breast cancers. Our results differ somewhat from the sensitivity rates of 84% for the radiologist and 82% for the CAD system reported by Freer and Ulissey (22). A comparison of the various studies about the evaluation of the effect of CAD on screening mammographic interpretations can be found in Table 5. Variations in reported CAD sensitivity rates may be attributable to differences in the specific mammographic features of the screening-detected breast cancers in any given study (microcalcifications vs mass vs architectural distortion). CAD algorithms are known to have a higher sensitivity for marking

Figure 1

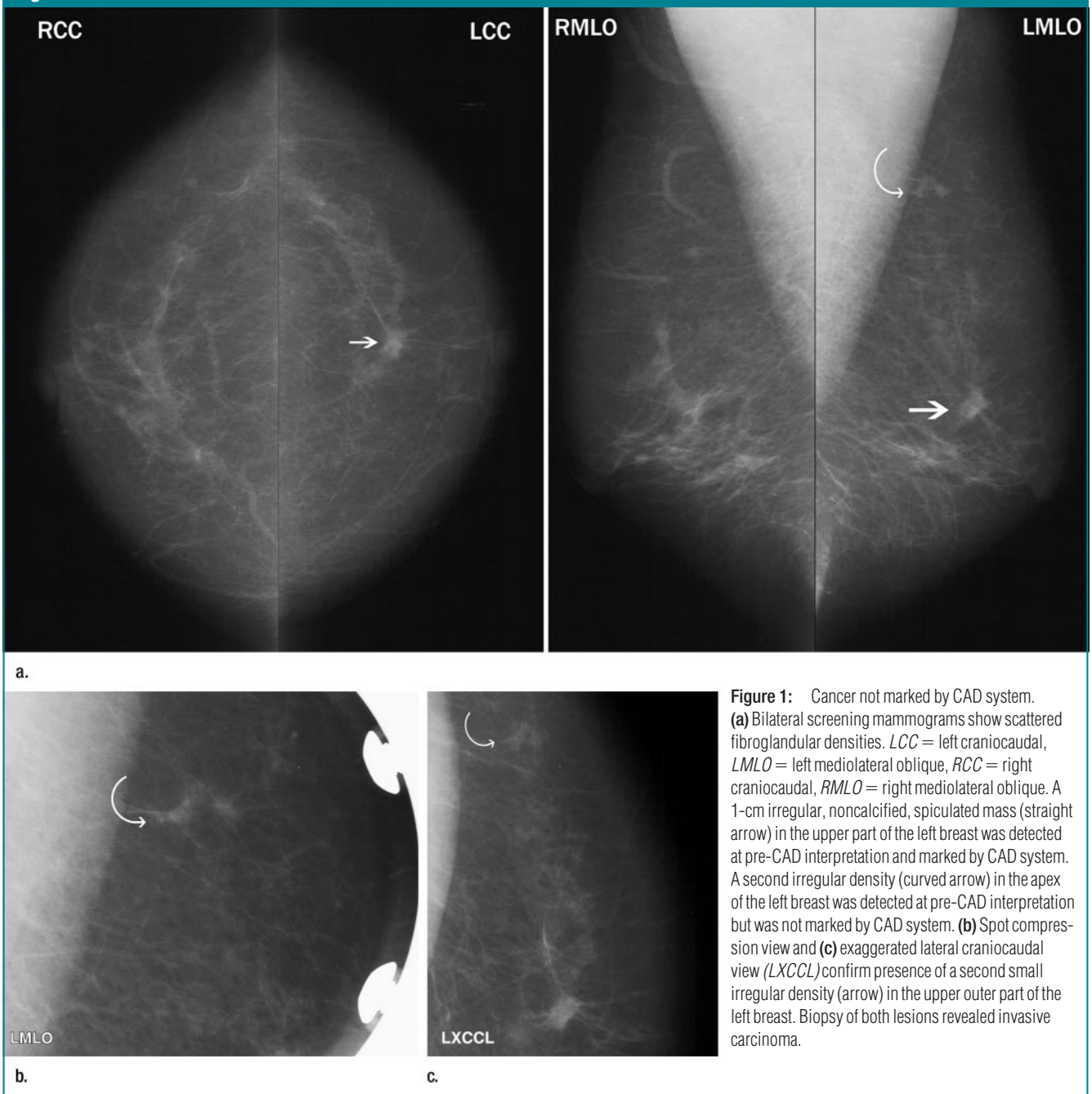

Figure 1: Cancer not marked by CAD system. **(a)** Bilateral screening mammograms show scattered fibroglandular densities. *LCC* = left craniocaudal, *LMLO* = left mediolateral oblique, *RCC* = right craniocaudal, *RMLO* = right mediolateral oblique. A 1-cm irregular, noncalcified, spiculated mass (straight arrow) in the upper part of the left breast was detected at pre-CAD interpretation and marked by CAD system. A second irregular density (curved arrow) in the apex of the left breast was detected at pre-CAD interpretation but was not marked by CAD system. **(b)** Spot compression view and **(c)** exaggerated lateral craniocaudal view *(LXCCL)* confirm presence of a second small irregular density (arrow) in the upper outer part of the left breast. Biopsy of both lesions revealed invasive carcinoma.

Figure 2

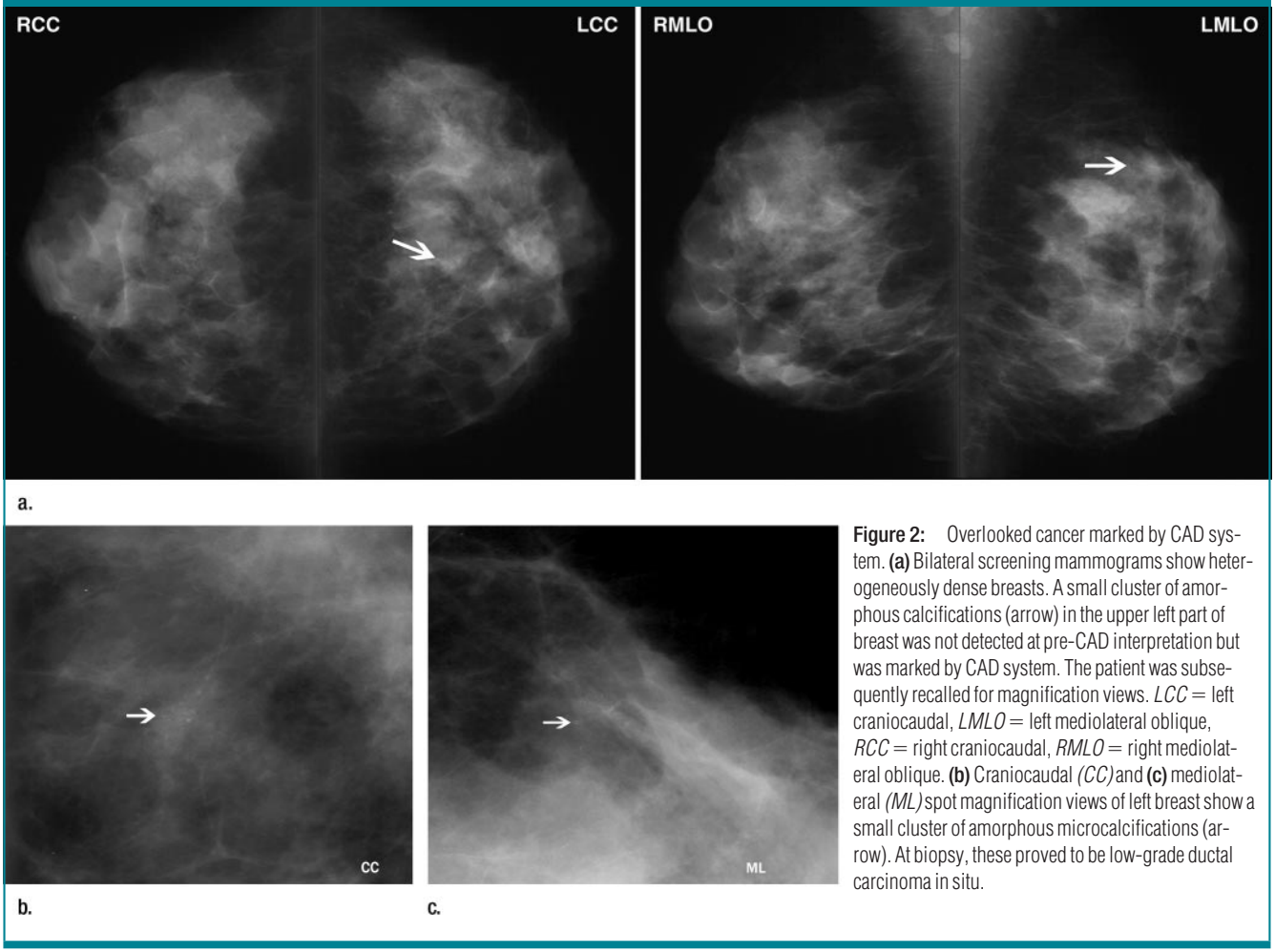

Figure 2: Overlooked cancer marked by CAD system. **(a)** Bilateral screening mammograms show heterogeneously dense breasts. A small cluster of amorphous calcifications (arrow) in the upper left part of breast was not detected at pre-CAD interpretation but was marked by CAD system. The patient was subsequently recalled for magnification views. *LCC* = left craniocaudal, *LMLO* = left mediolateral oblique, *RCC* = right craniocaudal, *RMLO* = right mediolateral oblique. **(b)** Craniocaudal *(CC)* and **(c)** mediolateral *(ML)* spot magnification views of left breast show a small cluster of amorphous microcalcifications (arrow). At biopsy, these proved to be low-grade ductal carcinoma in situ.

malignant calcifications than for marking malignant masses (14,15,22,26,27). In addition, CAD sensitivity values have been shown to vary according to the specifications of the CAD software used (28). Breast cancer detection rates and sensitivity values would also be expected to vary according to the expertise of the individual radiologist and to be influenced by differences in the prevalence and incidence rates of breast cancers in the populations screened. A higher prevalence of disease would be expected in our screening patient population (mean age, 60 years) compared with that (mean age, 49 years) in the study reported by Freer and Ulissey (22).

In most studies, the reported increases in breast cancer detection that occur with CAD-assisted mammographic interpretations are principally attributable to the high sensitivity of CAD algorithms in the marking of clustered microcalcifications. In our study, five (62%) of eight additional breast cancers detected with CAD were clustered malignant microcalcifications, whereas seven (88%) of eight breast cancers detected with CAD in the study by Freer and Ulissey (22) were malignant microcalcifications.

Table 4

Mammographic Findings, Histologic Types, and Size of Eight Breast Cancers Not Perceived by Radiologists at Pre-CAD Interpretations

Mammographic Finding	No. of Mammograms*	Histologic Types†	Size (cm)
Mass or density	2 (25)	1 ILC, 1 IDC	1.1, 1.2
Architectural distortion	1 (12)	1 ILC	1.6
Microcalcifications	5 (62)	5 DCIS	0.6–1.5

* Numbers in parentheses are percentages.

† DCIS = ductal carcinoma in situ (three were low grade and two were intermediate grade), IDC = invasive ductal carcinoma, ILC = invasive lobular carcinoma.

The results of studies in which the investigators evaluated the sensitivity of CAD algorithms for detection of malignant masses have been less impressive. The mass detection algorithm of current CAD systems encompasses a wide spectrum of mammographic abnormalities, ranging from dense spiculated masses and architectural distortion to ill-defined masses and focal asymmetric densities. In some reports, breast cancers in which the predominant mammographic finding was a mass, asymmetric density, or architectural distortion were all combined into a single "mass" category for the purpose of study analysis (14,22). In these studies, CAD sensitivity values for mass detection ranged from 67% to 75%. Using more advanced CAD algorithms, Castellino et al (28) and Petrick et al (29) achieved mass detection sensitivity values of 86% and 87%, respectively. By comparison, we found that the sensitivity of CAD was 65% (47 of 72) for malignant masses and 80% (eight of 10) for malignant architectural distortions. The 80% accuracy rate for CAD detection of architectural distortion is similar to the 83% and 85% CAD accuracy rates reported in two previous studies (15,27). In contrast, Baker et al (30) found that fewer than half of 45 malignant lesions characterized by architectural distortion were correctly marked by either one of two commercially available CAD systems.

The majority of CAD-generated marks are determined to be associated with benign findings. In particular, false-positive CAD marks are frequently seen with vascular calcifications and pseudolesions caused by overlapping fibroglandular tissue. For the CAD system used in this study, the average number of false-positive marks has been reported to be 0.5–0.7 per image (22,30). Although most experienced radiologists can readily recognize and easily ignore these false-positive marks, the low specificity of the CAD system has potential negative effects on interpretations of screening mammograms. Most notably, these include an increase in the number of patients recalled for additional imaging and an increase in the number of breast biopsies with benign results (a decrease in the PPV for biopsy). In our study, CAD image review resulted in an absolute increase in the recall rate of 0.93%, from 9.84% to 10.77%. Although the statistical significance of this change was not determined in our study, the increase was considered an acceptable trade-off for the increase in breast cancer detection that occurred. The desirable goal for the recall rate at screening mammography, as currently suggested by the American College of Radiology, is less than 10% (24). The use of CAD to supplement screening mammographic interpretations resulted in 21 additional patient biopsies. The ratio of benign to malignant histologic diagnoses, however, was very similar to that for the biopsies performed on the basis of pre-CAD interpretations, and the PPV was minimally affected (41.0% before CAD vs 40.8% after CAD).

Our study was limited by the potential for observational bias that may have been introduced because interpreting radiologists were aware of their participation in the study. Readers' vigilance for detection of microcalcifications may have been reduced because of their awareness that the reported sensitivity of CAD for detection of grouped microcalcifications is quite high. Reliance on the CAD system to mark calcifications may have introduced a bias against the results without CAD. Because we did not record CAD-marked regions of interest that were not considered actionable by the interpreting radiologist, we

Table 5

Summary of Studies about Evaluation of Effect of CAD on Breast Cancer Detection Rates, Recall Rates, and PPV for Biopsy

Study and Year	Study Design	No. of Mammograms	Cancers Detected per 1000 Screening Mammograms		Change (%)	Recall Rate (%)*		PPV for Biopsy (%)†	
			Without CAD	With CAD		Without CAD	With CAD	Without CAD	With CAD
Warren Burhenne et al (14), 2000	Retrospective‡	286§	NA	NA	20.2‖	NA	NA	NA	NA
		1083#	NA	NA		8.3	7.6	NA	NA
Freer and Ulissey (22), 2001	Prospective**	12 860	3.20	3.80	19.5	6.5 (830/12 960)	7.7 (986/12 860)	38.3 (41/107)	38.3 (49/128)
Brem et al (21), 2003	Retrospective‡	177	NA	NA	21.2‖	NA	NA	NA	NA
Present study	Prospective**	21 349	4.92	5.29	7.62	9.84 (2101/21 349)	10.77 (2300/21 349)	41.0 (105/256)	40.8 (113/277)

Note.—NA = not applicable.

* Numbers in parentheses are number of patients recalled for additional diagnostic imaging/total number of screening mammograms interpreted.

† Numbers in parentheses are the number of breast cancers found at biopsy/total number of patients who underwent biopsy.

‡ CAD analysis of prior screening mammograms with retrospectively visible overlooked breast cancers.

§ Retrospectively visible overlooked breast cancer cases.

‖ Estimated increase in sensitivity of breast cancer detection resulting from CAD analysis of overlooked breast cancers.

Currently diagnosed breast cancer cases.

** Comparison of screening mammographic interpretations with and without CAD.

could not determine whether the CAD system correctly marked findings of breast cancers that were dismissed by the radiologist as benign.

In conclusion, we found that the use of CAD analysis in the interpretation of 21 349 screening mammograms resulted in a 7.62% increase in the breast cancer detection rate. This was accomplished with an absolute increase in the recall rate of 0.93% (relative increase, 9.47%) and with essentially no effect on the PPV for biopsy (41.0% before CAD and 40.8% after CAD). The improvement in detection sensitivity was principally attributable to the presence of CAD marks on clusters of malignant calcifications that were not initially perceived by the radiologist. The CAD system marked 76.1% of the total breast cancers, and the radiologists detected 92.9%. The lower sensitivity of CAD in marking malignant masses and asymmetric densities underscores the fact that the absence of a CAD mark on any potential mammographic finding should not dissuade the radiologist from recalling a patient for additional imaging.

Acknowledgment: The authors acknowledge and thank Jayawant N. Mandrekar, PhD, for his contribution to the data review and statistical analysis.

References

1. Tabar L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age: new results from the Swedish Two-County Trial. Cancer 1995;75:2507–2517.
2. Tabar L, Vitak B, Chen HH, Yen MF, Duffy SW, Smith RA. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer 2001;91:1724–1731.
3. Duffy SW, Tabar L, Chen HH, et al. The effect of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer 2002;95:458–469.
4. 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 2002;137(5 pt 1):347–360.
5. Vainio H, Bianchini F. IARC handbooks of cancer prevention. Vol 7, Breast cancer screening. Lyon, France: International Agency for Research on Cancer, 2002; 128.
6. 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.
7. Yankaskas BC, Schell MJ, Bird RE, Desrochers DA. Reassessment of breast cancers missed during routine screening mammography: a community-based study. AJR Am J Roentgenol 2001; 177:535–541.
8. Bird RE, Wallace TW, Yankaskas BC. Analysis of cancers missed at screening mammography. Radiology 1992;184:613–617.
9. Goergen SK, Evans J, Cohen GP, MacMillan JH. Characteristics of breast carcinomas missed by screening radiologists. Radiology 1997;204:131–135.
10. Ganott MA, Harris KM, Klaman HM, Keeling TL. Analysis of false-negative cancer cases identified with a mammography audit. Breast J 1999;5:166–175.
11. Kerlikowske K, Grady D, Barclay J, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. J Natl Cancer Inst 1998;90:1801–1809.
12. Harvey JA, Fajardo LL, Innis CA. Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation—1993 ARRS President's Award. AJR Am J Roentgenol 1993;161:1167–1172.
13. Jones RD, McLean L, Young JR, Simpson W, Neilson F. Proportion of cancers detected at the first incident screen which were false negative at the prevalent screen. Breast 1996;5:339–343.
14. 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. [Published correction appears in Radiology 2000;216:306.]
15. 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.
16. Ikeda DM, Birdwell RL, O'Shaughnessy KF, Brenner RJ, Sickles EA. Analysis of 172 subtle findings on prior normal mammograms in women with breast cancer detected at follow-up screening. Radiology 2003;226:494–503.
17. Anderson ED, Muir BB, Walsh JS, Kirkpatrick AE. The efficacy of double reading mammograms in breast screening. Clin Radiol 1994;49:248–251.
18. Thurfjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. Radiology 1994;191:241–244.
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. Harvey SC, Geller B, Oppenheimer RG, Pinet M, Riddell L, Garra B. Increase in cancer detection and recall rates with independent double interpretation of screening mammography. AJR Am J Roentgenol 2003;180:1461–1467.
21. Brem RF, Baum J, Lechner M, et al. Improvement in sensitivity of screening mammography with computer-aided detection: a multiinstitutional trial. AJR Am J Roentgenol 2003;181:687–693.
22. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology 2001;220:781–786.
23. Mammography Quality Standards Act (MQSA) of 1992. Rockville, Md: U.S. Department of Health and Human Services, Food and Drug Administration, 1992.
24. American College of Radiology (ACR). ACR BI-RADS-Mammography. 4th ed. ACR Breast Imaging Reporting and Data System, Breast Imaging Atlas. Reston, Va: American College of Radiology, 2003.
25. Winsberg F, Elkin M, Macy J Jr, Bordaz V, Weymouth W. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. Radiology 1967;89:211–215.
26. Vyborny CJ, Doi T, O'Shaughnessy KF, Romsdahl HM, Schneider AC, Stein AA. Breast cancer: importance of spiculation in computer-aided detection. Radiology 2000;215:703–707.
27. Evans WP, Warren Burhenne LJ, Laurie L, O'Shaughnessy KF, Castellino RA. Invasive lobular carcinoma of the breast: mammographic characteristics and computer-aided detection. Radiology 2002;225:182–189.
28. Castellino RA, Roehrig J, Zhang W. Improved computer-aided detection (CAD) algorithms for screening mammography (abstr). Radiology 2000;217(P):400.
29. Petrick N, Sahiner B, Chan HP, Helvie MA, Paquerault S, Hadjiiski LM. Breast cancer detection: evaluation of a mass-detection algorithm for computer-aided diagnosis—experience in 263 patients. Radiology 2002;224:217–224.
30. Baker JA, Rosen EL, Lo JY, Gimenez EI, Walsh R, Soo MS. Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion. AJR Am J Roentgenol 2003;181:1083–1088.