



Published in final edited form as:

AJR Am J Roentgenol. 2012 September ; 199(3): W392–W401. doi:10.2214/AJR.11.7255.

Assessing the Standalone Sensitivity of Computer-aided Detection (CADe) with Cancer Cases from the Digital Mammographic Imaging Screening Trial (DMIST)

Elodia B. Cole, MS,

Medical University of South Carolina Department of Radiology and Radiological Sciences 96 Jonathan Lucas St., MSC 323 Charleston, SC 29425

Zheng Zhang, PhD,

Brown University Center for Statistical Science Box G-S121-7 121 S. Main St. Providence, RI 02912

Helga S. Marques, MS,

Brown University Center for Statistical Science Box G-S121-7 121 S. Main St. Providence, RI 02912

Robert M. Nishikawa, PhD,

Department of Radiology and the Committee on Medical Physics The University of Chicago 5841 South Maryland Ave., MC2026 Chicago, IL 60637

R. Edward Hendrick, PhD,

University of Colorado – Denver, School of Medicine Department of Radiology, C278 12700 E. 19th Ave. Aurora, CO 80045

Martin J. Yaffe, PhD,

Sunnybrook Health Sciences Centre 2075 Bayview Ave., S-Wing, Room S657 Toronto, ON M4N 3M5 CANADA

Wittaya Padungchaichote, MD,

Lopburi Cancer Center Lopburi 15000, Thailand

Cherie Kuzmiak, DO,

University of North Carolina Department of Radiology CB #5120 170 Manning Dr. Chapel Hill, NC 27599

Jatuporn Chayakulkheeree, MD,

Memorial Hospital Chulalongkorn University Bangkok 10330, Thailand

Emily F. Conant, MD,

University of Pennsylvania Department of Radiology/1 Silverstein 3400 Spruce St. Philadelphia, PA 19104

Laurie Fajardo, MD,

Department of Radiology Room 3970 JPP University of Iowa Hospitals and Clinics 200 Hawkins Dr. Iowa City, IA 52242

Janet Baum, MD,

Cambridge Health Alliance Dept of Radiology 1493 Cambridge St Cambridge, MA 02139

*Corresponding Author: Elodia B. Cole M.S., Department of Radiology and Radiological Science, Medical University of South Carolina, 96 Jonathan Lucas St., Suite 210, MSC 323, Charleston, SC 29425-3230. Phone: (843) 792-6965, Fax: (843) 792-1889, coleeb@musc.edu.

Constantine Gatsonis, PhD, and
Brown University Center for Statistical Science Box G-S121-7 121 S. Main St. Providence, RI
02912

Etta Pisano, MD
Medical University of South Carolina Department of Radiology and Radiological Sciences 96
Jonathan Lucas St., MSC 323 Charleston, SC 29425

Abstract

Objective—To assess the sensitivities and false detection rates of two CADe systems when applied to digital or screen-film mammograms in detecting the known breast cancer cases from the DMIST breast cancer screening population.

Materials and Methods—Available screen-film and digital mammograms of 161 breast cancer cases from DMIST were analyzed by two CADe systems, iCAD SecondLook (iCAD) and R2 ImageChecker (R2). Three experienced breast imaging radiologists reviewed the CADe marks generated for each available cancer case, recording the number and locations of CADe marks and whether each CADe mark location corresponded with the known location of the cancer.

Results—For the 161 cancer cases included in this study, the sensitivities of the DMIST reader without CAD were 0.43 (69/161, 95% CI 0.35 to 0.51) for digital and 0.41 (66/161, 95% CI 0.33 to 0.49) for film-screen mammography. The sensitivities of iCAD were 0.74 (119/161, 95% CI 0.66 to 0.81) for digital and 0.69 (111/161, 95% CI 0.61 to 0.76) for screen-film mammogram, both significantly higher than the DMIST study sensitivities ($p < 0.0001$ for both). The average number of false CADe marks per case of iCAD was 2.57 (SD 1.92) for digital and 3.06 (SD 1.72) for screen-film mammography. The sensitivity of R2 was 0.74 (119/161, 95% CI 0.66 to 0.81) for digital, and 0.60 (97/161, 95% CI 0.52 to 0.68) for screen-film mammography, both significantly higher than the DMIST study sensitivities ($p < 0.0001$ for both). The average number of false CADe marks per case of R2 was 2.07 (SD 1.57) for digital and 1.52 (SD 1.45) for screen-film mammogram.

Conclusions—Our results suggest the use of CADe in interpretation of digital and screen-film mammograms could lead to improvements in cancer detection.

INTRODUCTION

In the Digital Mammographic Imaging Screening Trial (DMIST), the reported sensitivities for both digital and screen-film mammography were 0.41 based on 455 days of follow-up.¹ While low, similar sensitivities for digital and screen-film mammography have been previously reported, ranging from 0.40 to 0.77 for digital mammography and from 0.31 to 0.71 for screen-film mammography in recently published results from six other controlled laboratory studies,^{1-12,13} while the National Breast Cancer Surveillance Consortium (NBCSC) reported an average screening sensitivity of 81.2% in their analysis of screening mammography records for over 3.8 million women between 1996 and 2007¹⁴ based on 365 days of follow-up.³ There is wide variation in reported radiologist sensitivities in these studies that is likely dependent not only on the skills of the participating radiologists, but also on the image case sets, the number of rounds of screening, and definitions of sensitivity¹⁵⁻¹⁸. Even so, there is a need for development of tools that improve the detection of breast cancer; thus the impetus for computer-aided detection software development. We use the abbreviation CADe to refer to computer-aided detection software, as opposed to the more ubiquitous CAD, which has been used to reference both computer-aided detection algorithms and computer-aided diagnostic algorithms. While the former is designed to prompt the radiologist to regions of a mammogram that may have been overlooked, the

latter, now frequently referred to as CADx in the literature, is designed to aid radiologists in distinguishing cancerous lesions from non-cancerous lesions.

It has been noted previously that the cases used to train and ultimately test a CADe system are the primary drivers on how well that technology will perform in cancer detection.¹⁹

While CADe publications have gone into great detail about the types and sizes of lesions included, the complexity of the dataset is hard to determine simply from these metrics. CADe systems were designed specifically to reduce the number of cancers missed by radiologists by marking suspicious regions of interest, prompting radiologists to inspect the annotated region. While the gold standard methodology for evaluation of a CADe system should involve radiologists reviewing clinical cases without and then with CADe, it is also useful to assess the cancer detection capability of the CADe system alone, without a radiologist reader. The study objectives detailed in this paper were: i) to estimate the sensitivity of each CADe system to known cancers when applied to digital and screen-film mammography; ii) to compare the standalone sensitivities of each CADe system to the original DMIST study sensitivities for digital and screen-film interpretation; iii) to report the average number of false positive marks per case for each CADe system; and, iv) to estimate the effect of tumor and patient characteristics on sensitivity.

MATERIALS AND METHODS

Appropriate institutional review board approval was obtained prior to conducting this study. Informed consent and HIPAA consent were obtained from all evaluable subjects prior to image acquisition for DMIST. All data were handled in a HIPAA compliant manner. Digital and screen-film mammograms of the available DMIST pathologically-proven cancer cases, i.e., those detected within 455 days of acquisition of the paired digital and screen-film mammograms were included in this study. The screen-film and digital mammograms were obtained from the American College of Radiology Imaging Network (ACRIN) image archive. A total of 329 case pairs, including both digital and screen-film images, were available for inclusion in this study out of a total of 335 DMIST cancer cases.

CADe Systems Tested

The CADe systems for screen-film and digital mammography of two manufacturers were tested in this study: iCAD Secondlook v1.4 (digital and screen-film), R2 ImageChecker Cenova v1.0 (digital), and R2 ImageChecker v 8.0 (screen-film). The results from the default CAD sensitivity settings for each system were reported as the main results. In addition to the default CAD sensitivity setting for R2 screen-film, two additional CAD sensitivity settings were applied to a single digitized film set, and results are included in the appendix. While any film mammogram can be processed using any film CADe system, not all digital mammograms are suitable for processing with all digital CADe systems. At the time of this study, the iCAD Secondlook v1.4 used in the study was compatible with images from the Fuji CR, Hologic Selenia, and GE Senographe 2000D systems. The R2 ImageChecker Cenova CADe system tested was capable of analyzing images from the Fischer SenoScan, Fuji CR, Hologic Selenia, and GE Senographe 2000D digital mammography systems. The maximum available number of cases for each of our two digital CADe systems was limited by these compatibility constraints. In total, 245 cases were available for the iCAD digital system and all 329 cases were available for the R2 digital system.

Film Preparation

De-identified screen-film mammograms were digitized using manufacturer specified digitizer. CADe reports were printed to paper. Twenty-seven of the 329 available screen-film mammograms (8.2%) were rejected by the iCAD screen-film digitizer and therefore could not be digitized, resulting in 302 screen-film cases analyzed by iCAD. Twenty-six cases of 329 (7.9%) available screen-film mammograms were rejected by the R2 film digitizer and therefore could not be digitized, resulting in 303 screen-film cases analyzed by R2.

Digital Image Preparation

Digital mammograms were processed with CADe and displayed on manufacturer specific mammography workstations. The application of CADe algorithms was not successful for all digital images. In total, there were 179 evaluable digital cases for iCAD and 227 digital cases for R2. The CADe-generated electronic reports were displayed on a Hologic Secureview 6.0 workstation for R2 (Figure 1) and a Sectra IDS5.MX review workstation for iCAD (Figure 2).

Common Cases by Both Vendors

There are 161 common cases that were fully evaluated by both vendors, on both film screen and digital mammograms. We focused our data analysis on those common cases, and presented the patient and lesion characteristics of those cases on Table 1.

CADe Scoring Procedure

A radiologist with 26 years of mammography experience annotated acetate overlays with the screen-film mammograms for each case (Figure 3), denoting the locations of known cancers as seen on DMIST images and/or validated by DMIST pathology reports. The annotating radiologist specified whether cancer was visible for each modality.

Three different radiologists reviewed the generated CADe marks and compared these to the known cancer locations depicted on the overlays. All three readers completed breast imaging fellowships and had an average of 4 years (range 2-8 years) experience in mammography. An average of 48.3% (range 25%-70%) of their time was spent interpreting screening mammograms in their respective clinical practices. One radiologist scored the iCAD digital cases, a second radiologist scored the R2 digital cases, and a third radiologist scored all digitized screen-film CADe for both R2 and iCAD. The radiologists recorded the number of CADe marks generated, the location of each CADe mark, and determine whether the location of each CADe mark corresponded to a known cancer as recorded on pathology reports and the case overlay. Marks that did not correspond to a known cancer location were recorded as false positives by the radiologists.

Statistical Methods

All statistics were performed at the case level to allow for comparison to DMIST results. The study objectives were: i) to estimate the sensitivity of each CADe system to known cancers when applied to digital and screen-film mammography; ii) to compare the standalone sensitivities of each CADe system to the original DMIST study sensitivities for digital and screen-film interpretation; iii) to report the average number of false positive marks per case for each CADe system; and, iv) to estimate the effect of tumor and patient characteristics on sensitivity. Sensitivity and the associated 95% exact confidence interval for each CADe system were calculated and McNemar's test was used to test the difference between digital and screen-film within each CAD system. McNemar's test and 95% exact confidence intervals were used for the second objective. Mean, standard deviation, median,

and number of false positive marks per case were calculated for the third objective and paired t-test were used to test the difference between digital and screen-film within each CAD system. To address the fourth objective, sensitivities for subsets defined by age, breast density, menopausal status, cancer histology, tumor size, and lesion type were calculated by simple counts as well as univariate logistic regression-based estimates. 95% confidence intervals and p-values for comparing covariate effects on sensitivity were drawn from the same logistic regression model.

RESULTS

We present the results for the default CADe sensitivity settings for screen-film and digital for iCAD and R2 for the 161 common cancer cases.

Standalone Performance

There were 161 common cases across both screen-film and digital modalities for each CAD vendor, allowing for modality comparison within each vendor. The overall sensitivity of iCAD digital was 119/161 (0.74) [95% CI: (0.66, 0.81)] and 111/161 (0.69) [95% CI: (0.61, 0.76)] for screen-film. This difference in sensitivity between iCAD digital and screen-film 0.05 [95% CI: (-0.03, 0.13)] was not statistically significant ($p=0.26$). The average number of false CADe marks per case for iCAD digital (2.57 ± 1.92) was significantly lower than that for iCAD screen-film (3.06 ± 1.72) ($p=0.0008$). The overall sensitivity of R2 digital was 119/161 (0.74) [95% CI: (0.66, 0.81)] and 97/161 (0.60) [95% CI: (0.52, 0.68)] for R2 screen-film. The difference in sensitivity between R2 digital and R2 screen-film 0.14 [95% CI: (0.05, 0.23)] was statistically significant ($p=0.003$). The average number of false CADe marks per case for R2 digital (2.07 ± 1.57) was significantly higher than for R2 screen-film (1.52 ± 1.45) ($p < 0.0001$) (Table 2).

CADe Detection of Cancers Not Detected in the DMIST Study

In standalone mode, each CADe system for both screen-film and digital mammography was able to detect significantly more cancer cases than were detected by the original clinical radiologist readers in DMIST for the same cases. Of the 161 common cases, the original DMIST study has sensitivity of 0.43 (69/161, 95% CI 0.35 to 0.51) for digital and 0.41 (66/161, 95% CI 0.33 to 0.49) for screen-film. For digital mammogram, both CADe systems' sensitivity for digital (0.74) was significantly higher than the DMIST's sensitivity ($p < 0.0001$). For screen-film, sensitivities for iCAD(0.69) and R2(0.60) were both significantly higher than the DMIST's sensitivity ($p < 0.0001$ for both) (Table 3).

In addition, CADe system has shown ability to detect the cases that were not detected by any method in the DMIST study. Of the 46/161 cases that were detected by neither screen-film nor digital readers in the primary DMIST study, 25 (54.3%) were marked by the iCAD system when applied to digital, and 18(39.1%) when applied to screen-film. The comparable numbers for R2 system is 27(58.7%) for digital and 12 (26.1%) for screen-film.

Lesion and Subject Characteristics Impact on Sensitivity

CADe's sensitivity on lesion's histology type (DCIS or invasive), size and type (mass, calcification, asymmetric density and arch distortion), and the subject's age, breast density, and menopausal status was performed (Table 4) and compared using univariate logistic regression models (Table 5). CADe sensitivity did not appear to depend on subject's characteristics, although the sensitivity of R2 screen-film was lower on pre-menopausal women (49%) than on post-menopausal women (64%), but the difference was not significant ($p=0.09$). Lesion histology type and size didn't influence the sensitivity significantly, but lesion type seemed to have some effect. In general, CADe had the highest

sensitivity for calcification (83% for both iCAD and R2 digital, 73% for both iCAD and R2 screen-film), but R2 screen-film has a low sensitivity of 55% for mass, significantly lower than the 73% for calcification on the same system ($p = 0.038$). In addition, asymmetric density has low sensitivity of 50% on both iCAD and R2 digital, which are significantly lower than the 83% for calcification ($p=0.017$).

DISCUSSION

In this study, we evaluated the standalone sensitivity of commercial CADE systems to breast cancers that were found at screening or follow-up in the DMIST study for both digital and screen-film mammography. We found that the CADE systems tested were able to detect significantly more cancers than were found on initial screening by the original radiologists in DMIST for both digital and screen-film mammography; for each manufacturer, standalone sensitivity was higher for digital than for screen-film; and the standalone sensitivities of CADE systems was not influenced by lesion characteristics (histology, lesion size, lesion type) or subject characteristics (age, breast density, menopausal status).

The overall average standalone CADE sensitivities reported here, ranging from 0.60 to 0.74, are low compared to other retrospective standalone CADE studies. Bolivar et al reported a digital R2 CADE sensitivity of 93% for a study that included cancers seen on both MLO and CC views.²⁰ The et al reported iCAD CADE sensitivity of 94% for cancers seen in at least one screening view.²¹ Yang et al reported a sensitivity of 96.1% for cases that included only a single cancer per subject detected by radiologists on digital mammograms.²² Our study differed in that we included mammographically occult lesions. Besides mammographic visibility of lesions, the characteristics of lesions and subjects have been shown to impact radiologist sensitivity. In our study, there was no statistical difference in CADE sensitivity based on breast density, lesion type, cancer histology, subject age, tumor size, and menopausal status. The et al. reported no statistical difference in CADE sensitivity for their study based on histopathology or tumor size.²¹ Bolivar et al reported no difference in digital CADE sensitivity based on breast density, lesion type, or histopathology, but they did find a significant difference based on tumor size for masses.²⁰

One unique benefit of using the DMIST dataset for CADE evaluation was the ability to assess the performance of CADE algorithms for both digital and screen-film mammography. In our study, we used a single CADE system for both digital and screen-film with iCAD, finding no difference in standalone CADE sensitivities between iCAD digital and screen-film ($p=0.26$). For R2 digital and screen-film cases, two different versions of R2 CADE algorithms were used. There was a significantly higher standalone CADE sensitivity for R2 digital than for R2 screen-film ($p=0.003$).

There were a few of limitations to this study. First, the methodology for scoring CADE mark localization of known cancer locations could have introduced bias. Three different radiologists conducted the scoring for screen-film (both iCAD and R2), iCAD digital, and R2 digital. While we attempted to minimize this potential bias by providing the readers with annotated overlays that directly matched the films generated by a fourth radiologist, the radiologists that scored the digital CAD did have to use their judgment as to whether the CADE mark shown on the computer screen overlapped the annotation on the overlay. Second, because our study included only cancer cases, we could not assess specificity. While higher sensitivity was realized under certain conditions, there is an unknown impact on specificity. We are in the process of completing analysis of a retrospective radiologist reader study that will allow us to assess the impact of CADE on radiologist sensitivity and specificity with digital mammography. Third, the results provide the performance of CADE at a given point in time. The algorithms tested were current at the time this study was

performed in 2008-2009 timeframe. Certainly newer versions of these algorithms would be available today resulting in perhaps different results. For reference we do provide results from R2 CADe at different sensitivity settings in the appendix. Standalone CADe sensitivity assessment is a critical first step in determining if a CADe algorithm is ready for clinical assessment. Our results show that CADe marked between 26.1%-58.7% more cancers than were detected by human readers alone, suggesting that the use of CADe with both screen-film and digital mammography could lead to improved breast cancer detection.

Appendix
Appendix:
Results

The following tables represent the full R2 screen-film dataset including all three sensitivity settings tested. R2-1 is the default sensitivity setting, with R2-2 having higher sensitivity than the default setting, and R2-3 having the highest sensitivity setting allowed by the system.

Table A-1

Patient and Lesion Characteristics for the R2 CADe Screen-Film Full Dataset

	SCREEN-FILM		
	R2-1(N=303) n (%)	R2-2(N=300) n (%)	R2-3(N=291) n (%)
AGE (years)			
Age<50	64 (21.1)	65 (21.7)	61 (21.0)
50<=Age<=64	150 (49.5)	147 (49.0)	145 (49.8)
Age>=65	89 (29.4)	88 (29.3)	85 (29.1)
BREAST DENSITY			
Dense Breasts	150 (49.5)	148 (49.3)	146 (50.2)
Fatty Breasts	153 (50.5)	152 (50.7)	145 (49.8)
MENOPAUSAL STATUS			
Pre or Peri-Menopausal	90 (30.1)	91 (30.7)	87 (30.3)
Post-Menopausal	209 (69.9)	205 (69.3)	200 (69.7)
HISTOLOGY			
DCIS	88 (29.1)	86 (28.8)	83 (28.6)
Invasive	214 (70.9)	213 (71.2)	207 (71.4)
TUMOR SIZE (mm)			
Size<=5	43 (19.7)	42 (19.4)	39 (18.6)
5<Size<=10	56 (25.7)	56 (25.9)	55 (26.2)
Size>10	119 (54.6)	118 (54.6)	116 (55.2)
LESION TYPE			
Mass	129 (48.9)	127 (48.7)	123 (48.2)
Asym Density	21 (8.0)	21 (8.0)	21 (8.2)
Calcification	99 (37.5)	98 (37.5)	96 (37.6)

	SCREEN-FILM		
	R2-1(N=303) n (%)	R2-2(N=300) n (%)	R2-3(N=291) n (%)
Arch Distortion	15 (5.7)	15 (5.7)	15 (5.9)

Table A-2

Sensitivity Estimates and False Positive Marks per Case for the R2 CADe Screen-Film Full Dataset

		Screen-Film		
		R2-1	R2-2	R2-3
Sensitivity	n/N	176/303	190/300	184/291
	Estimate	0.58	0.63	0.63
	95% CI	(0.52,0.64)	(0.58,0.69)	(0.57,0.69)
False Positives	N	303	300	291
	Mean	1.5	2.0	2.0
	SD	1.42	1.65	1.64
	Range	(0, 7)	(0,8)	(0,8)
	Median	1	2	2

n = number of correctly detected cancers

N = total number of cancer cases analyzable for each CADe system

95% CI = 95% exact confidence interval

SD =standard deviation

Table A-3
Sensitivities by Patient and Lesion Characteristics for the R2 CADe Screen-film Full Datasets

	SCREEN-FILM						
	R2-1		R2-2		R2-3		
	Sensitivity	95% CI	Sensitivity	95% CI	Sensitivity	95% CI	
AGE(years)	Age<50	0.53 (34/ 64)	(0.41, 0.65)	0.58 (38/ 65)	(0.46, 0.70)	0.59 (36/ 61)	(0.46, 0.71)
	50 Age 64	0.59 (88/ 150)	(0.51, 0.66)	0.63 (93/ 147)	(0.55, 0.71)	0.64 (93/ 145)	(0.56, 0.72)
	Age 65	0.61 (54/ 89)	(0.50, 0.70)	0.67 (59/ 88)	(0.57, 0.76)	0.65 (55/ 85)	(0.54, 0.74)
BREAST DENSITY	Dense	0.60 (90/ 150)	(0.52, 0.68)	0.64 (95/ 148)	(0.56, 0.71)	0.65 (95/ 146)	(0.57, 0.72)
	Fatty	0.56 (86/ 153)	(0.48, 0.64)	0.63 (95/ 152)	(0.55, 0.70)	0.61 (89/ 145)	(0.53, 0.69)
MENOPAUSAL STATUS	Pre/Peri	0.51 (46/ 90)	(0.41, 0.61)	0.58 (53/ 91)	(0.48, 0.68)	0.57 (50/ 87)	(0.47, 0.67)
	Post	0.61 (127/ 209)	(0.54, 0.67)	0.65 (134/ 205)	(0.59, 0.72)	0.66 (131/ 200)	(0.59, 0.72)
	DCIS	0.57 (50/ 88)	(0.46, 0.67)	0.63 (54/ 86)	(0.52, 0.72)	0.61 (51/ 83)	(0.51, 0.71)
HISTOLOGY	Invasive	0.59 (126/ 214)	(0.52, 0.65)	0.64 (136/ 213)	(0.57, 0.70)	0.64 (133/ 207)	(0.57, 0.70)
	Size 5	0.53 (23/ 43)	(0.39, 0.68)	0.64 (27/ 42)	(0.49, 0.77)	0.67 (26/ 39)	(0.51, 0.80)
TUMOR SIZE(mm)	5<Size 10	0.48 (27/ 56)	(0.36, 0.61)	0.55 (31/ 56)	(0.42, 0.68)	0.56 (31/ 55)	(0.43, 0.69)
	Size>10	0.66 (78/ 119)	(0.57, 0.74)	0.69 (81/ 118)	(0.60, 0.76)	0.68 (79/ 116)	(0.59, 0.76)
	Mass	0.60 (78/ 129)	(0.52, 0.69)	0.66 (84/ 127)	(0.57, 0.74)	0.67 (83/ 123)	(0.59, 0.75)
LESION TYPE	Asym Density	0.52 (11/ 21)	(0.32, 0.72)	0.57 (12/ 21)	(0.36, 0.76)	0.57 (12/ 21)	(0.36, 0.76)
	Calcification	0.73 (72/ 99)	(0.63, 0.81)	0.77 (75/ 98)	(0.67, 0.84)	0.76 (73/ 96)	(0.67, 0.84)
	Arch Distortion	0.53 (8/ 15)	(0.29, 0.76)	0.67 (10/ 15)	(0.41, 0.85)	0.67 (10/ 15)	(0.41, 0.85)

Table A-4

Univariate Logistic Regression Models by Patient and Lesion Characteristics for the R2 CADe Screen-film Full Datasets

	SCREEN-FILM											
	R2-1				R2-2				R2-3			
	OR	95% CI(OR)	p-value	OR	95% CI(OR)	p-value	OR	95% CI(OR)	p-value	OR	95% CI(OR)	p-value
AGE												
Age<50												
50 Age 64	1.25	(0.70, 2.26)	0.4538	1.22	(0.67, 2.22)	0.5071	1.24	(0.67, 2.29)	0.4882			
Age 65	1.36	(0.71, 2.61)	0.3520	1.45	(0.74, 2.81)	0.2768	1.27	(0.65, 2.51)	0.4844			
50 Age 64	1.09	(0.64, 1.86)	0.7600	1.18	(0.68, 2.06)	0.5575	1.03	(0.59, 1.79)	0.9308			
BREAST DENSITY												
Dense												
Fatty	0.86	(0.54, 1.35)	0.5038	0.93	(0.58, 1.49)	0.7615	0.85	(0.53, 1.37)	0.5142			
MENOPAUSAL STATUS												
Pre/Peri												
Post	1.48	(0.90, 2.44)	0.1218	1.35	(0.82, 2.25)	0.2417	1.40	(0.84, 2.35)	0.1961			
HISTOLOGY												
DCIS												
Invasive	1.09	(0.66, 1.80)	0.7408	1.05	(0.62, 1.76)	0.8629	1.13	(0.67, 1.91)	0.6540			
TUMOR SIZE (mm)												
5<Size 10	0.81	(0.37, 1.79)	0.6031	0.69	(0.30, 1.57)	0.3743	0.65	(0.28, 1.52)	0.3149			
Size>10	1.65	(0.81, 3.36)	0.1638	1.22	(0.58, 2.55)	0.6048	1.07	(0.49, 2.31)	0.8681			
5<Size 10	2.04	(1.07, 3.90)	0.0302	1.77	(0.92, 3.40)	0.0889	1.65	(0.85, 3.20)	0.1359			
LESION TYPE												
Asym Density	0.72	(0.28, 1.82)	0.4855	0.68	(0.27, 1.75)	0.4252	0.64	(0.25, 1.65)	0.3578			
Calcification	1.74	(0.99, 3.07)	0.0542	1.67	(0.92, 3.02)	0.0912	1.53	(0.84, 2.79)	0.1663			
Arch Distortion	0.75	(0.26, 2.19)	0.5949	1.02	(0.33, 3.18)	0.9676	0.96	(0.31, 3.01)	0.9494			
Asym Density	2.42	(0.92, 6.35)	0.0718	2.45	(0.92, 6.53)	0.0744	2.38	(0.89, 6.36)	0.0838			
Arch Distortion	1.04	(0.28, 3.92)	0.9550	1.50	(0.38, 5.95)	0.5641	1.50	(0.38, 5.95)	0.5640			
Calcification	0.43	(0.14, 1.30)	0.1335	0.61	(0.19, 1.98)	0.4133	0.63	(0.20, 2.03)	0.4398			

OR = odds ratio

REFERENCES

1. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med*. 2005; 353:1773–1783. [PubMed: 16169887]
2. 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*. 2002; 179:671–677. [PubMed: 12185042]
3. 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*. 2001; 218:873–880. [PubMed: 11230669]
4. Skaane P, Balleyguier C, Diekmann F, et al. Breast lesion detection and classification: comparison of screen-film mammography and full-field digital mammography with soft-copy reading--observer performance study. *Radiology*. 2005; 237:37–44. [PubMed: 16100086]
5. Skaane P, Kshirsagar A, Stapleton S, Young K, Castellino RA. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. *AJR Am J Roentgenol*. 2007; 188:377–384. [PubMed: 17242245]
6. 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*. 2004; 232:197–204. [PubMed: 15155893]
7. 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*. 2003; 229:877–884. [PubMed: 14576447]
8. Pisano ED, Gatsonis CA, Yaffe MJ, et al. American College of Radiology Imaging Network digital mammographic imaging screening trial: objectives and methodology. *Radiology*. 2005; 236:404–412. [PubMed: 15961755]
9. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology*. 2008; 246:376–383. [PubMed: 18227537]
10. Heddson B, Ronnow K, Olsson M, Miller D. Digital versus screen-film mammography: a retrospective comparison in a population-based screening program. *Eur J Radiol*. 2007; 64:419–425. [PubMed: 17383841]
11. Del Turco MR, Mantellini P, Ciatto S, et al. Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. *AJR Am J Roentgenol*. 2007; 189:860–866. [PubMed: 17885057]
12. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P. Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. *Eur Radiol*. 2008; 18:183–191. [PubMed: 17680246]
13. Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol*. 2009; 50:3–14. [PubMed: 19037825]
14. Breast Cancer Surveillance Consortium. [Accessed 08/17, 2010] Performance Measures for 3,884,059 Screening Mammography Examinations from 1996 to 2997 by Age. Available at:http://breastscreening.cancer.gov/data/performance/screening/perf_age.html
15. Balleyguier C, Kinkel K, Fermanian J, et al. Computer-aided detection (CAD) in mammography: does it help the junior or the senior radiologist? *Eur J Radiol*. 2005; 54:90–96. [PubMed: 15797297]
16. Hukkinen K, Vehmas T, Pamilo M, Kivisaari L. Effect of computer-aided detection on mammographic performance: experimental study on readers with different levels of experience. *Acta Radiol*. 2006; 47:257–263. [PubMed: 16613306]
17. Paquerault S, Samuelson FW, Petrick N, Myers KJ, Smith RC. Investigation of reading mode and relative sensitivity as factors that influence reader performance when using computer-aided detection software. *Acad Radiol*. 2009; 16:1095–1107. [PubMed: 19523855]

18. Karssemeijer N, Bluekens AM, Beijerinck D, et al. Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program. *Radiology*. 2009; 253:353–358. [PubMed: 19703851]
19. Nishikawa RM, Giger ML, Doi K, et al. Effect of case selection on the performance of computer-aided detection schemes. *Med Phys*. 1994; 21:265–269. [PubMed: 8177159]
20. Bolivar AV, Gomez SS, Merino P, et al. Computer-aided detection system applied to full-field digital mammograms. *Acta Radiol*. 2010; 51:1086–1092. [PubMed: 20883182]
21. The JS, Schilling KJ, Hoffmeister JW, Friedmann E, McGinnis R, Holcomb RG. Detection of breast cancer with full-field digital mammography and computer-aided detection. *AJR Am J Roentgenol*. 2009; 192:337–340. [PubMed: 19155392]
22. Yang SK, Moon WK, Cho N, et al. Screening mammography-detected cancers: sensitivity of a computer-aided detection system applied to full-field digital mammograms. *Radiology*. 2007; 244:104–111. [PubMed: 17507722]

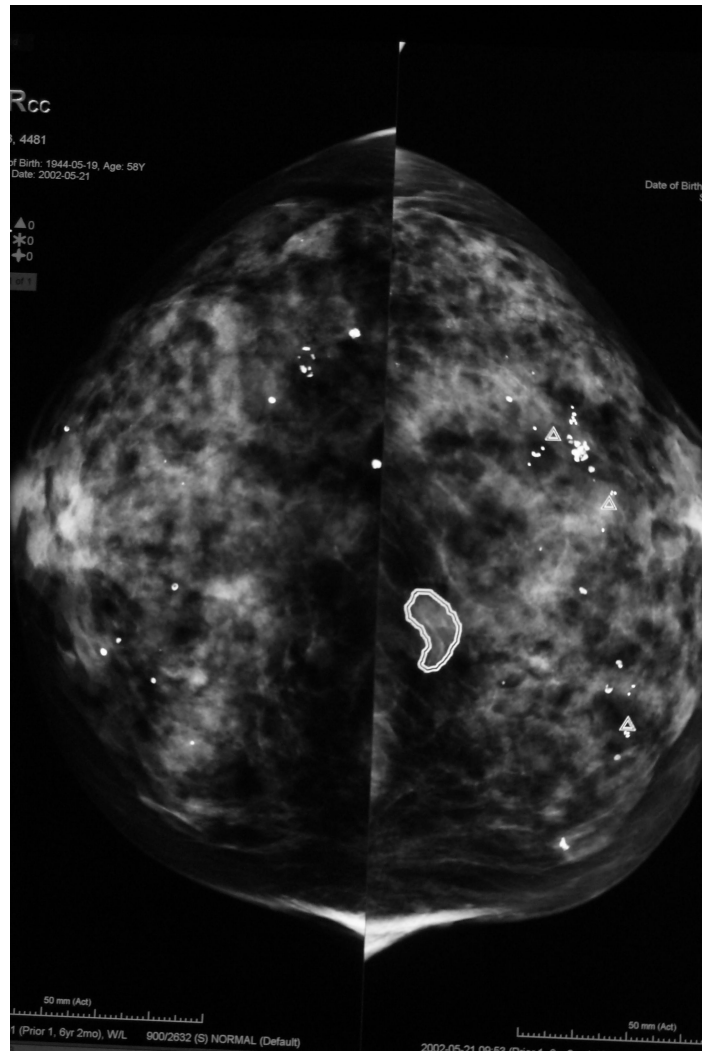


Figure 1. Representative image of CAD marks and digital mammographic image R2. Δ symbol indicates calcifications, while masses are circumscribed by a freeform shape.

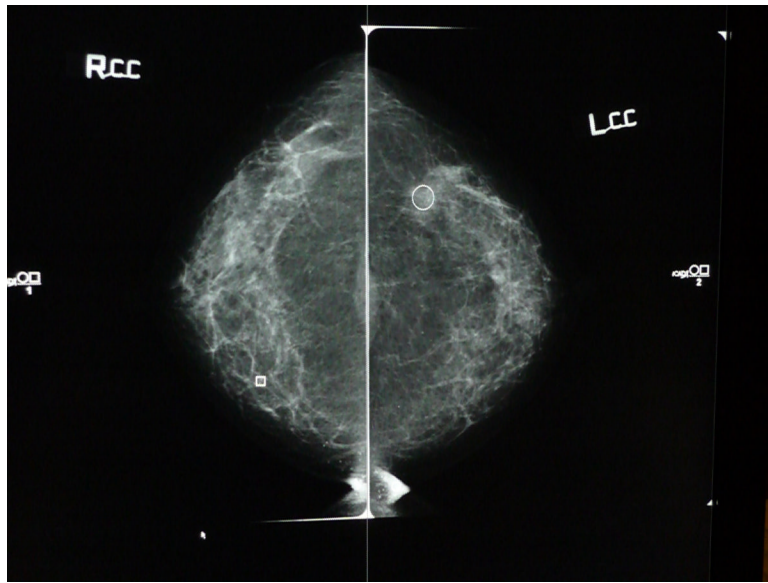


Figure 2. Screen shot of CC views of digital mammogram with iCAD report overlay. Square represents calcifications and circle represents mass finding.

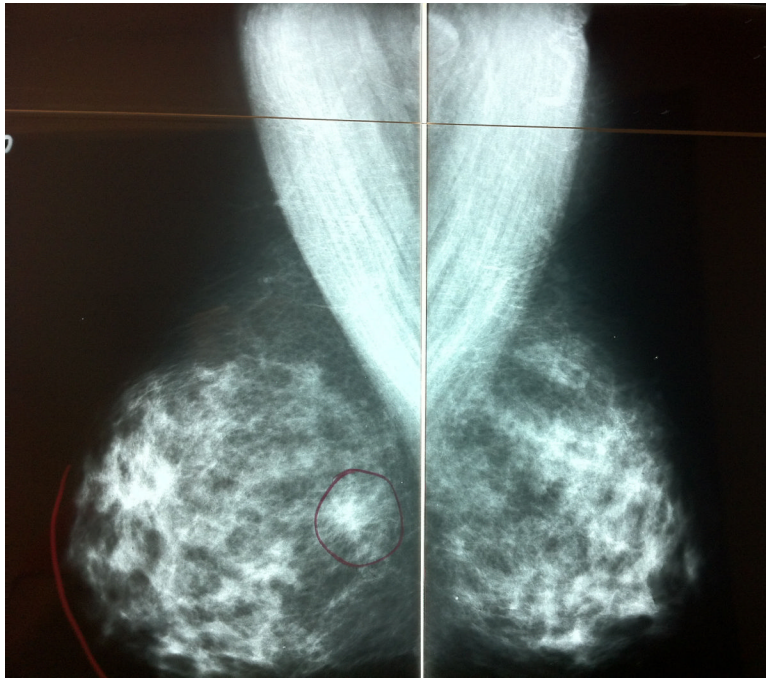


Figure 3.
Example of screen-film mammogram with annotated acetate overlay denoting location of cancer.

Table 1

Patient and Lesion Characteristics for the Common Cases (N=161)

	n (%)
AGE (years)	
Age<50	30 (18.6)
50<=Age<=64	81 (50.3)
Age>=65	50 (31.1)
BREAST DENSITY	
Dense Breasts	75 (46.6)
Fatty Breasts	86 (53.4)
MENOPAUSAL STATUS	
Pre or Peri-Menopausal	43 (26.7)
Post-Menopausal	114 (70.8)
Missing	4 (2.5)
HISTOLOGY	
DCIS	46 (28.6)
Invasive	115 (71.4)
TUMOR SIZE (mm)	
Size<=5	21 (13.0)
5<Size<=10	31 (19.3)
Size>10	69 (42.9)
Missing	40 (24.8)
LESION TYPE	
Mass	65 (40.4)
Asym Density	12 (7.5)
Calcification	60 (37.3)
Arch Distortion	8 (5.0)
Missing	16 (9.9)

Table 2

Sensitivity and False Positive Marks for the Common Cases

		iCAD	R2
Sensitivity	DG	0.74(119/161) 95% CI: (0.66,0.81)	0.74(119/161) 95% CI: (0.66,0.81)
	SF	0.69(111/161) 95% CI: (0.61,0.76)	0.60(97/161) 95% CI: (0.52,0.68)
	DGvSF	0.05 95% CI: (-0.03,0.13) p=0.2559	0.14 95% CI: (0.05,0.23) p=0.0026
False Positive	DG	2.57 ± 1.92	2.07 ± 1.57
	SF	3.06 ± 1.72	1.52 ± 1.45
	DGvSF	-0.49 95% CI: (-0.77,-0.21)p=0.0008	0.55 95% CI: (0.33, 0.78) p<0.0001

Table 3

Comparison of Sensitivities of CADe system and DMIST Study for the Common Cases

	iCAD		R2	
	DG	SF	DG	SF
CAD	0.74(119/161) 95% CI: (0.66,0.81)	0.69(111/161) 95% CI: (0.61,0.76)	0.74(119/161) 95% CI: (0.66,0.81)	0.60(97/161) 95% CI: (0.52,0.68)
DMIST	0.43(69/161) 95% CI: (0.35,0.51)	0.41(66/161) 95% CI: (0.33,0.49)	0.43(69/161) 95% CI: (0.35,0.51)	0.41(66/161) 95% CI: (0.33,0.49)
CAD v DMIST	0.31 95% CI: (0.22,0.40) p<.0001	0.28 95% CI: (0.19,0.37) p<.0001	0.31 95% CI: (0.22,0.40) p<.0001	0.19 95% CI: (0.10,0.28) p<.0001

Table 4

Sensitivities by Patient and Lesion Characteristics for the Common Cases

	iCAD				R2			
	DG		SF		DG		SF	
	sensitivity	95% CI	sensitivity	95% CI	sensitivity	95% CI	sensitivity	95% CI
AGE(years)								
Age<50	0.73 (22/ 30)	(0.55, 0.86)	0.63 (19/ 30)	(0.45, 0.78)	0.80 (24/ 30)	(0.62, 0.91)	0.53 (16/ 30)	(0.36, 0.70)
50 Age 64	0.72 (58/ 81)	(0.61, 0.80)	0.73 (59/ 81)	(0.62, 0.81)	0.74 (60/ 81)	(0.63, 0.82)	0.62 (50/ 81)	(0.51, 0.72)
Age 65	0.78 (39/ 50)	(0.64, 0.87)	0.66 (33/ 50)	(0.52, 0.78)	0.70 (35/ 50)	(0.56, 0.81)	0.62 (31/ 50)	(0.48, 0.74)
BREAST DENSITY								
Dense	0.76 (57/ 75)	(0.65, 0.84)	0.67 (50/ 75)	(0.55, 0.76)	0.77 (58/ 75)	(0.67, 0.85)	0.60 (45/ 75)	(0.49, 0.70)
Fatty	0.72 (62/ 86)	(0.62, 0.81)	0.71 (61/ 86)	(0.61, 0.80)	0.71 (61/ 86)	(0.61, 0.80)	0.60 (52/ 86)	(0.50, 0.70)
MENOPAUSAL STATUS								
Pre or Peri-Menopausal	0.72 (31/ 43)	(0.57, 0.83)	0.63 (27/ 43)	(0.48, 0.76)	0.74 (32/ 43)	(0.59, 0.85)	0.49 (21/ 43)	(0.34, 0.63)
Post-Menopausal	0.75 (85/ 114)	(0.66, 0.82)	0.71 (81/ 114)	(0.62, 0.79)	0.74 (84/ 114)	(0.65, 0.81)	0.64 (73/ 114)	(0.55, 0.72)
HISTOLOGY								
DCIS	0.74 (34/ 46)	(0.59, 0.85)	0.59 (27/ 46)	(0.44, 0.72)	0.74 (34/ 46)	(0.59, 0.85)	0.63 (29/ 46)	(0.48, 0.76)
Invasive	0.74 (85/ 115)	(0.65, 0.81)	0.73 (84/ 115)	(0.64, 0.80)	0.74 (85/ 115)	(0.65, 0.81)	0.59 (68/ 115)	(0.50, 0.68)
TUMOR SIZE (mm)								
Size 5	0.71 (15/ 21)	(0.49, 0.87)	0.76 (16/ 21)	(0.54, 0.90)	0.71 (15/ 21)	(0.49, 0.87)	0.48 (10/ 21)	(0.28, 0.68)
5<Size 10	0.74 (23/ 31)	(0.56, 0.87)	0.68 (21/ 31)	(0.50, 0.82)	0.81 (25/ 31)	(0.63, 0.91)	0.55 (17/ 31)	(0.37, 0.71)
Size>10	0.75 (52/ 69)	(0.64, 0.84)	0.71 (49/ 69)	(0.59, 0.80)	0.75 (52/ 69)	(0.64, 0.84)	0.67 (46/ 69)	(0.55, 0.77)
LESION TYPE								
Mass	0.69 (45/ 65)	(0.57, 0.79)	0.69 (45/ 65)	(0.57, 0.79)	0.72 (47/ 65)	(0.60, 0.82)	0.55 (36/ 65)	(0.43, 0.67)
Asym Density	0.50 (6/ 12)	(0.24, 0.76)	0.75 (9/ 12)	(0.45, 0.92)	0.50 (6/ 12)	(0.24, 0.76)	0.75 (9/ 12)	(0.45, 0.92)
Calcification	0.83 (50/ 60)	(0.72, 0.91)	0.73 (44/ 60)	(0.61, 0.83)	0.83 (50/ 60)	(0.72, 0.91)	0.73 (44/ 60)	(0.61, 0.83)
Arch Distortion	0.88 (7/ 8)	(0.46, 0.98)	0.63 (5/ 8)	(0.28, 0.87)	0.75 (6/ 8)	(0.38, 0.94)	0.63 (5/ 8)	(0.28, 0.87)

Table 5
Univariate Logistic Regression Models by Patient and Lesion Characteristics for the Common Cases

	iCAD				R2								
	DG		SF		DG		SF						
	OR	95% CI(OR)	p-value	OR	95% CI(OR)	p-value	OR	95% CI(OR)	p-value				
AGE(years)													
Age<50	50 Age 64	0.92	(0.36, 2.35)	0.8570	1.55	(0.64, 3.78)	0.3323	0.71	(0.26, 1.99)	0.5196	1.41	(0.61, 3.29)	0.4246
	Age 65	1.29	(0.45, 3.68)	0.6354	1.12	(0.44, 2.89)	0.8087	0.58	(0.20, 1.72)	0.3281	1.43	(0.57, 3.57)	0.4466
50 Age 64	Age 65	1.41	(0.62, 3.21)	0.4184	0.72	(0.34, 1.55)	0.4064	0.82	(0.37, 1.79)	0.6121	1.01	(0.49, 2.09)	0.9752
BREAST DENSITY													
Dense	Fatty	0.82	(0.40, 1.66)	0.5736	1.22	(0.63, 2.38)	0.5600	0.72	(0.35, 1.46)	0.3570	1.02	(0.54, 1.92)	0.9520
MENOPAUSAL STATUS													
Pre/Peri	Post	1.13	(0.52, 2.50)	0.7536	1.45	(0.69, 3.05)	0.3204	0.96	(0.43, 2.15)	0.9259	1.87	(0.92, 3.79)	0.0852
HISTOLOGY													
DCIS	Invasive	1.00	(0.46, 2.18)	1.0000	1.91	(0.93, 3.91)	0.0776	1.00	(0.46, 2.18)	1.0000	0.85	(0.42, 1.72)	0.6469
TUMOR SIZE(mm)													
Size 5	5<Size 10	1.15	(0.33, 3.98)	0.8255	0.66	(0.19, 2.30)	0.5108	1.67	(0.45, 6.12)	0.4413	1.34	(0.44, 4.06)	0.6095
	Size>10	1.22	(0.41, 3.65)	0.7177	0.77	(0.25, 2.37)	0.6435	1.22	(0.41, 3.65)	0.7176	2.20	(0.82, 5.93)	0.1193
5-Size 10	Size>10	1.06	(0.40, 2.82)	0.9007	1.17	(0.47, 2.91)	0.7413	0.73	(0.26, 2.09)	0.5626	1.65	(0.69, 3.92)	0.2591
LESION TYPE													
Mass	Asym Density	0.44	(0.13, 1.55)	0.2029	1.33	(0.33, 5.45)	0.6890	0.38	(0.11, 1.34)	0.1340	2.42	(0.60, 9.75)	0.2153
	Calcification	2.22	(0.94, 5.25)	0.0686	1.22	(0.56, 2.66)	0.6131	1.91	(0.80, 4.57)	0.1431	2.21	(1.04, 4.70)	0.0384*
	Arch Distortion	3.11	(0.36, 26.99)	0.3032	0.74	(0.16, 3.40)	0.6997	1.15	(0.21, 6.23)	0.8721	1.34	(0.30, 6.09)	0.7026
Asym Density	Calcification	5.00	(1.34, 18.71)	0.0168*	0.92	(0.22, 3.82)	0.9049	5.00	(1.34, 18.71)	0.0168*	0.92	(0.22, 3.82)	0.9051
	Arch Distortion	7.00	(0.65, 75.73)	0.1093	0.56	(0.08, 3.86)	0.5522	3.00	(0.42, 21.30)	0.2719	0.56	(0.08, 3.86)	0.5525
Calcification	Arch Distortion	1.40	(0.15, 12.67)	0.7647	0.61	(0.13, 2.83)	0.5242	0.60	(0.11, 3.41)	0.5647	0.61	(0.13, 2.83)	0.5244

OR: odds ratio