Original Investigation | LESS IS MORE

Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection

Constance D. Lehman, MD, PhD; Robert D. Wellman, MS; Diana S. M. Buist, PhD; Karla Kerlikowske, MD; Anna N. A. Tosteson, ScD; Diana L. Miglioretti, PhD; for the Breast Cancer Surveillance Consortium

IMPORTANCE After the US Food and Drug Administration (FDA) approved computer-aided detection (CAD) for mammography in 1998, and the Centers for Medicare and Medicaid Services (CMS) provided increased payment in 2002, CAD technology disseminated rapidly. Despite sparse evidence that CAD improves accuracy of mammographic interpretations and costs over $400 million a year, CAD is currently used for most screening mammograms in the United States.

OBJECTIVE To measure performance of digital screening mammography with and without CAD in US community practice.

DESIGN, SETTING, AND PARTICIPANTS We compared the accuracy of digital screening mammography interpreted with (n = 495 818) vs without (n = 129 807) CAD from 2003 through 2009 in 323 973 women. Mammograms were interpreted by 271 radiologists from 66 facilities in the Breast Cancer Surveillance Consortium. Linkage with tumor registries identified 3159 breast cancers in 323 973 women within 1 year of the screening.

MAIN OUTCOMES AND MEASURES Mammography performance (sensitivity, specificity, and screen-detected and interval cancers per 1000 women) was modeled using logistic regression with radiologist-specific random effects to account for correlation among examinations interpreted by the same radiologist, adjusting for patient age, race/ethnicity, time since prior mammogram, examination year, and registry. Conditional logistic regression was used to compare performance among 107 radiologists who interpreted mammograms both with and without CAD.

RESULTS Screening performance was not improved with CAD on any metric assessed. Mammography sensitivity was 85.3% (95% CI, 83.6%-86.9%) with and 87.3% (95% CI, 84.5%-89.7%) without CAD. Specificity was 91.6% (95% CI, 91.0%-92.2%) with and 91.4% (95% CI, 90.6%-92.0%) without CAD. There was no difference in cancer detection rate (4.1 in 1000 women screened with and without CAD). Computer-aided detection did not improve intraradiologist performance. Sensitivity was significantly decreased for mammograms interpreted with vs without CAD in the subset of radiologists who interpreted both with and without CAD (odds ratio, 0.53; 95% CI, 0.29-0.97).

CONCLUSIONS AND RELEVANCE Computer-aided detection does not improve diagnostic accuracy of mammography. These results suggest that insurers pay more for CAD with no established benefit to women.

Invited Commentary
page 1837

Author Affiliations: Department of Radiology, Massachusetts General Hospital, Boston (Lehman); Group Health Research Institute, Seattle, Washington (Wellman, Buist, Miglioretti); Departments of Medicine and Epidemiology and Biostatistics, University of California, San Francisco, San Francisco (Kerlikowske); Norris Cotton Cancer Center, Geisel School of Medicine at Dartmouth, Dartmouth College, Lebanon, New Hampshire (Tosteson); Department of Public Health Sciences, School of Medicine, University of California, Davis (Miglioretti).

Corresponding Author: Constance D. Lehman, MD, PhD, Department of Radiology, Massachusetts General Hospital, Avon Comprehensive Breast Evaluation Center, 55 Fruit St, WAC 240, Boston, MA 02114 (clehman@mgh.harvard.edu).

JAMA Intern Med. 2015;175(11):1828-1837. doi:10.1001/jamainternmed.2015.5231
Published online September 28, 2015.

Computer-aided detection (CAD) for mammography is intended to assist radiologists in identifying subtle cancers that might otherwise be missed. Computer-aided detection marks potential areas of concern on the mammogram, and the radiologist determines whether the area warrants further evaluation. Although CAD for mammography was approved by the US Food and Drug Administration (FDA) in 1998,[1] by 2001, less than 5% of screening mammograms were interpreted with CAD in the United States. However, in 2002, the Centers for Medicare and Medicaid Services (CMS) increased reimbursement for CAD, and by 2008, 74% of all screening mammograms in the Medicare population were interpreted with CAD.[2,3]

Measuring the true impact of CAD on the accuracy of mammographic interpretation has proved challenging. Findings on potential benefits and harms are inconsistent and contradictory.[4-19] Study designs include reader studies[4-7] of enriched case sets; prospective "sequential reading" clinical studies[8-12] in which a radiologist records a mammogram interpretation without CAD assistance, then immediately reviews and records an interpretation with CAD assistance; and retrospective observational studies[13-16] using historical controls. One large European trial[17] used a randomized clinical trial design to compare mammographic interpretations by a single reader with CAD compared with double readings without CAD.

Comparisons of mammography interpretations with vs without CAD in US community practice have not supported improved performance with CAD.[18,19] However, these studies were limited by relatively small numbers and a focus on older women. Another limitation was that CAD technology was studied relatively early in its adoption, so examinations were interpreted during the early part of radiologists' learning curves and included examinations with outdated film screen mammography. Our study addresses these limitations by using a large database of more than 495 000 full-field digital screening mammograms interpreted with CAD, accounting for radiologists' early learning curves, and adjusting for patient and radiologist variables. We also assessed performance within a subset of radiologists who interpreted with and without CAD during the study period.

Methods

Data Source

Data were pooled from 5 mammography registries that participate in the Breast Cancer Surveillance Consortium (BCSC)[20] funded by the National Cancer Institute: (1) the San Francisco Mammography Registry, (2) the New Mexico Mammography Advocacy Project, (3) the Vermont Breast Cancer Surveillance System, (4) the New Hampshire Mammography Network, and (5) the Carolina Mammography Registry. Each mammography registry links women to a state tumor registry or regional Surveillance Epidemiology and End Results program that collects population-based cancer data. Each registry and the BCSC Statistical Coordinating Center have institutional review board approval for either active or passive consenting processes or a waiver of consent to enroll participants, link data, and perform analytic studies. All procedures are Health Insurance Portability and Accountability Act compliant, and all registries and the Statistical Coordinating Center have received a Federal Certificate of Confidentiality and other protection for the identities of women, physicians, and facilities that participate in this research.

Participants

We included digital screening mammography examinations interpreted by 271 radiologists with (n = 495 818) or without CAD (n = 129 807) between January 1, 2003, and December 31, 2009, among 323 973 women aged 40 to 89 years with information on race, ethnicity, and time since last mammogram. Of the radiologists, 82 never used CAD, 82 always used CAD, and 107 sometimes used CAD. The latter 107 radiologists contributed 45 990 examinations interpreted without using CAD and 337 572 interpreted using CAD. The median percentage of examinations interpreted using CAD among the 107 radiologists was 93%, and the interquartile range was 31%.

Data Collection

Methods used to identify and assess screening mammograms, patient characteristics, and outcomes have been described previously.[20,21] Briefly, screening mammograms were defined as bilateral mammograms designated as "routine screening" by the radiologist. Mammographic assessments followed the Breast Imaging Reporting and Data System (BI-RADS) of 0, additional imaging; 1, negative; 2, benign finding; 3, probably benign finding; 4, suspicious abnormality; or 5, abnormality highly suspicious for malignant neoplasm.[22]

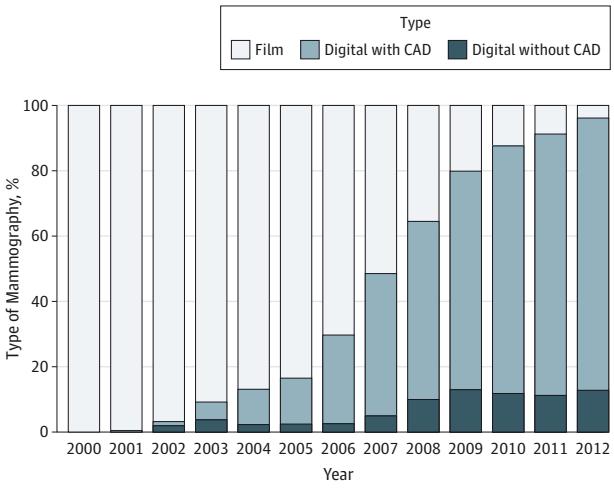
Woman-level characteristics including menopausal status, race/ethnicity, and first-degree family history were captured through self-administered questionnaires at each examination. Breast density was recorded by the radiologist at the time of the mammogram using the BI-RADS standard terminology of almost entirely fat, scattered fibroglandular densities, heterogeneously dense, and extremely dense.[23]

Outcomes

We calculated sensitivity, specificity, cancer detection rates, and interval cancer rates. We defined positive mammograms as those with BI-RADS assessments of 0, 4 or 5, or 3 with a recommendation for immediate follow-up. Negative mammogram results were defined as BI-RADS assessments 1 or 2, or 3 without a recommendation for immediate follow-up. All women were followed for breast cancer from their mammogram up until their next screening mammogram or 12 months, whichever came first. Breast cancer diagnoses included ductal carcinoma in situ (DCIS) or invasive breast cancer within this follow-up period.

False-negative examination results were defined as mammograms with a negative assessment but a breast can-

Figure 1. Screening Mammography Patterns From 2000 to 2012 in US Community Practices in the Breast Cancer Surveillance Consortium (BCSC)



Data are provided from the larger BCSC population including all screening mammograms (5.2 million mammograms) for the indicated time period.

cer diagnosis within the follow-up period. True-positive examination results were defined as those with a positive examination assessment and breast cancer diagnosis. False-positive examination results were examinations with a positive assessment but no cancer diagnosis. True-negative examination results had a negative assessment and no cancer diagnosis. Sensitivity was calculated as the number of true-positive mammogram results over the total number of breast cancers. For calculations of sensitivity, radiologists who interpreted no mammograms associated with cancer during the study period (n = 136) were excluded. Specificity was calculated as the number of true-negative mammogram results over the total number of mammograms without a breast cancer diagnosis. Cancer detection rate was defined as the number of true-positive examination results over the total number of mammograms, and interval cancer rate was the number of false-negative examination results over the total number of mammograms, reported per 1000 mammograms.[24]

Statistical Analysis

All analyses were conducted using the screening examination as the unit of analysis and allowing women to contribute multiple examinations during the study period; however, only 1 screening examination was associated with a breast cancer diagnosis. Distributions of breast cancer risk factors, demographic characteristics of examinations, and mammographic density and assessments were computed separately by CAD use vs no use.

We evaluated the diffusion of digital screening mammography with and without CAD in the larger BCSC population from 2002 through 2012 including 5.2 million screening mammograms.

Mammography performance measures were modeled using logistic regression, including normally distributed, radiologist-specific random effects to account for the correlation among examinations read by the same radiologist. Random effects were allowed to vary by CAD use or nonuse during the reading. Performance measures were estimated at the median of the random effects distribution. Adjusted, radiologist-specific relative performance was measured by an odds ratio (OR) with 95% CIs comparing CAD use to no CAD, adjusting for patient age at diagnosis, time since last mammogram and year of examination, and the BCSC registry.

Receiver operating characteristic (ROC) curves were estimated from 135 radiologists who interpreted at least 1 mammogram associated with a cancer diagnosis using a hierarchical logistic regression model that allowed the threshold and accuracy parameters to depend on whether CAD was used during examination interpretation. We assumed a constant accuracy among radiologists for examinations interpreted under the same condition (with or without CAD) and allowed the threshold for recall to vary across radiologists through normally distributed, radiologist-specific random effects that varied by whether the radiologist used CAD during the reading.[25] We estimated the normalized partial area under the summary ROC curves across the observed range of false-positive rates from this model.[26] We plotted the true-positive rate vs the false-positive rate and superimposed the estimated ROC curves.

Two separate main sensitivity analyses were conducted in subsets of total examinations: (1) to account for a possible learning curve for using CAD, we excluded the first year of each radiologist's CAD use; and (2) to estimate the within-radiologist effect of CAD, we limited analysis to the 107 radiologists who interpreted mammograms during the study period with and without CAD, using conditional logistic regression and adjusting for patient age, time since last mammogram, and race/ethnicity.

Two-sided statistical tests were used with $P < .05$ considered statistically significant. All analyses were conducted by one of us (R.D.W.) using SAS statistical software (version 9.2; SAS Institute Inc for Windows 7).

Results

Increase in Digital Screening Mammagraphy and CAD Use

Digital screening mammography and CAD use increased from 2000 to 2012. In 2003, only 5% of all screening mammograms in the BCSC were digital with CAD; by 2012, 83% of all screening mammograms were acquired digitally and interpreted with CAD assistance (**Figure 1**).

Among 323 973 women ages 40 to 89 years, 625 625 digital screening mammography examinations were performed (495 818 interpreted with CAD and 129 807 without CAD) between 2003 and 2009 by 271 radiologists. Breast cancer was diagnosed in 3159 women within 12 months of the screening mammogram and prior to the next screening mammogram. Women undergoing screening mammography with and without CAD assistance were similar in age,

Table 1. Characteristics of Women Undergoing Digital Screening Mammography With and Without CAD

Characteristic	No. (%)		
	CAD	No CAD	Overall
Age, y			
40-49	147 486 (29.8)	36 503 (28.1)	183 989 (29.4)
50-59	158 780 (32.0)	44 766 (34.5)	203 546 (32.5)
60-69	108 329 (21.9)	29 914 (23.0)	138 243 (22.1)
70-79	60 545 (12.2)	14 656 (11.3)	75 201 (12.0)
80-89	20 678 (4.2)	3968 (3.1)	24 646 (3.9)
Menopausal status			
Premenopausal	120 559 (30.3)	34 688 (32.9)	155 247 (30.9)
Postmenopausal, currently taking HT	33 764 (8.5)	6338 (6.0)	40 102 (8.0)
Postmenopausal, not currently taking HT	243 105 (61.2)	64 335 (61.1)	307 440 (61.2)
Missing	98 390 (19.8)	24 446 (18.8)	122 836 (19.6)
Race/ethnicity			
Non-Hispanic			
White	410 385 (82.8)	63 306 (48.8)	473 691 (75.7)
Black	28 533 (5.8)	7985 (6.2)	36 518 (5.8)
Asian/Pacific Islander	35 081 (7.1)	43 991 (33.9)	79 072 (12.6)
American Indian or Alaska Native	1194 (0.2)	590 (0.5)	1784 (0.3)
Other	7228 (1.5)	2497 (1.9)	9725 (1.6)
Hispanic	13 397 (2.7)	11 438 (8.8)	24 835 (4.0)
First-degree family history of breast cancer			
No	412 071 (83.6)	110 544 (87.4)	522 615 (84.4)
Yes	80 800 (16.4)	15 902 (12.6)	96 702 (15.6)
Missing	2947 (0.6)	3361 (2.6)	6308 (1.0)
Time since last mammogram			
No prior mammogram	12 518 (2.5)	6750 (5.2)	19 268 (3.1)
1 y	361 842 (73.0)	74 687 (57.5)	436 529 (69.8)
2 y	68 905 (13.9)	31 131 (24.0)	100 036 (16.0)
≥3 y	52 553 (10.6)	17 239 (13.3)	69 792 (11.2)
BI-RADS density			
Almost entirely fat	52 875 (12.4)	8833 (11.4)	61 708 (12.2)
Scattered fibroglandular densities	175 579 (41.1)	33 473 (43.1)	209 052 (41.4)
Heterogeneously dense	167 506 (39.2)	30 104 (38.7)	197 610 (39.1)
Extremely dense	31 252 (7.3)	5305 (6.8)	36 557 (7.2)
Missing	68 606 (13.8)	52 092 (40.1)	120 698 (19.3)

Abbreviations: BI-RADS, Breast Imaging Reporting and Data System; CAD, computer-aided detection; hormonal therapy; HT, hormonal therapy.

menopausal status, family history of breast cancer, time since last mammogram, and breast density. Women undergoing screening mammography with CAD were more likely to be non-Hispanic white than women whose mammograms were interpreted without CAD (**Table 1**).

Performance Measures for Mammography Interpreted With and Without CAD

Overall

Diagnostic accuracy was not improved with CAD on any performance metric assessed. Sensitivity of mammography was 85.3% (95% CI, 83.6%-86.9%) with and 87.3% (95% CI, 84.5%-89.7%) without CAD. Sensitivity of mammography for invasive cancer was 82.1% (95% CI, 80.0%-84.0%) with and 85.0% (95% CI, 81.5%-87.9%) without CAD; for DCIS, sensitivity was 93.2% (95% CI, 91.1%-94.9%) with and 94.3% (95% CI, 89.4%-97.1%) without CAD. Specificity of mammography was 91.6% (95% CI, 91.0%-92.2%) with and 91.4% (95% CI, 90.6%-92.0%) without CAD. There was no difference in overall cancer detection rate (4.1 cancers per 1000 women screened with CAD and without CAD) or in invasive cancer detection rate (2.9 vs 3.0 per 1000 women screened with CAD and without CAD). However, the DCIS detection rate was higher in patients whose mammograms were assessed with CAD compared with those whose mammograms were assessed without CAD (1.2 vs 0.9 per 1000; 95% CI, 1.0-1.9; $P < .03$) (**Table 2**).

To allow for the possibility that performance improved after the first year of CAD use by a radiologist, and to account for any possible learning curve, we excluded the first year of mammographic interpretations with CAD for individual radiologists and found no differences for any of our performance measurements (data not shown).

Table 2. Performance Measures of Digital Screening Mammography With and Without CAD

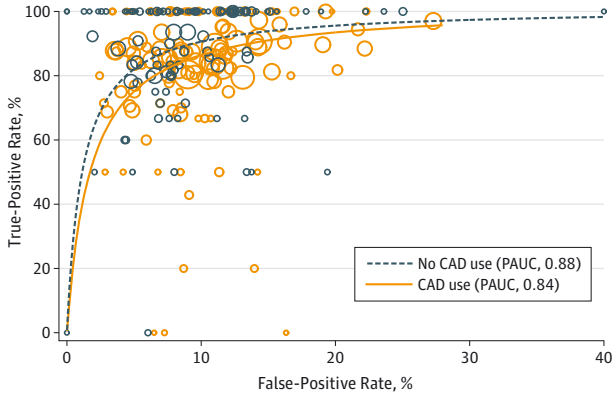
Measure	CAD, No.		No CAD, No.		Mean (95% CI)			
	Events	Exams	Events	Exams	CAD	No CAD	AOR (95% CI)[a]	*P* Value
Cancers detected per 1000 exams								
Total	2145	495 818	558	129 807	4.1 (3.8-4.4)	4.1 (3.6-4.6)	0.99 (0.84-1.15)	.86
Invasive	1485	495 818	408	129 807	2.9 (2.7-3.1)	3.1 (2.7-3.5)	0.92 (0.77-1.08)	.30
DCIS	660	495 818	150	129 807	1.2 (1.0-1.3)	1.0 (0.7-1.2)	1.39 (1.03-1.87)	.03
Interval cancers per 1000 exams								
Total	375	495 818	81	129 807	0.8 (0.7-0.8)	0.6 (0.5-0.8)	1.14 (0.87-1.50)	.33
Invasive	327	495 818	72	129 807	0.7 (0.6-0.7)	0.6 (0.4-0.7)	1.09 (0.82-1.46)	.54
DCIS	48	495 818	9	129 807	0.1 (0.1-0.1)	0.0 (0.0-0.2)	1.59 (0.72-3.51)	.25
Sensitivity[b]								
Total	2145	2520	558	639	85.3 (83.6-86.9)	87.3 (84.5-89.7)	0.81 (0.60-1.10)	.18
Invasive	1485	1812	408	480	82.1 (80.0-84.0)	85.0 (81.5-87.9)	0.83 (0.59-1.17)	.28
DCIS	660	708	150	159	93.2 (91.1-94.9)	94.3 (89.4-97.1)	0.88 (0.37-2.07)	.76
Specificity[b]								
Total	444 356	493 298	118 025	129 168	91.6 (91.0-92.2)	91.4 (90.6-92.0)	1.02 (0.94-1.11)	.58
Invasive	444 404	494 006	118 034	129 327	91.5 (90.9-92.1)	91.3 (90.5-91.9)	1.02 (0.94-1.11)	.58
DCIS	444 683	495 110	118 097	129 648	91.4 (90.7-92.0)	91.0 (90.3-91.7)	1.04%(0.96-1.13)	.36
Recall rate per 100 exams	51 087	495 818	11 701	129 807	8.7 (8.1-9.4)	9.1 (8.4-9.8)	0.96%(0.89-1.04)	.35

Abbreviations: AOR, adjusted odds ratio; CAD, computer-aided detection; DCIS, ductal carcinoma in situ; exam, examination.

[a] Odds ratio for CAD vs No CAD adjusted for site, age group, race/ethnicity, time since prior mammogram, and calendar year of the examination using mixed-effects model with random effect for examination reader and varying with CAD use.

[b] The 95% CIs for sensitivity and specificity are given as percentages.

Figure 2. Receiver Operating Characteristic Curves for Digital Screening Mammography With and Without the Use of CAD, Estimated From 135 Radiologists Who Interpreted at Least 1 Examination Associated With Cancer



Each circle represents the true-positive or false-positive rate for a single radiologist, for examinations interpreted with (orange) or without (blue) computer-aided detection (CAD). Circle size is proportional to the number of mammograms associated with cancer interpreted by that radiologist with or without CAD. PAUC indicates partial area under the curve.

From the ROC analysis, the accuracy of mammographic interpretations with CAD was significantly lower than for those without CAD (P = .002). The normalized partial area under the summary ROC curve was 0.84 for interpretations with CAD and 0.88 for interpretations without CAD (**Figure 2**). In this subset of 135 radiologists who interpreted at least 1 mammogram associated with a cancer diagnosis, sensitivity of mammography was 84.9% (95% CI, 82.9%-86.9%) with and 89.3%% (95% CI, 86.9%-91.7%) without CAD. Specificity of mammography was 91.1% (95% CI, 90.4%-91.8%) with and 91.3% (95% CI, 90.5%-92.1%) without CAD.

Differences by Age, Breast Density, Menopausal Status, and Time Since Last Mammogram

We found no differences in diagnostic accuracy of mammographic interpretations with and without CAD in any of the subgroups assessed, including patient age, breast density, menopausal status, and time since last mammogram (**Table 3**).

Intraradiologist Performance Measures for Mammography With and Without CAD

Among 107 radiologists who interpreted mammograms both with and without CAD, intraradiologist performance was not improved with CAD, and CAD was associated with decreased sensitivity. Sensitivity of mammography was 83.3% (95% CI, 81.0%-85.6%) with and 89.6% (95% CI, 86.0%-93.1%) without CAD. Specificity of mammography was 90.7% (95% CI, 89.8%-91.7%) with and 89.6% (95% CI, 88.6%-91.1%) without CAD. The OR for specificity between mammograms interpreted with CAD and those interpreted without CAD by the same radiologist was 1.02 (95% CI, 0.99-1.05). Sensitivity was significantly decreased for mammograms interpreted

Table 3. Performance Measures of Digital Screening Mammography With and Without CAD, by Examination-Level Patient Characteristics

Measure	CAD		No CAD		Mean (95% CI)		AOR (95% CI)[a]	*P* Value
	Events	Exams	Events	Exams	CAD	No CAD		
By Age, y								
Cancers detected per 1000 exams								
40-49	419	147 486	107	36 503	2.7 (2.4-3.1)	2.6 (2.0-3.4)	1.12 (0.80-1.57)	.50
50-73	1358	295 392	383	82 000	4.3 (4.0-4.7)	4.5 (4.0-5.2)	0.94 (0.79-1.11)	.46
Sensitivity[b]								
40-49	419	515	107	126	81.6 (77.4-85.2)	89.9 (74.2-96.5)	0.62 (0.24-1.61)	.32
50-73	1358	1581	383	437	85.9 (84.1-87.6)	87.6 (84.2-90.4)	0.87 (0.57-1.32)	.50
Specificity[b]								
40-49	127 519	146 971	32 228	36 377	88.7 (87.8-89.6)	89.1 (88.1-90.0)	0.99 (0.90-1.09)	.80
50-73	267 865	293 811	75 251	81 563	92.3 (91.7-92.9)	92.2 (91.5-92.8)	1.01 (0.92-1.10)	.88
By BI-RADS Breast Density								
Cancers detected per 1000 exams								
Almost entirely fat	147	52 875	34	8833	2.8 (2.4-3.3)	3.8 (2.7-5.4)	0.58 (0.33-1.03)	.06
Scattered fibroglandular densities	717	175 579	135	33 473	3.8 (3.3-4.2)	4 (3.4-4.8)	0.86 (0.69-1.07)	.17
Heterogeneously dense	783	167 506	123	30,104	4.5 (4.1-5.0)	3.9 (3.1-4.9)	1.11 (0.86-1.44)	.43
Extremely dense	102	31 252	20	5305	2.7 (1.9-3.7)	1.7 (0.5-5.4)	1.72 (0.46-6.34)	.42
Sensitivity[b]								
Almost entirely fat	147	163	34	36	90.2 (84.4-94.0)	100 (91.4-100)	1.26 (1.01-1.56)	.04
Scattered fibroglandular densities	717	810	135	151	89.0 (86.1-91.4)	89.4 (83.3-93.4)	0.73 (0.37-1.41)	.34
Heterogeneously dense	783	949	123	143	82.5 (79.9-84.8)	86.0 (79.2-90.8)	0.86 (0.48-1.56)	.62
Extremely dense	102	144	20	26	72.1 (60.5-81.4)	77.8 (51.4-92.0)	0.85 (0.24-3.04)	.80
Specificity[b]								
Almost entirely fat	49 864	52 712	8330	8797	95.8 (95.2-96.3)	94.7 (93.8-95.4)	0.77 (0.63-0.94)	.01
Scattered fibroglandular densities	158 575	174 769	30 230	33 322	92.1 (91.4-92.7)	92.0 (91.2-92.7)	1.01 (0.92-1.12)	.80
Heterogeneously dense	146 180	166 557	26, 510	29 961	89.2 (88.3-90.1)	89.2 (88.1-90.2)	1.03 (0.92-1.15)	.63
Extremely dense	27 930	31 108	4724	5279	91.4 (90.2-92.4)	89.5 (88.0-90.8)	1.30 (1.09-1.55)	.003
By Menopausal Status								
Cancers detected per 1000 exams								
Premenopausal	401	120 559	117	34 688	3.3 (3 0-3.7)	2.9 (2.1-3.9)	1.16 (0.85-1.57)	.36
Postmenopausal, currently taking HT	204	33 764	44	6338	6 (5.3-6.9)	6.2 (3.9-10.0)	0.88 (0.47-1.65)	.69
Postmenopausal, not currently taking HT	1217	243 105	304	64 335	4.7 (4.3-5.1)	4.6 (4.0-5.3)	0.96 (0.80-1.15)	.66
Sensitivity[b]								
Premenopausal	401	484	117	141	83.0 (79.0-86.4)	84.2 (74.5-90.7)	1.06 (0.52-2.17)	.87
Postmenopausal, currently taking HT	204	243	44	51	84.0 (78.7-88.1)	86.3 (73.6-93.4)	0.93 (0.33-2.62)	.89
Postmenopausal, not currently taking HT	1217	1408	304	343	86.4 (84.5-88.1)	90.3 (84.7-94.0)	0.72 (0.41-1.27)	.26
Specificity[b]								
Premenopausal	102 940	120 075	30 505	34 547	87.9 (86.9-88.8)	88.3 (87.4-89.2)	0.99 (0.90-1.09)	.84
Postmenopausal, currently taking HT	30 129	33 521	5701	6287	90.7 (89.8-91.6)	91.3 (90.1-92.3)	0.88 (0.75-1.04)	.13
Postmenopausal, not currently taking HT	222 887	241 697	59 263	63 992	93.2 (92.6-93.7)	92.7 (92.1-93.3)	1.05 (0.96-1.16)	.27

(continued)

Table 3. Performance Measures of Digital Screening Mammography With and Without CAD, by Examination-Level Patient Characteristics (continued)

Measure	CAD		No CAD		Mean (95% CI)		AOR (95% CI)[a]	*P* Value
	Events	Exams	Events	Exams	CAD	No CAD		
By Time Since Prior Mammogram								
Cancers detected per 1000 exams								
No prior mammogram	78	12 518	28	6750	6.0 (4.5-8.1)	4.1 (2.8-6.0)	1.53 (0.90-2.59)	.11
1 y	1354	361 842	278	74 687	3.5 (3.2-3.8)	3.5 (3.0-4.1)	0.98 (0.82-1.18)	.84
2 y	308	68 905	142	31 131	4.1 (3.5-4.8)	4.4 (3.5-5.5)	0.75 (0.55-1.04)	.08
≥3 y	405	52 553	110	17 239	7.5 (6.7-8.5)	6.0 (4.7-7.6)	1.17 (0.87-1.57)	.31
Sensitivity[b]								
No prior mammogram	78	86	28	35	100.0 (95.9-100.0)	83.2 (47.4-96.5)	NE	
1 y	1354	1636	278	331	83.2 (80.7-85.4)	84.0 (79.6-87.6)	0.93 (0.64-1.35)	.69
2 y	308	350	142	156	88.0 (84.1-91.0)	95.2 (80.5-99.0)	0.45 (0.20-1.00)	.049
≥3 y	405	448	110	117	90.4 (87.3-92.8)	100.0 (96.9-100)	0.95 (0.83-1.08)	.42
Specificity[b]								
No prior mammogram	9333	12 432	5638	6715	75.9 (73.3-78.4)	81.6 (78.3-84.5)	0.89 (0.73-1.09)	.25
1 y	328 519	360 206	68 463	74 356	92.7 (92.1-93.2)	92.4 (91.7-93.0)	1.05 (0.96-1.14)	.27
2 y	61 616	68 555	28 710	30 975	91.1 (90.2-91.8)	92.2 (91.4-93.0)	0.94 (0.83-1.07)	.35
≥3 y	44 888	52 105	15 214	17 122	88.0 (86.9-89.0)	89.1 (87.9-90.3)	1.06 (0.93-1.20)	.42

Abbreviations: AOR, adjusted odds ratio; BI-RADS, Breast Imaging Reporting and Data System; CAD, computer-aided detection; exam, examination; HT, hormonal therapy; NE, could not be estimated.

[a] Odds ratio for CAD vs No CAD adjusted for Breast Cancer Surveillance Consortium registry, age group, race/ethnicity, time since prior mammogram and calendar year of the examination using mixed-effects model with random effect for examination reader.

[b] The 95% CIs for sensitivity and specificity are given as percentages.

with vs without CAD in the subset of radiologists who interpreted both with and without CAD assistance (OR, 0.53 [95% CI, 0.29-0.97]).

Discussion

We found no evidence that CAD applied to digital mammography in US community practice improves screening mammography performance on any performance measure or in any subgroup of women. In fact, mammography sensitivity was decreased in the subset of radiologists who interpreted mammograms with and without CAD. This study builds on prior studies[18,19] by demonstrating that radiologists' early learning curve and patient characteristics do not account for the lack of benefit from CAD.

Whether CAD provides added value to women undergoing screening mammography is a topic of strong debate.[27-36] The lack of consensus may be partly explained by wide variation in CAD use and inherent biases in the methods used to study the impact of CAD on screening mammography. Early studies[37,38] supporting the efficacy of CAD were laboratory based and measured the ability of CAD programs to mark cancers on selected mammograms. The reported "high sensitivities" of CAD from these studies did not translate to higher cancer detection in clinical practice. In clinical practice, most positive marks by CAD must be reviewed and discounted by a radiologist to avoid unacceptably high rates of false-positive results and unnecessary biopsies, and to practice within acceptable performance parameters recommended by the American College of Radiology.[24] The most optimistic view of CAD is that it improves mammography sensitivity by 20%.[8,28,30,32] If this were true, cancer detection rates of 4 to 5 per 1000 without CAD would increase to 5 to 6 per 1000 with CAD. In other words, for every 1000 women whose screening mammograms were interpreted with CAD, 1 cancer would be identified that was missed by the unassisted radiologist interpretation. To achieve that single true-positive CAD marking in 1000 women, CAD would render 2000 to 4000 false-positive marks. Thus, under this scenario, a radiologist would need to recommend diagnostic evaluation for the single CAD mark of the otherwise missed cancer, while discounting thousands of false-positive CAD marks.

Consistent with reports of a prior BCSC cohort study[18] and Surveillance, Epidemiology, and End Results-Medicare data[2] which primarily evaluated film-screen mammography, we found higher rates of DCIS lesions detected with CAD on digital mammography, but no differences in *sensitivity* for cancer (whether for DCIS or invasive) and no differences in rates of invasive cancers detected. A meta-analysis[39] in 2008 of 10 studies of CAD applied to screening mammography concluded that CAD significantly increased recall rates with no significant improvement in cancer detection rates compared with readings without CAD. The largest recent reader study of digital mammography obtained during the Digital Mammography Imaging Screening Trial (DMIST)[5] found no impact of CAD on radiologist interpretations of mammograms. In that report,[5] the authors concluded that radiologists overall were not in-

fluenced by CAD markings and CAD had no impact, either beneficial or detrimental, on mammography interpretations.

Our study had sufficiently large numbers to compare interpretations of mammograms read by radiologists who practiced at some sites with CAD and at other sites without CAD. We are concerned that, in these comparisons, sensitivity was lower in CAD-assisted mammograms. Prior reports have confirmed that not all cancers are marked by CAD and that cancers are overlooked more often if CAD fails to mark a visible lesion. In a large reader study, Taplin et al[7] reported that visible, noncalcified lesions that went unmarked by CAD were significantly less likely to be assessed as abnormal by radiologists. However, our finding of lower sensitivity with CAD was in a subgroup analysis and should be interpreted with caution.

Given the observational methods of our study, we could not compare mammography performance among women who had their mammograms interpreted both with and without CAD. It is possible that CAD was used preferentially in women whose mammograms were more challenging. However, given the large sample size we were able to control for multiple key factors known to influence mammography performance, including patient age, breast density, menopausal status, and time since last mammogram. We also were not able to control for radiologist characteristics, such as experience, and thus compared performance with and without CAD in the same radiologists, to address across-radiologist variability.

Our study found no beneficial impact of CAD on mammography interpretation. However, CAD may offer advantages beyond interpretation, such as improved workflow or reduced search time for faint calcifications. Future research on potential applications of CAD may emphasize the contribution of CAD to guide decision-making about treatment of a radiologist-detected lesion, with the worthy goals of reducing unnecessary biopsy of a mammography lesion with specific benign features or supporting biopsy of a lesion with specific malignant features. Finally, CAD might improve mammography performance when appropriate training is provided on how to use it to enhance performance. Nevertheless, given that the evidence of the current application of CAD in community practice does not show an improvement in diagnostic accuracy, we question the policy of continuing to charge for a technology that provides no established benefits to women.

Gross et al[40] reported that the costs of breast cancer screening exceed $1 billion annually in the Medicare fee-for-service population. Consistent with our findings, they found wide variation in CAD use and very limited effectiveness and encouraged attention to more appropriate and evidence-based application of new technologies in breast cancer screening programs. Despite its overall lack of improvement on interpretive performance, CAD has become routine practice in mammography interpretations in the United States. Seventeen years have passed since the FDA approved the use of CAD in screening mammography, and 14 years have passed since Congress mandated Medicare coverage of CAD. Ten years ago, the Institute of Medicine stated that more information on CAD applied to mammography was needed before making conclusions about its effect on interpretation.[41] The US FDA estimates that 38.8 million mammograms are performed each year in the United States. In the BCSC database, 80% of mammograms are performed for screening and by 2012, 83% of screening mammograms in the BCSC were digital examinations interpreted with CAD. Current CMS reimbursement for CAD is roughly $7 per examination, and many private insurers pay more than $20 per examination for CAD, translating to over $400 million per year in current US health care expenditures, with no added value and in some cases decreased performance.

Conclusions

In the era of Choosing Wisely and clear commitments to support technology that brings added value to the patient experience, while aggressively reducing waste and containing costs,[42] CAD is a technology that does not seem to warrant added compensation beyond coverage of the mammographic examination. The results of our comprehensive study lend no support for continued reimbursement for CAD as a method to increase mammography performance or improve patient outcomes.

ARTICLE INFORMATION

Group Information: Members of the Breast Cancer Surveillance Consortium are listed at the end of this article.

Accepted for Publication: August 7, 2015.

Published Online: September 28, 2015. doi:10.1001/jamainternmed.2015.5231.

Author Contributions: Drs Lehman and Miglioretti had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Lehman, Buist, Kerlikowske, Miglioretti.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Lehman, Wellman, Buist.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Wellman, Miglioretti.
Obtained funding: Lehman, Buist, Kerlikowske, Tosteson, Miglioretti.
Administrative, technical, or material support: Lehman, Buist, Kerlikowske, Tosteson.
Study supervision: Lehman, Buist, Miglioretti.

Conflict of Interest Disclosures: Dr Lehman has received grant support from General Electric (GE) Healthcare and is a member of the Comparative Effectiveness Research Advisory Board for GE Healthcare. No other disclosures are reported.

Funding/Support: This original research was supported by the National Cancer Institute (NCI) (P01CA154292). Data collection was further supported by the NCI-funded Breast Cancer Surveillance Consortium (BCSC) (HHSN261201100031C, P01CA154292), and PROSPR U54CA163303. The collection of cancer and vital status data used in this study was supported in part by several state public health departments and cancer registries throughout the United States. For a full description of these sources, please see the NCI website, BCSC Cancer Registry Acknowledgement: http://breastscreening.cancer.gov/work/acknowledgement.html.

Role of the Funder/Sponsor: The NCI had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Breast Cancer Surveillance Consortium Group Information: Linn Abraham, MS, Group Health Cooperative; Rob Arao, MS, Group Health Cooperative; Andrew Avins, MD, Kaiser Permanente Division of Research; Steve Balch, MS, MBA, Group Health Cooperative; Thad Benefield, MS, University of North Carolina, Chapel Hill; Erin Aiello Bowles, MPH, Group Health Cooperative; Mark Bowman, University of Vermont; Susan

Brandzel, MPH, Group Health Cooperative; Diana Buist, PhD, MPH, Group Health Cooperative; David Burian, BA, University of California, San Francisco; Elyse Chiapello, BASc, University of California, San Francisco; Rachael Chicoine, BS, University of Vermont; Firas Dabbous, MS, University of Illinois at Chicago; Tammy Dodd, Group Health Cooperative; Therese Dolecek, PhD, MS, University of Illinois at Chicago; Scottie Eliassen, MS, Dartmouth College; Kevin Filocamo, Group Health Cooperative; Pete Frawley, Group Health Cooperative; Hongyuan Gao, MS, Group Health Cooperative; Charlotte Gard, PhD, MS, Consultant New Mexico State University; Berta Geller, PhD, University of Vermont; Martha Goodrich, MS, Dartmouth College; Mikael Anne Greenwood-Hickman, MPH, University of North Carolina, Chapel Hill; Cindy Groseclose, University of Vermont; Louise Henderson, PhD, MSPH, University of North Carolina, Chapel Hill; Deirdre Hill, PhD, University of New Mexico; Michael Hofmann, MS, University of California, San Francisco; Rebecca Hubbard, PhD, University of Pennsylvania; Erika Holden, Group Health Cooperative; Tiffany Hoots, University of North Carolina, Chapel Hill; Kathleen Howe, AA, University of Vermont; Laura Ichikawa, MS, Group Health Cooperative; Doug Kane, MS, Group Health Cooperative; Karla Kerlikowske, MD, University of California, San Francisco; Jenna Khan, MPH, University of Illinois at Chicago; Gabe Knop, University of North Carolina, Chapel Hill; Suzanne Kolb, MPH University of Washington; Casey Luce, MSPH, Group Health Cooperative; Lin Ma, MS University of California, San Francisco; Terry Macarol, Advocate Health Care; John Mace, PhD, University of Vermont; Jennifer Maeser, MS, University of Washington; Kathy Malvin, BA, University of California, San Francisco; Katie Marsh, MPH, University of North Carolina, Chapel Hill; Diana Miglioretti, PhD, University of California, Davis; Anne Marie Murphy, PhD, Metropolitan Chicago Breast Cancer Task Force; Ellen O'Meara, PhD, Group Health Cooperative; Tracy Onega, PhD, MA, MS, Dartmouth College; Tiffany Pelkey, BA, University of Vermont; Dusty Quick, University of Vermont; Garth Rauscher, PhD, University of Illinois at Chicago; KatieRose Richmire, Group Health Cooperative; Scott Savioli, MA, Dartmouth College; Deborah Seger, Group Health Cooperative; Jennette Sison, MPH, University of California, San Francisco; Brian Sprague, PhD, University of Vermont; Wm. Thomas Summerfelt, PhD, Advocate Health Care; Katherine Tossas-Milligan, MS, University of Illinois at Chicago; Anna Tosteson, ScD, Dartmouth College; Rod Walker, MS, Group Health Cooperative; Julie Weiss, MS, Dartmouth College; Rob Wellman, MS, Group Health Cooperative; Karen Wernli, PhD, Group Health Cooperative; Heidi Whiting, MS Group Health Cooperative; Bonnie Yankaskas, PhD, University of North Carolina, Chapel Hill; Weiwei Zhu, MS, Group Health Cooperative.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the NCI or the National Institutes of Health.

Additional Contributions: We thank the participating women, mammography facilities, and radiologists for the data they have provided for this study. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at http://breastscreening.cancer.gov.

REFERENCES

1. US Food and Drug Administration. Summary of Safety and Effectiveness Data. R2 Technologies (P970058) 1998. P970058. http://www.fda.gov/ohrms/dockets/98fr/123098b.txt.

2. Fenton JJ, Xing G, Elmore JG, et al. Short-term outcomes of screening mammography using computer-aided detection: a population-based study of Medicare enrollees. *Ann Intern Med*. 2013;158(8):580-587.

3. Rao VM, Levin DC, Parker L, Cavanaugh B, Frangos AJ, Sunshine JH. How widely is computer-aided detection used in screening and diagnostic mammography? *J Am Coll Radiol*. 2010;7(10):802-805.

4. Gilbert FJ, Astley SM, McGee MA, et al. Single reading with computer-aided detection and double reading of screening mammograms in the United Kingdom National Breast Screening Program. *Radiology*. 2006;241(1):47-53.

5. Cole EB, Zhang Z, Marques HS, Edward Hendrick R, Yaffe MJ, Pisano ED. Impact of computer-aided detection systems on radiologist accuracy with digital mammography. *AJR Am J Roentgenol*. 2014;203(4):909-916.

6. Ciatto S, Del Turco MR, Risso G, et al. Comparison of standard reading and computer aided detection (CAD) on a national proficiency test of screening mammography. *Eur J Radiol*. 2003;45(2):135-138.

7. Taplin SH, Rutter CM, Lehman CD. Testing the effect of computer-assisted detection on interpretive performance in screening mammography. *AJR Am J Roentgenol*. 2006;187(6):1475-1482.

8. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*. 2001;220(3):781-786.

9. Birdwell RL, Bandodkar P, Ikeda DM. Computer-aided detection with screening mammography in a university hospital setting. *Radiology*. 2005;236(2):451-457.

10. Ko JM, Nicholas MJ, Mendel JB, Slanetz PJ. Prospective assessment of computer-aided detection in interpretation of screening mammography. *AJR Am J Roentgenol*. 2006;187(6):1483-1491.

11. Morton MJ, Whaley DH, Brandt KR, Amrami KK. Screening mammograms: interpretation with computer-aided detection—prospective evaluation. *Radiology*. 2006;239(2):375-383.

12. Georgian-Smith D, Moore RH, Halpern E, et al. Blinded comparison of computer-aided detection with human second reading in screening mammography. *AJR Am J Roentgenol*. 2007;189(5):1135-1141.

13. Gur D, Sumkin JH, Rockette HE, et al. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. *J Natl Cancer Inst*. 2004;96(3):185-190.

14. Cupples TE, Cunningham JE, Reynolds JC. Impact of computer-aided detection in a regional screening mammography program. *AJR Am J Roentgenol*. 2005;185(4):944-950.

15. Romero C, Varela C, Muñoz E, Almenar A, Pinto JM, Botella M. Impact on breast cancer diagnosis in a multidisciplinary unit after the incorporation of mammography digitalization and computer-aided detection systems. *AJR Am J Roentgenol*. 2011;197(6):1492-1497.

16. Gromet M. Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. *AJR Am J Roentgenol*. 2008;190(4):854-859.

17. Gilbert FJ, Astley SM, Gillan MG, et al; CADET II Group. Single reading with computer-aided detection for screening mammography. *N Engl J Med*. 2008;359(16):1675-1684.

18. Fenton JJ, Taplin SH, Carney PA, et al. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med*. 2007;356(14):1399-1409.

19. Fenton JJ, Abraham L, Taplin SH, et al; Breast Cancer Surveillance Consortium. Effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst*. 2011;103(15):1152-1161.

20. National Cancer Institute Breast Cancer Surveillance Consortium. http://breastscreening.cancer.gov/.

21. Yankaskas BC, Taplin SH, Ichikawa L, et al. Association between mammography timing and measures of screening performance in the United States. *Radiology*. 2005;234(2):363-373.

22. D'Orsi CJ, Sickles EA, Mendelson EB, et al. *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology; 2013.

23. Sickles EA, D'Orsi CJ, Bassett LW, et al. ACR BI-RADS Mammography. In: *ACR BI-RADS, Breast Imaging Reporting and Data System*. 5th ed. Reston, VA: American College of Radiology; 2013.

24. Sickles EA, D'Orsi CJ. ACR BI-RADS Follow-up and Outcome Monitoring. In: *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology; 2013.

25. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20(19):2865-2884.

26. Pepe MS, ed. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. New York, NY: Oxford University Press; 2003.

27. Elmore JG, Carney PA. Computer-aided detection of breast cancer: has promise outstripped performance? *J Natl Cancer Inst*. 2004;96(3):162-163.

28. Feig SA, Sickles EA, Evans WP, Linver MN. Re: Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. *J Natl Cancer Inst*. 2004;96(16):1260-1261.

29. Hall FM. Breast imaging and computer-aided detection. *N Engl J Med*. 2007;356(14):1464-1466.

30. Feig SA, Birdwell RL, Linver MN. Computer-aided screening mammography. *N Engl J Med*. 2007;357(1):84.

31. Gur D. Computer-aided screening mammography. *N Engl J Med*. 2007;357(1):83-84.

32. Birdwell RL. The preponderance of evidence supports computer-aided detection for screening mammography. *Radiology*. 2009;253(1):9-16.

33. Philpotts LE. Can computer-aided detection be detrimental to mammographic interpretation? *Radiology*. 2009;253(1):17-22.

34. Berry DA. Computer-assisted detection and screening mammography: where's the beef? *J Natl Cancer Inst*. 2011;103(15):1139-1141.

35. Nishikawa RM, Giger ML, Jiang Y, Metz CE. Re: effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst*. 2012;104(1):77-78.

36. Levman J. Re: effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst*. 2012;104(1):77-78.

37. Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA. Computer-aided detection of clustered microcalcifications: an improved method for grouping detected signals. *Med Phys*. 1993;20(6):1661-1666.

38. Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Giger ML, Doi K. Improving breast cancer diagnosis with computer-aided diagnosis. *Acad Radiol*. 1999;6(1):22-33.

39. Taylor P, Potts HW. Computer aids and human second reading as interventions in screening mammography: two systematic reviews to compare effects on cancer detection and recall rate. *Eur J Cancer*. 2008;44(6):798-807.

40. Gross CP, Long JB, Ross JS, et al. The cost of breast cancer screening in the Medicare population. *JAMA Intern Med*. 2013;173(3):220-226.

41. Nass S, Ball J, eds. *Improving Breast Imaging Quality Standards. Institute of Medicine and National Research Council of the National Academies*. Washington, DC: The National Academies Press; 2005.

42. American Board of Internal Medicine Foundation. Choosing Wisely campaign. http://www.abimfoundation.org/Initiatives/Choosing-Wisely.aspx. Accessed September 1, 2015,

Invited Commentary

Is It Time to Stop Paying for Computer-Aided Mammography?

Joshua J. Fenton, MD, MPH

Computer-aided detection (CAD) is a technology designed to address the problem of screening mammography's imperfect sensitivity. Now used on over 90% of US mammograms, CAD essentially acts like an automated second reader by marking potentially suspicious spots for radiologists to review before making final recommendations. Early studies suggested that CAD could increase cancer detection rates by 20%.[1] But subsequent research suggested little, if any, impact of CAD on cancer detection and raised concerns that CAD may increase recall and biopsy rates.[2,3]

Related article page 1828

However, most clinical studies to date have assessed CAD impacts when used with film mammograms. Digital mammography has now largely supplanted film mammography in the United States. When used in the context of digital mammography, does CAD yield net benefits to women? A study by Lehman et al[4] in this issue of *JAMA Internal Medicine* addresses this important question.

In an observational study of 323 973 women undergoing digital screening mammography in diverse US practices, Lehman et al[4] found that CAD use was not associated with any improvement in sensitivity, specificity, positive predictive value, cancer detection rates, or other proximal screening outcomes. Indeed, among radiologists who interpreted digital mammograms with and without CAD, sensitivity was worse with CAD, contrary to CAD's design.

While earlier evaluations suggested that community radiologists often overreacted to CAD output, leading to higher rates of diagnostic investigation,[2,5] Lehman et al[4] found little, if any, impact of CAD in modern digital mammography practice. It is possible that, with years of CAD use, many radiologists have learned that the yield of CAD is minimal so they now largely ignore CAD output. It is also conceivable that improvements in digital mammography technology have swamped any incremental impacts of CAD on interpretation that may have been previously detectable. This observational study may be confounded by unmeasured radiologist or mammography facility factors, although earlier research adjusting for these factors also found no benefits of CAD.[2] Like all subgroup analyses, analyses of outcomes among subsets of radiologists must be interpreted cautiously. Despite these limitations, this study[4] is another large-sample, real-world evaluation of CAD's interpretive outcomes suggesting that CAD yields no clinically significant benefits in typical mammography practice.

The field of implementation science should take interest in interventions like CAD that are widely adopted in advance of strong evidence of effectiveness. What made CAD so alluring to patients, practitioners, or both, and why were payers willing (at least initially) to finance CAD? How is it that CAD is applied on 90% of US mammograms when it yields no clear benefits to patients?

The first essential step for broad CAD adoption was US Food and Drug Administration (FDA) approval in 1998. Because CAD is a device rather than a drug, the evidence bar for FDA approval was comparatively low. Its approval was based on small studies of CAD's "safety" and "effectiveness." Effectiveness, for example, was demonstrated by studies in which radiologists read sets of mammograms with enriched breast cancer prevalence, suggesting that CAD could prompt increased cancer detection. In addition, Congressional members, lobbied by industry, pressured the FDA to approve CAD.[6]

Even so, CAD was still a longshot. Use of CAD required film mammograms to be fed into machines to digitize images for computer analysis, and CAD output had to be viewed on dedicated devices separate from actual mammograms. Mammography was already a loss-leader for many radiology practices, yet CAD added unreimbursed technician and radiologist effort. Without reimbursement, few mammography facilities could justify the capital costs for CAD installation. At the time, establishing reimbursement for new preventive services, such as CAD, required Congressional amendment of the Medicare statute. While Congress had previously added Medicare benefits for preventive services, such as prostate cancer screening, these efforts required strong Congressional sponsorship and auspicious political winds.[7]

CAD ended up having both. Representing Silicon Valley (home of CAD's leading manufacturer), California Congress-