Radiology

ORIGINAL RESEARCH • BREAST IMAGING

Breast Cancer Detection with Standalone AI versus Radiologist Interpretation of Unilateral Surveillance Mammography after Mastectomy

Su Min Ha, MD, PhD[1,2,3] • *Janie M. Lee, MD, MSc*[4] • *Myoung-jin Jang, PhD*[5] • *Hong-Kyu Kim, MD, PhD*[6] • *Jung Min Chang, MD, PhD*[1,2]

Author affiliations, funding, and conflicts of interest are listed at the end of this article.

See also the editorial by Philpotts in this issue.

Radiology 2025; 315(1):e242955 • https://doi.org/10.1148/radiol.242955 • Content codes: BR AI

Background: Limited data are available regarding the accuracy of artificial intelligence (AI) algorithms trained on bilateral mammograms for second breast cancer surveillance in patients with a personal history of breast cancer treated with unilateral mastectomy.

Purpose: To compare the performance of standalone AI for second breast cancer surveillance on unilateral mammograms with that of radiologists reading mammograms without AI assistance.

Materials and Methods: In this retrospective institutional database study, patients who were diagnosed with breast cancer between January 2001 and December 2018 and underwent postmastectomy surveillance mammography from January 2011 to March 2023 were included. Radiologists' mammogram interpretations without AI assistance were collected from these records and compared with AI interpretations of the same mammograms. The reference standards were histologic examination and 1-year follow-up data. The cancer detection rate per 1000 screening examinations, sensitivity, and specificity of standalone AI and the radiologists' interpretations without AI were compared using the McNemar test.

Results: Among the 4184 asymptomatic female patients (mean age, 52 years), 111 (2.7%) had contralateral second breast cancer. The cancer detection rate (17.4 per 1000 examinations [73 of 4184]; 95% CI: 13.7, 21.9) and sensitivity (65.8% [73 of 111]; 95% CI: 56.2, 74.5) were greater for standalone AI than for radiologists (14.6 per 1000 examinations [61 of 4184]; 95% CI: 11.2, 18.7; $P = .01$; 55.0% [61 of 111]; 95% CI: 45.2, 64.4; $P = .01$). The specificity was lower for standalone AI than for radiologists (91.5% [3725 of 4073]; 95% CI: 90.6, 92.3 vs 98.1% [3996 of 4073]; 95% CI: 97.6, 98.5; $P < .001$). AI detected 16 of 50 (32%) cancers missed by radiologists; however, 34 of 111 (30.6%) breast cancers were missed by both radiologists and AI.

Conclusion: Standalone AI for surveillance mammography showed higher sensitivity with lower specificity for contralateral breast cancer detection in patients treated with unilateral mastectomy than radiologists without AI assistance.

Supplemental material is available for this article.

Breast cancer is the most common malignancy in women worldwide (1). The number of breast cancer survivors is increasing due to advances in both breast cancer screening and cancer treatment (2). Patients with a personal history of breast cancer (PHBC) are at risk for developing a second breast cancer, including an ipsilateral recurrence or a new primary cancer in the treated or contralateral breast. Therefore, surveillance mammography is recommended for patients with a PHBC, with observational studies and meta-analyses showing reduced breast cancer mortality and improved quality of life with surveillance (3). However, surveillance mammography in patients with a PHBC has lower sensitivity (69.9% vs 86.9%) and higher interval cancer rates (3.4 vs 0.8 per 1000 examinations) than general screening mammography (95% of examinations from individuals without a history of breast cancer) (4,5). Some interval breast cancers at mammography result from radiologists' failure to detect abnormalities (6) along with the presence of dense breast tissue (3). These cancers are associated with poor prognostic markers, including larger tumor size, negative hormone receptor status, and positive lymph node status (5). Hence, there is a need for more precise and effective tools for screening and surveillance mammography to improve patient outcomes.

In recent years, various mammographic artificial intelligence (AI) systems have been developed and validated (7–9). Indeed, several retrospective studies have shown promising results with AI systems for breast cancer detection from screening mammography (10–16). AI systems can be used for concurrent clinical decision support to improve radiologists' overall accuracy in reading mammograms (7–9). In the recent Mammography Screening with Artificial Intelligence or, MASAI, trial, standalone AI systems served as independent readers in a double-reading setting (17). Although AI has shown promise in improving screening mammography interpretation, evidence regarding the interpretation of surveillance mammograms in patients with a PHBC is sparse, with only one study to date (18) that investigated AI for reading mammography in patients treated with conservation surgery. In addition, although AI algorithms are typically trained on bilateral mammography examinations, no study has investigated AI accuracy for second breast cancer surveillance in patients treated with unilateral mastectomy. Given the growing population of patients with a PHBC undergoing surveillance mammography, investigating how this unique group can benefit from AI is highly important.

Abbreviations

AI = artificial intelligence, PHBC = personal history of breast cancer

Summary

Standalone artificial intelligence (AI) could detect contralateral breast cancers in patients treated with unilateral mastectomy, showing higher cancer detection rates and sensitivity than radiologists without AI assistance.

Key Results

- In a retrospective study of 4184 female patients treated with unilateral mastectomy, standalone artificial intelligence (AI) showed a higher cancer detection rate (17.4 vs 14.6 per 1000 examinations; $P = .01$) and sensitivity (65.8% vs 55.0%; $P = .01$) but lower specificity (91.5% vs 98.1%; $P < .001$) than radiologists detecting contralateral breast cancer on mammograms.
- AI identified 16 of 50 (32%) cancers missed by radiologists; however, 34 of 111 (30.6%) total breast cancers were missed by both radiologists and AI.

The purpose of this study was to compare the performance of a standalone AI system for second breast cancer surveillance in patients treated with unilateral mastectomy with that of radiologists reading mammograms without AI assistance.

Materials and Methods

Study Sample

This retrospective study was approved by the institutional review board of the Seoul National University Hospital, and the requirement for informed consent was waived. A retrospective search of the Seoul National University Hospital database was performed to identify consecutive patients with primary unilateral breast cancer diagnosis of ductal carcinoma in situ or a stage I–III invasive cancer, according to the American Joint Committee on Cancer staging manual, eighth edition (19), who underwent definitive treatment with mastectomy (but not bilateral mastectomy) between January 2001 and December 2018. Annual mammography was offered along with supplemental annual or semiannual whole-breast US or MRI for the first 5 years after breast cancer treatment and annually thereafter for patients with PHBC (20). Additional inclusion criteria were receipt of postoperative surveillance with unilateral digital mammography from January 2011 to March 2023. Patients with no reference standard for breast cancer status and patients with a history of benign mass excision or reduction mammoplasty in the contralateral breast that could interfere with accurate lesion assessment were excluded. Since surveillance mammography was performed for each patient more than once during the study period, the earliest available surveillance mammogram was used for analysis in patients without a second cancer diagnosis; otherwise, the mammogram immediately preceding the second breast cancer diagnosis at the institution was used (Fig 1).

Imaging Protocol

All imaging data were obtained as part of routine clinical practice and stored in the institutional picture archiving and communication system. The two-dimensional mammographic imaging data were acquired using a full-field mammography unit (Selenia Dimensions, Hologic; Senographe 2000D, GE HealthCare). Unilateral digital mammography included two views (mediolateral oblique and craniocaudal) of the breast. At the time of unilateral mammogram acquisition, AI was not applied; thus, the radiologists (2–30 years of experience in breast imaging, including S.M.H. and J.M.C., with 10 and 17 years of experience, respectively) who interpreted the mammograms did not have AI results available.

Image Analysis

The radiologists determined breast density, mammographic findings, and final assessment category during routine clinical care using the American College of Radiology Breast Imaging Reporting and Data System (21,22). Mammographic examinations with heterogeneously or extremely dense breasts were classified as having dense tissue (23). Previous mammographic, US, or MRI examinations were available for comparison at the time of mammogram reading. When mammography was scheduled on the same day as supplemental US, the radiologists reported the mammographic findings first, independently of the US evaluation; only afterward did they report the US findings and

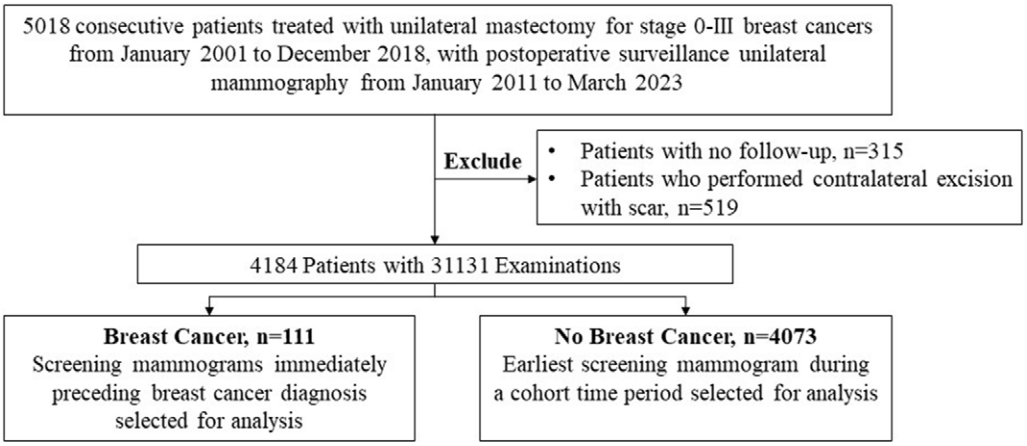

Figure 1: Flowchart of the final study sample. Since screening mammography was performed more than once during a cohort time period (January 2011 to March 2023), the earliest screening mammogram was used for analysis in patients without a second cancer diagnosis, and the mammogram immediately preceding the second breast cancer diagnosis was used for analysis in patients with a diagnosis and treatment of second breast cancer.

Table 1: Study Population Characteristics

Variable	All Patients (*n* = 4184)	Patients with Contralateral Breast Cancer (*n* = 111)	Patients without Contralateral Breast Cancer (*n* = 4073)
Mean age at first breast cancer diagnosis (y)	52.0 ± 10.9 (21–86)	51.1 ± 10.3 (28–80)	52.0 ± 10.9 (21–86)
Median age at first breast cancer diagnosis (y)	51 (44–59)	50 (43–58)	51 (44–59)
Mammographic breast density*			
Almost entirely fatty, BI-RADS A	111 (2.7)	2 (1.8)	109 (2.7)
Scattered fibroglandular density, BI-RADS B	1289 (30.8)	31 (27.9)	1258 (30.9)
Heterogeneously dense, BI-RADS C	2397 (57.3)	67 (60.4)	2330 (57.2)
Extremely dense, BI-RADS D	387 (9.2)	11 (9.9)	376 (9.2)
Mammographic breast density (dichotomous)*			
Nondense, BI-RADS A and B	1400 (33.5)	33 (29.7)	1367 (33.6)
Dense, BI-RADS C and D	2784 (66.5)	78 (70.3)	2706 (66.4)
BRCA1/2 genetic testing			
Untested	3854 (92.1)	83 (74.8)	3771 (92.6)
Tested			
Negative result	293 (7.0)	21 (18.9)	272 (6.7)
Positive result	37 (0.8)	7 (6.3)	30 (0.7)

Note.—Data are presented as means ± SDs with ranges in parentheses, as medians with IQRs in parentheses, or as numbers of patients with percentages in parentheses. BI-RADS = Breast Imaging Reporting and Data System.

* Breast density categories are according to the American College of Radiology BI-RADS.

make a final assessment with recommendations on the basis of both modalities and guidance (Appendix S1).

AI Software Application

A commercially available AI software (Lunit INSIGHT for Mammography, version 1.1.7.1; Lunit) that was validated through a multinational study was used (7). The model training was described in Appendix S1 (24). Since the AI system was used for standard four-view mammography at the study institution but not routinely applied for unilateral mammography, AI was retrospectively applied to all postmastectomy surveillance mammograms included in this study. The AI algorithm provided prediction scores for each view independently using an image-level classifier. Analytic results were presented as a separate gray-scale image, with overall per-breast abnormality scores for each mediolateral oblique and craniocaudal view ranging from 0 to 100 and a heat map of abnormal areas that could indicate malignancy.

Data Collection and Reference Standards

The clinical digital mammography reports were extracted from medical records, and AI abnormality scores were collected. For radiologist interpretations, lesions with recalled findings (excluding technical recalls) from initial mammography reports (Breast Imaging Reporting and Data System categories of 0, 3, 4, or 5) were considered positive. For the standalone AI, lesions with abnormality scores of 10 or greater were considered positive. A true-positive result was defined as the heat map marking a site corresponding to a breast cancer diagnosis, and a false-positive result was defined as a heat map annotation at a different site. For breasts with one lesion, the highest score from both views was used as the examination-level score. If more than one lesion was identified in a breast, the highest abnormality score across lesions was used as the final examination-level score. Two breast radiologists (S.M.H. and J.M.C., with 10 and 17 years of experience, respectively) retrospectively reviewed all mammography reports and AI scores in consensus, along with cancer status data of lesions identified at surveillance mammography, to assess whether the interpreting radiologists or AI software identified or missed a cancer lesion. Mammograms obtained after wire localization and preoperative MRI or US images were used as reference standards to determine the location of the cancer. Additional data were also collected from medical records, including age, breast density, delay between initial breast cancer surgery and surveillance mammography, and *BRCA* gene mutation status. Histologic examinations and at least 1 year of follow-up data were used as the reference standards. Patients diagnosed with ductal carcinoma in situ or invasive carcinoma were considered to have contralateral breast recurrences. For malignant lesions, pathologic tumor size, histologic type, TNM stage, estrogen receptor status, progesterone receptor status, and human epidermal growth factor receptor type 2 status were collected (Appendix S1). High-risk lesions, including atypical ductal hyperplasia, lobular cancer in situ, and complex sclerosing lesions, were considered benign. Lesions that were diagnosed as benign at biopsy or surgery or those that did not change at follow-up for 1 year or more were also considered benign. Finally, the mode of detection (mammography, breast US, or MRI) for contralateral breast cancers was recorded for investigation.

Statistical Analysis

All statistical analyses were conducted using individual mammography examinations as the unit of analysis. The cancer detection rate, sensitivity, specificity, recall rate, positive predictive value, and negative predictive value were calculated for the entire study sample (Appendix S1). The 95% CIs were estimated as the Clopper-Pearson exact CIs for proportions. Sensitivity and specificity were compared using the McNemar test. A generalized estimating equation was used to compare the positive predictive value and the negative predictive value. Performance comparisons between standalone AI and radiologist interpretation were performed for mammographic density subgroups (dense vs nondense) using post

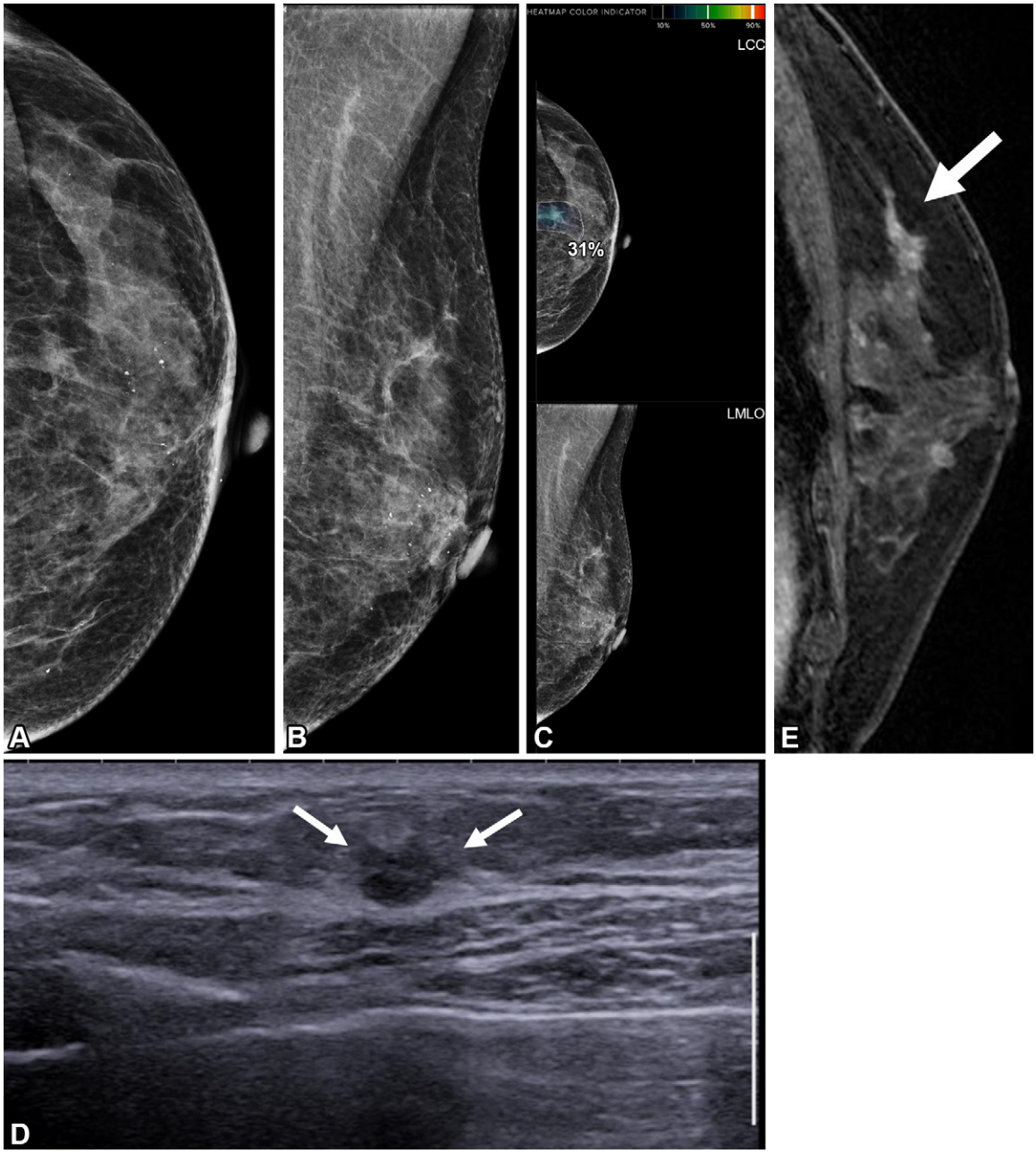

Figure 2: Images in a 65-year-old patient with contralateral second breast cancer 6.3 years after right mastectomy. **(A)** Left craniocaudal and **(B)** mediolateral oblique mammograms assessed as benign. **(C)** The artificial intelligence algorithm outlined a lesion with a score of 31 on the craniocaudal view. **(D)** Supplemental axial US scan, which was obtained on the same day as the mammogram, shows an irregular hypoechoic mass (arrows) in the left upper central breast. **(E)** Preoperative sagittal contrast-enhanced T1-weighted MRI scan shows the corresponding enhancing mass (arrow) in the left upper center. The mass was proven to be ductal carcinoma in situ. LCC = left craniocaudal, LMLO = left mediolateral oblique.

hoc analyses. Statistical analysis was performed using SAS, version 9.4 (SAS Institute). Post hoc power analyses were conducted using PASS 2022 (NCSS Statistical Software) (Appendix S1). A two-sided $P < .05$ indicated a statistically significant difference.

Results

Patient Characteristics

Among the 5018 eligible female patients with breast cancer treated with unilateral mastectomy, 834 patients without reference standard data ($n = 315$) or with prior surgery in the contralateral breast that could interfere with lesion assessment ($n = 519$) were excluded. Ultimately, 4184 asymptomatic consecutive patients (mean age at imaging, 52.0 years ± 10.9 [SD]; range, 21–86 years) were included in the analysis (Fig 1). Mammographic breast density was classified as nondense in 1400 (33.5%) patients and dense in 2784 (66.5%) patients. Additional information is provided in Table 1.

Diagnostic Performance of Standalone AI versus Radiologists

The total of 4184 patients included 111 patients with breast cancer and 4073 patients with no breast cancer diagnosis over a median follow-up time of 6 years (IQR, 3.4–9.7 years). The cancer detection rate was 17.4 per 1000 examinations (73 of 4184; 95% CI: 13.7, 21.9) for standalone AI versus 14.6 per 1000 examinations

Table 2: Comparison of Diagnostic Performance between Radiologists and Standalone AI

Parameter	Radiologists	Standalone AI	*P* Value
Total			
Cancer detection rate per 1000 examinations	14.6 (61/4184) [11.2, 18.7]	17.4 (73/4184) [13.7, 21.9]	.01
Sensitivity (%)	55.0 (61/111) [45.2, 64.4]	65.8 (73/111) [56.2, 74.5]	.01
Specificity (%)	98.1 (3996/4073) [97.6, 98.5]	91.5 (3725/4073) [90.6, 92.3]	<.001
Recall rate (%)	3.3 (138/4184) [2.8, 3.9]	10.1 (421/4184) [9.2, 11.0]	<.001
PPV (%)	44.2 (61/138) [35.8, 52.9]	17.3 (73/421) [13.8, 21.3]	<.001
NPV (%)	98.8 (3996/4046) [98.4, 99.1]	99.0 (3725/3763) [98.6, 99.3]	.047
Patients with dense breast			
Cancer detection rate per 1000 examinations	15.4 (43/2784) [11.2, 20.7]	19.0 (53/2784) [14.3, 24.8]	.01
Sensitivity (%)	55 (43/78) [43.4, 66.4]	68 (53/78) [56.4, 78.1]	.01
Specificity (%)	97.8 (2646/2706) [97.2, 98.3]	90.3 (2443/2706) [89.1, 91.4]	<.001
Recall rate (%)	3.7 (103/2784) [3.0, 4.5]	11.4 (316/2784) [10.2, 12.6]	<.001
PPV (%)	41.7 (43/103) [32.1, 51.9]	16.8 (53/316) [12.8, 21.4]	<.001
NPV (%)	98.7 (2646/2681) [98.2, 99.1]	99.0 (2443/2468) [98.5, 99.3]	.04
Patients with nondense breast			
Cancer detection rate per 1000 examinations	12.9 (18/1400) [7.6, 20.2]	14.3 (20/1400) [8.7, 22.0]	.69
Sensitivity (%)	55 (18/33) [36.4, 71.9]	61 (20/33) [42.1, 77.1]	.69
Specificity (%)	98.8 (1350/1367) [98.0, 99.3]	93.8 (1282/1367) [92.4, 95.0]	<.001
Recall rate (%)	2.5 (35/1400) [1.7, 3.5]	7.5 (105/1400) [6.2, 9.0]	<.001
PPV (%)	51 (18/35) [34.0, 68.6]	19.0 (20/105) [12.0, 27.9]	<.001
NPV (%)	98.9 (1350/1365) [98.2, 99.4]	99.0 (1282/1295) [98.3, 99.5]	.60

Note.—Data in parentheses are numerators and denominators; data in brackets are 95% CIs. *P* values were obtained from McNemar tests for sensitivity and specificity and from generalized estimating equations for positive predictive value (PPV) and negative predictive value (NPV). The 95% CIs for proportions were calculated using the Clopper-Pearson exact CI. AI = artificial intelligence.

(61 of 4184; 95% CI: 11.2, 18.7; $P = .01$) for radiologist interpretations alone. The sensitivity was greater with standalone AI (65.8% [73 of 111]; 95% CI: 56.2, 74.5) than with radiologist interpretation (55.0% [61 of 111]; 95% CI: 45.2, 64.4; $P = .01$). The specificity of standalone AI (91.5% [3725 of 4073]; 95% CI: 90.6, 92.3) was lower than that of radiologist interpretation (98.1% [3996 of 4073]; 95% CI: 97.6, 98.5; $P < .001$). The recall rate was higher with standalone AI (10.1% [421 of 4184]; 95% CI: 9.2, 11.0) than with radiologist interpretation (3.3% [138 of 4184]; 95% CI: 2.8, 3.9; $P < .001$). The positive predictive value (17.3% [73 of 421]; 95% CI: 13.8, 21.3) was lower with standalone AI than with radiologist interpretation (44.2% [61 of 138]; 95% CI: 35.8, 52.9; $P < .001$). Stratification according to density revealed similar performance in patients with dense breasts to that in the total patients. The number of second cancers among patients with nondense breasts was both relatively small and had wide CIs, with no evidence of differences in the cancer detection rate, sensitivity, and negative predictive value between radiologists and standalone AI (Table 2).

Characteristics of Contralateral Breast Cancers

Among the 111 patients with contralateral breast cancer, 62.2% ($n = 69$) had invasive cancer and 37.8% ($n = 42$) had ductal carcinoma in situ. The median invasive cancer size was 1.0 cm (IQR, 0.6–1.5 cm). Most cancers (97.3%; $n = 108$) were stage 0, I, or II, with 63.1% ($n = 70$) that were estrogen receptor or progesterone receptor positive and 93.7% ($n = 104$) that did not have lymph node metastasis.

Among the 111 detected cancers, 61 cancers were detected at surveillance mammography by the interpreting radiologists. For the other 50 cancers, the mammograms were interpreted as negative by radiologists; among these initially missed cancers, 43 were detected with other breast imaging modalities: supplemental US (0–10.2 months after mammography, $n = 36$), supplemental MRI (0–5.8 months after negative mammography, $n = 5$), and chest CT ($n = 2$). The remaining seven cancers were detected with palpable symptom presentation (4.8–6.3 months). Among all the cancers, 57 (51.4%) were detected by both the radiologists and AI, and 34 (30.6%) were detected by neither. A total of 20 (18.0%) cancers were detected by either AI only ($n = 16$; 14.4%) (Fig 2) or radiologists only ($n = 4$; 3.6%) (Table 3). Among the 16 cancers detected by AI but not by radiologists, 69% ($n = 11$) were invasive (median size, 1.5 cm; IQR, 0.6–3.4 cm), 63% ($n = 10$) were stage I or II, 69% ($n = 11$) were estrogen receptor or progesterone receptor positive, and 88% ($n = 14$) were without lymph node metastasis. The four invasive cancers identified by radiologists but missed by AI had lesion types of asymmetry ($n = 3$) or a mass ($n = 1$) at mammography (Table S1, Fig 3).

Discussion

Given that the use of artificial intelligence (AI) in breast cancer screening can enhance diagnostic performance and efficiency, implementing AI for evaluating surveillance mammography in patients treated with mastectomy may enable earlier detection of contralateral second breast cancer and facilitate timely surgery, ultimately improving survival outcomes. Our study demonstrated that a standalone AI system can achieve a higher cancer detection rate (17.4 vs 14.6 per 1000 examinations) and sensitivity (65.8% vs 55.0%) for contralateral breast cancers in patients treated with unilateral mastectomy than can radiologists' mammographic interpretation alone. Although AI identified 16

Table 3: Characteristics of Cancers

Characteristic	Total No. of Cancers (*n* = 111)	Both Radiologists and AI Detected Cancer (*n* = 57)	Radiologists Missed and AI Detected Cancer (*n* = 16)	Radiologists Detected and AI Missed Cancer (*n* = 4)	Both Radiologists and AI Missed Cancer (*n* = 34)
Histologic type					
Noninvasive	42 (37.8)	25 (44)	5 (31)	0	12 (35)
Invasive	69 (62.2)	32 (56)	11 (69)	4 (100)	22 (65)
Invasive ductal	60/69 (87)	26/32 (81)	9/11 (82)	4/4 (100)	21/22 (96)
Invasive lobular	7/69 (10)	5/32 (16)	1/11 (9)	0	1/22 (5)
Mixed ductal and lobular	1/69 (2)	0	1/11 (9)	0	0
Adenoid cystic carcinoma	1/69 (2)	1/32 (3)	0	0	0
Size of invasive cancer (cm)	1.0 (0.6–1.5)	1.0 (0.7–1.5)	1.5 (0.6–3.4)	1.1 (0.7–2.4)	1.1 (0.5–1.3)
Lymph node status					
No metastasis	104 (93.7)	53 (93)	14 (88)	4 (100)	33 (97)
Metastasis	7 (6.3)	4 (7)	2 (12)	0	1 (3)
Cancer type					
DCIS	42 (37.8)	25 (44)	5 (31)	0	12 (35)
Invasive cancer, stage I	55 (49.5)	27 (47)	6 (38)	3 (75)	19 (56)
Invasive cancer, stage II	11 (9.9)	5 (9)	4 (25)	1 (25)	1 (3)
Invasive cancer, stage III	1 (0.9)	0	1 (6)	0	0
Invasive cancer, stage IV	2 (1.8)	0	0	0	2 (6)
Invasive cancer stage I or II	66 (59.5)	32 (56)	10 (63)	4 (100)	20 (59)
Invasive cancer stage III or IV	3 (2.7)	0	1 (6)	0	2 (6)
Immunohistochemical subtype					
ER or PR positive	70 (63.1)	35 (61)	11 (69)	2 (50)	22 (65)
HER2 positive	16 (14.4)	12 (21)	1 (6)	0	3 (9)
Triple negative	21 (18.9)	8 (14)	3 (19)	2 (50)	8 (24)
Data not available	4 (3.6)	2 (4)	1 (6)	0	1 (3)

Note.—Data are presented as numbers with percentages in parentheses or medians with IQRs in parentheses. AI = artificial intelligence, DCIS = ductal carcinoma in situ, ER = estrogen receptor, HER2 = human epidermal growth factor receptor type 2, PR = progesterone receptor.

of 50 cancers missed by radiologists, suggesting improved cancer detection ability, 34 of 111 total breast cancers were undetected by either radiologists or AI, indicating that there is additional room for improvement in detecting asymptomatic contralateral breast cancers.

Efforts to improve earlier cancer detection in patients with a PHBC to date have focused on the use of supplemental US, MRI, or contrast-enhanced mammography (25–27). Although the results of these studies on supplemental surveillance imaging are promising, these modalities also introduce additional burdens such as increased time, cost, and use of contrast agents.

The advancement of AI technology in breast imaging has broadened in scope, extending beyond cancer detection in screening mammography (28). It can also be applied to density measurement (29), risk assessment (30,31), and determining the need for supplemental screening (32). The use of AI in the posttreatment surveillance setting has been tested in only one other study thus far (18). Specifically, Yoon et al (18) evaluated digital breast tomosynthesis and AI performance in addition to digital mammography in patients treated with breast-conserving surgery. In their retrospective reader study, significantly lower sensitivity and higher specificity were reported for ipsilateral breast evaluation, and higher specificity was reported for contralateral breast evaluation with AI support compared with that achieved with digital mammography without AI (18). There was no benefit of AI aid in terms of sensitivity, perhaps due to the limited sample of patients with cancer in that study (*n* = 6). Our findings provide support for the use of AI in mammographic surveillance for patients treated with unilateral mastectomy, revealing that AI substantially increases cancer detection. The recall rate of AI was higher than that of radiologist interpretation; however, this may be attributed to the remarkably low recall rate of radiologist interpretation in our study compared with that reported in prior studies (5). Our study cohort consisted of patients who underwent unilateral mammography with prior mammography, which may have contributed to the lower recall rate. In addition, the image-level data analysis used in our analysis (24) could also have influenced the number of AI detections. Indeed, AI systems can extract various types of information from mammographic images with different levels of analysis (ie, the case level, image level, and lesion level) (33).

Standalone AI detected 16 of the 50 cancers that were missed by radiologists. The cancers were mostly stage I or II, estrogen receptor– or progesterone receptor–positive invasive cancers without lymph node metastasis, indicating that AI may support early detection. However, the sensitivity of standalone AI for surveillance is relatively low (65.8%) compared with that of prior meta-analyses in screening patients without PHBC (approximately 80%) (16). This result may be due to the high prevalence of patients with dense breasts among those with a PHBC (66.5%), which may lower the performance of radiologists or AI in reading mammograms. Indeed, 38 of the 111 cancers in our study with

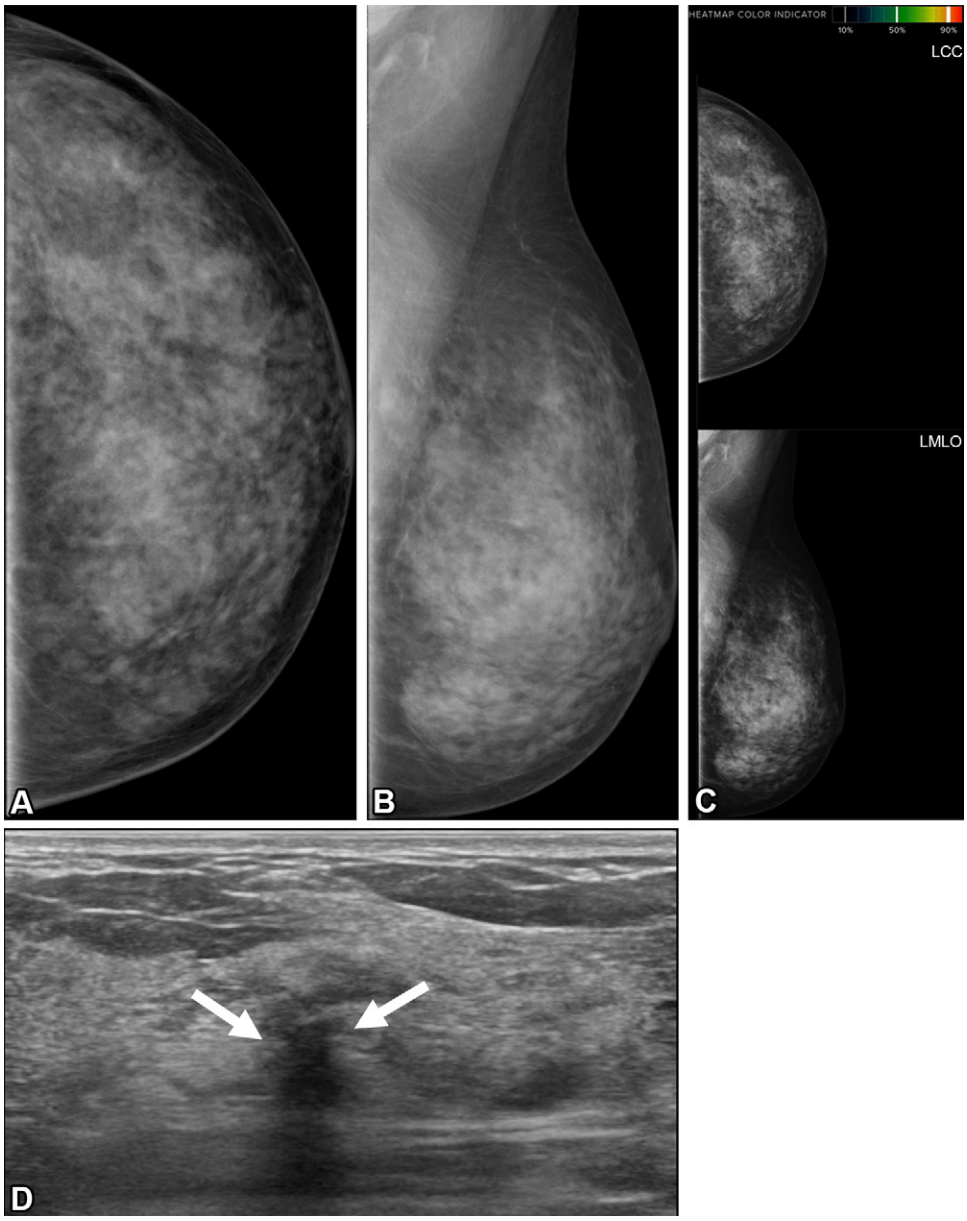

Figure 3: Images in a 51-year-old patient with contralateral second breast cancer 2.8 years after right mastectomy. **(A)** Left craniocaudal and **(B)** mediolateral oblique mammograms assessed as negative. **(C)** When artificial intelligence was applied, no abnormalities were noted. **(D)** Supplemental axial US scan, which was obtained on the same day as the mammogram, shows an irregular hypoechoic mass (arrows) in the left upper outer breast. Pathology results at biopsy revealed invasive ductal carcinoma (pT1N0, 1.0 cm, estrogen receptor negative, progesterone receptor negative, human epidermal growth factor receptor type 2 negative, histologic grade II). LCC = left craniocaudal, LMLO = left mediolateral oblique.

false-negative results (34.2%; four missed by AI and 34 missed by both radiologists and AI) were in patients with dense breasts (66%; 25 of 38); these cancers were mostly stage 0 or I (89%; 34 of 38) or cancers measuring 1 cm or less (66%; 25 of 38) and were detected at concurrent or subsequent supplemental breast US or MRI surveillance (79%; 30 of 38). These results imply that negative results obtained from the current version of the AI algorithm may not be appropriate for standalone use and that supplemental multimodality imaging may continue to improve early second breast cancer detection (34). Additionally, considering that false-negative assessments by radiologists and AI differ and that AI yielded 193 false-positive results to find 16 additional cancers, an ideal approach for the current algorithm could involve the use of an AI system as an interactive tool during radiologist interpretation. This approach could help prevent overlooking cancer findings or interpretation errors while simultaneously reducing false-positive findings caused by either radiologists or AI. Further prospective studies on the interpretation of surveillance mammography with AI aid are warranted.

Our study had several limitations. First, it was a retrospective study conducted at a single institution. Second, radiologists' mammography interpretations were made in comparison with those of prior mammograms, whereas the AI system did not consider prior mammograms. Third, patients with postsurgical changes were excluded. It is possible that the relative low specificity of AI may be even lower if it is applied to evaluate postsurgical breasts. Fourth, we included only two-dimensional mammography, and three-dimensional mammography such as digital breast

tomosynthesis or synthetic mammography were not included. Finally, AI results were obtained from a system of a single vendor that was not specifically designed for unilateral mammography evaluation. Thus far, there are no commercially available or even experimental AI algorithms specifically dedicated to unilateral mammography, and it is uncertain when such algorithms might be developed. Thus, the use of existing AI software for patients who have had mastectomies could be a more pragmatic and efficient approach for improving contralateral breast cancer detection rather than waiting until dedicated AI systems are developed for patients treated with mastectomy. Larger prospective studies using various AI software algorithms applied to both digital mammography and digital breast tomosynthesis examinations are needed to validate our findings and to support AI use for the surveillance of patients treated with unilateral mastectomy.

In conclusion, it may be feasible to apply a commercially available artificial intelligence (AI) algorithm to the standalone detection of contralateral breast cancers in patients treated with unilateral mastectomy, with AI achieving a higher cancer detection and sensitivity than that of radiologist interpretation alone. However, the relatively high proportion of cancers missed by both AI and radiologists underscores the need for further studies to support the translation of this promising technology into routine clinical practice.

Deputy Editor: Linda Moy
Scientific Editor: Shannyn Wolfe (AJE)

Author affiliations:

[1] Department of Radiology, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

[2] Department of Radiology, Seoul National University College of Medicine, Seoul, Republic of Korea

[3] Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, Republic of Korea

[4] Department of Radiology, University of Washington School of Medicine, Seattle, Wash

[5] Medical Research Collaborating Center, Seoul National University Hospital, Seoul, Republic of Korea

[6] Department of Surgery, Seoul National University College of Medicine and Seoul National University Hospital, Seoul, Republic of Korea

Received October 2, 2024; revision requested November 11; final revision received January 14, 2025; accepted January 16.

Address correspondence to: J.M.C. (email: imchangjm@gmail.com).

Funding: This research was supported by the Seoul National University Hospital research fund (grant 04-2023-2160).

Author contributions: Guarantors of integrity of entire study, **S.M.H., H.K.K., J.M.C.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **S.M.H., H.K.K., J.M.C.**; clinical studies, **S.M.H., H.K.K., J.M.C.**; statistical analysis, **S.M.H., M.J.J.**; and manuscript editing, **S.M.H., J.M.L., H.K.K., J.M.C.**

Disclosures of conflicts of interest: **S.M.H.** No relevant relationships. **J.M.L.** Research grant from GE HealthCare. **M.J.J.** No relevant relationships. **H.K.K.** No relevant relationships. **J.M.C.** Research consultant for Hologic; member of *Radiology* editorial board.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71(3):209–249.
2. Houssami N, Ciatto S. Mammographic surveillance in women with a personal history of breast cancer: how accurate? How effective? Breast 2010;19(6):439–445.
3. Lawson MB, Herschorn SD, Sprague BL, et al. Imaging Surveillance Options for Individuals With a Personal History of Breast Cancer: AJR Expert Panel Narrative Review. AJR Am J Roentgenol 2022;219(6):854–868.
4. Lee CI, Abraham L, Miglioretti DL, et al. National Performance Benchmarks for Screening Digital Breast Tomosynthesis: Update from the Breast Cancer Surveillance Consortium. Radiology 2023;307(4):e222499.
5. Lee JM, Ichikawa LE, Wernli KJ, et al. Digital Mammography and Breast Tomosynthesis Performance in Women with a Personal History of Breast Cancer, 2007-2016. Radiology 2021;300(2):290–300.
6. Lamb LR, Mohallem Fonseca M, Verma R, Seely JM. Missed Breast Cancer: Effects of Subconscious Bias and Lesion Characteristics. RadioGraphics 2020;40(4):941–960.
7. Kim HE, Kim HH, Han BK, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. Lancet Digit Health 2020;2(3):e138–e148.
8. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. J Natl Cancer Inst 2019;111(9):916–922.
9. Pacilè S, Lopez J, Chone P, Bertinotti T, Grouin JM, Fillard P. Improving Breast Cancer Detection Accuracy of Mammography with the Concurrent Use of an Artificial Intelligence Tool. Radiol Artif Intell 2020;2(6):e190208.
10. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. Eur Radiol 2019;29(9):4825–4832.
11. Dembrower K, Wåhlin E, Liu Y, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. Lancet Digit Health 2020;2(9):e468–e474.
12. Lauritzen AD, Rodríguez-Ruiz A, von Euler-Chelpin MC, et al. An Artificial Intelligence-based Mammography Screening Protocol for Breast Cancer: Outcome and Radiologist Workload. Radiology 2022;304(1):41–49.
13. Wanders AJT, Mees W, Bun PAM, et al. Interval Cancer Detection Using a Neural Network and Breast Density in Women with Negative Screening Mammograms. Radiology 2022;303(2):269–275.
14. Larsen M, Aglen CF, Lee CI, et al. Artificial Intelligence Evaluation of 122 969 Mammography Examinations from a Population-based Screening Program. Radiology 2022;303(3):502–511.
15. Park GE, Kang BJ, Kim SH, Lee J. Retrospective Review of Missed Cancer Detection and Its Mammography Findings with Artificial-Intelligence-Based, Computer-Aided Diagnosis. Diagnostics (Basel) 2022;12(2):387.
16. Yoon JH, Strand F, Baltzer PAT, et al. Standalone AI for Breast Cancer Detection at Screening Digital Mammography and Digital Breast Tomosynthesis: A Systematic Review and Meta-Analysis. Radiology 2023;307(5):e222639.
17. Lång K, Josefsson V, Larsson AM, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol 2023;24(8):936–944.
18. Yoon JH, Kim EK, Kim GR, Han K, Moon HJ. Mammographic Surveillance After Breast-Conserving Therapy: Impact of Digital Breast Tomosynthesis and Artificial Intelligence-Based Computer-Aided Detection. AJR Am J Roentgenol 2022;218(1):42–51.
19. American Joint Committee on Cancer. AJCC cancer staging manual. 8th ed. Springer, 2017.
20. Ha SM, Lee JM, Kim SO, Moon WK, Chang JM. Semiannual Breast US or MRI Screening in Patients with a Personal History of Breast Cancer. Radiology 2023;307(5):e221660.
21. D'Orsi CJ, Bassett LW, Berg WA, et al. Mammography. In: Breast Imaging reporting and data system (BIRADS). 4th ed. American College of Radiology, 2003.
22. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS Atlas: Breast Imaging Reporting and Data System. 5th ed. American College of Radiology, 2013.
23. Harris E. FDA Updates breast density reporting standards, other mammogram rules. JAMA 2023;329(14):1142–1143.
24. Kim EK, Kim HE, Han K, et al. Applying Data-driven Imaging Biomarker in Mammography for Breast Cancer Screening: Preliminary Study. Sci Rep 2018;8(1):2762.
25. Cho N, Han W, Han BK, et al. Breast Cancer Screening With Mammography Plus Ultrasonography or Magnetic Resonance Imaging in Women 50 Years or Younger at Diagnosis and Treated With Breast Conservation Therapy. JAMA Oncol 2017;3(11):1495–1502.

26. Kim SY, Cho N, Kim SY, et al. Supplemental Breast US Screening in Women with a Personal History of Breast Cancer: A Matched Cohort Study. Radiology 2020;295(1):54–63.
27. Berg WA, Berg JM, Bandos AI, et al. Addition of Contrast-enhanced Mammography to Tomosynthesis for Breast Cancer Detection in Women with a Personal History of Breast Cancer: Prospective TOCEM Trial Interim Analysis. Radiology 2024;311(1):e231991.
28. Lamb LR, Lehman CD, Gastounioti A, Conant EF, Bahl M. Artificial Intelligence (AI) for Screening Mammography, From the *AJR* Special Series on AI Applications. AJR Am J Roentgenol 2022;219(3):369–380.
29. Lehman CD, Yala A, Schuster T, et al. Mammographic Breast Density Assessment Using Deep Learning: Clinical Implementation. Radiology 2019;290(1):52–58.
30. Lowry KP, Zuiderveld CC. Artificial Intelligence for Breast Cancer Risk Assessment. Radiol Clin North Am 2024;62(4):619–625.
31. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction. Radiology 2019;292(1):60–66.
32. Liu Y, Sorkhei M, Dembrower K, Azizpour H, Strand F, Smith K. Use of an AI Score Combining Cancer Signs, Masking, and Risk to Select Patients for Supplemental Breast Cancer Screening. Radiology 2024;311(1):e232535.
33. Hickman SE, Woitek R, Le EPV, et al. Machine Learning for Workflow Applications in Screening Mammography: Systematic Review and Meta-Analysis. Radiology 2022;302(1):88–104.
34. Ha SM, Jang MJ, Youn I, et al. Screening Outcomes of Mammography with AI in Dense Breasts: A Comparative Study with Supplemental Screening US. Radiology 2024;312(1):e233391.