UC Irvine
UC Irvine Previously Published Works

Title

Dynamic Changes of Convolutional Neural Network-based Mammographic Breast Cancer Risk Score Among Women Undergoing Chemoprevention Treatment.

Permalink

https://escholarship.org/uc/item/32t1z3sd

Journal

Authors

Manley, Haley

Mutasa, Simukayi

Chang, Peter

et al.

Publication Date

DOI

Peer reviewed

Dynamic Changes of Convolutional Neural Network-based Mammographic Breast Cancer Risk Score Among Women Undergoing Chemoprevention Treatment

Haley Manley[1], **Simukayi Mutasa**[2], **Peter Chang**[3], **Elise Desperito**[4], **Katherine Crew**[5], **Richard Ha**[6]

[1]Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX

[2]Department of Radiology, New York-Presbyterian/Columbia University Medical Center, New York, NY

[3]Department of Radiological Sciences and Center for Artificial Intelligence in Diagnostic Medicine, University of California, Irvine, Irvine, CA

[4]Department of Radiology, New York-Presbyterian/Columbia University Medical Center, New York, NY

[5]Departments of Medicine and Epidemiology, New York-Presbyterian/Columbia University Medical Center, New York, NY

[6]Department of Radiology and Breast Imaging Section, New York-Presbyterian/Columbia University Medical Center, New York, NY

Abstract

We investigated whether our convolutional neural network-based breast cancer risk model is modifiable by testing it on women who had undergone risk-reducing treatment with known chemoprevention agents. Compared with baseline, significantly more women in the treatment group had a decrease in the breast cancer risk score ($P < .01$), indicating that our convolutional neural network risk model is modifiable with potential utility in assessing the efficacy of chemoprevention strategies.

Introduction: We investigated whether our convolutional neural network (CNN)-based breast cancer risk model is modifiable by testing it on women who had undergone risk-reducing chemoprevention treatment.

Materials and Methods: We conducted a retrospective cohort study of patients diagnosed with atypical hyperplasia, lobular carcinoma in situ, or ductal carcinoma in situ at our institution from 2007 to 2015. The clinical characteristics, chemoprevention use, and mammography images were extracted from the electronic health records. We classified two groups according

to chemoprevention use. Mammograms were performed at baseline and subsequent follow-up evaluations for input to our CNN risk model. The 2 chemoprevention groups were compared for the risk score change from baseline to follow-up. The change categories included stayed high risk, stayed low risk, increased from low to high risk, and decreased from high to low risk. Unordered polytomous regression models were used for statistical analysis, with $P < .05$ considered statistically significant.

Results: Of 541 patients, 184 (34%) had undergone chemoprevention treatment (group 1) and 357 (66%) had not (group 2). Using our CNN breast cancer risk score, significantly more women in group 1 had shown a decrease in breast cancer risk compared with group 2 (33.7% vs. 22.9%; $P < .01$). Significantly fewer women in group 1 had an increase in breast cancer risk compared with group 2 (11.4% vs. 20.2%; $P < .01$). On multivariate analysis, an increase in breast cancer risk predicted by our model correlated negatively with the use of chemoprevention treatment ($P = .02$).

Conclusions: Our CNN-based breast cancer risk score is modifiable with potential utility in assessing the efficacy of known chemoprevention agents and testing new chemoprevention strategies.

Keywords

Breast cancer risk; Breast density; CNN; Deep learning; Tamoxifen

Introduction

Breast cancer is the leading cause of cancer-related mortality worldwide in women, with most cases occurring in the United States and Western Europe.[1] Increased mammographic breast density (BD), which describes the radiologically appearing white tissue on a mammogram, is a well-known breast cancer risk factor.[1] Using the 4 descriptors for BD on mammograms based on the BI-RADS (Breast Imaging-Reporting and Data System), women with an extensive degree of BD have a two- to sixfold greater risk of breast cancer compared with women with little BD.[1] Although BD is a risk factor for breast cancer, it has been difficult to establish screening guidelines for this population because > 50% of women aged < 50 years will have high BD but not all will be high risk themselves.[1] Furthermore, it is difficult to assess BD serially on an individual basis because BD can be vary owing to positional changes of the breast during mammogram acquisition.

Recently, a subset of machine learning processes termed deep learning (DL) and using an artificial neural network such as a convolutional neural network (CNN) has made great strides in medical imaging analysis. In contrast to traditional machine learning, which primarily relies on human-chosen feature analysis, neural networks depend on the input of raw data and allows the computer to automatically construct predictive statistical models through increasingly complex layers and self-optimization.[2–4] Our laboratory previously developed a novel CNN algorithm for breast cancer risk prediction using 1474 mammographic images.[5] Our results showed that both the CNN-based mammographic risk model and BD were significant independent predictors of breast cancer risk. The CNN risk model showed greater predictive potential compared with BD.[5] The conclusion of our study

indicated that a CNN algorithm can be used to stratify breast cancer risk, potentially better than, and independently of, the BD.[5]

Subsequently, Yala et al[6] conducted a larger retrospective study of 88,994 consecutive screening mammograms from 39,571 women. They used 3 models to assess breast cancer risk within 5 years: traditional risk assessment using the Tyrer-Cuzick model, version 8; a DL model using mammographic data; and a hybrid DL model using the Tyrer-Cuzick model and DL. Similar to the results of our study,[5] they showed that DL models substantially improved breast cancer risk discrimination compared with the Tyrer-Cuzick model, which includes BD as 1 of the risk factors.[6,7]

Most recently, Dembrower et al[8] developed a DL model to estimate breast cancer risk using mammographic images and tested the model on 2283 women. Their DL risk score output, which reflects the likelihood of developing breast cancer using standard mammographic views (craniocaudal and mediolateral oblique views) was used to estimate an individual's breast cancer risk. They compared the accuracy of the DL model with that of 2 different models based on BD, including BD measurements performed by automated software. Similar to the findings in our study, their DL model demonstrated a higher age-adjusted risk association for breast cancer compared with the best mammographic BD model.[8] In addition, the area under the curve for the DL model was greater than that for a model based on patient age and BD area. They concluded that a DL model could more accurately predict for breast cancer risk than could BD-based models.[8]

These 3 cited studies have provided strong evidence that mammographic features identified using DL can provide additional insight into an individual's risk of breast cancer beyond the mere quantification of the BD. The purpose of the present study was to further evaluate our CNN-based risk model to determine whether the model can be modified by testing it on patients who have undergone known risk-reducing chemoprevention treatment (eg, tamoxifen and aromatase inhibitors). If our CNN-based risk model is modifiable, it would have potential added utility for assessing the efficacy of treatment for patients receiving chemoprevention agents and testing new chemoprevention strategies.

Materials and Methods

Study Design and Study Population

We conducted a retrospective cohort study of patients who had received a diagnosis of atypical hyperplasia (AH), lobular carcinoma in situ (LCIS), or ductal carcinoma in situ (DCIS) at our institution from 2007 to 2015 to determine the association between chemoprevention uptake and change in CNN risk. The inclusion criteria for the present study were (1) a history of AH, LCIS, or DCIS without concurrent or previous invasive breast cancer; (2) for those with DCIS, estrogen receptor-positive (ER$^+$) and/or progesterone receptor-positive (PR$^+$) tumor status; and (3) 2 serial mammograms available at our institution after the diagnosis of AH, LCIS, or DCIS or after receipt of chemoprevention. Subjects with a history of bilateral mastectomy were excluded. All patients were considered eligible for chemoprevention use because of the diagnosis of AH, LCIS, or ER$^+$ and/or PR$^+$ DCIS. The institutional review board at our institution approved the present study.

Data Collection From Electronic Health Records

The patients' demographics, breast cancer risk factors, and medical information were collected through a medical record review and data extraction from the electronic health records (EHRs) at our institution. We captured data from the EHRs using diagnostic codes, breast pathology reports, and outpatient clinic notes. EHR data extraction also included the tumor registry, which identified incident cases of LCIS and DCIS. All patients with a diagnosis of AH or LCIS/DCIS were initially identified by the corresponding International Classification of Diseases (ICD)-9th or 10th revision, codes in the databases (codes 610.9/N60.99 and 233.0/D05.90, respectively), because LCIS and DCIS share the same ICD-9 and ICD-10 codes. If the patients had had > 1 diagnosis, they were classified by their most advanced breast lesion (DCIS more advanced than LCIS more advanced than AH). Any medical record documentation of invasive breast cancer was identified by the ICD-9 code 174.9. The tumor registry and pathology reports were used to identify those patients who had had invasive breast cancer before or concurrent with the diagnosis of AH, LCIS, or DCIS. These patients were also excluded.

Other covariates collected included age at the baseline mammogram, race and ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian, and other), menopausal status, body mass index, current or former hormone replacement therapy use (yes vs. no), and alcohol use (yes vs. no). Patients with missing information regarding menopausal status were considered postmenopausal if they were aged > 55 years. Subjects with missing information regarding the body mass index were classified as unknown.

The primary exposure was selective estrogen receptor modulator (SERM) or aromatase inhibitor (AI) use as documented in the medication list of the EHRs at any point and was dichotomized as yes versus never used. The chemoprevention agents used were also identified and categorized as tamoxifen, raloxifene, AIs (ie, anastrozole, exemestane), or multiple agents (ie, patients could have switched medications owing to toxicity).

CNN Architecture

For input, the mammogram from the contralateral unaffected side was used to limit the potentially confounding results from post-treatment changes of the affected side. We used our previously developed CNN-based risk model in the present study.[5] In brief, the CNN is based on a modified U-net architecture and implemented completely by a series of 3 × 3 convolutional kernels to limit overfitting. No pooling layers were used. Instead, downsampling was implemented simply using 3 × 3 strided convolutions to decrease the feature maps by 75%. All nonlinear functions were modeled by the rectified linear unit. Batch normalization was used between the convolutional rectified linear unit layers to limit the drift of layer activations during training. In successively deep layers, the number of feature channels will gradually increase from 16, 32, 64, 128, to 256, reflecting the increasing representational complexity. Each mammogram was background nulled and normalized using contrast-adaptive histogram normalization and resized to an input image of 256 × 256. The data augmentation used in the present study involved a number of real-time modifications, including affine warps to the source images at the time of training. Training was implemented using the Adam optimizer, an algorithm for the first-order

gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower order moments. Parameters were initialized using the heuristic described by He et al.[9] L2 regularization was implemented to limit overfitting of data by limiting the squared magnitude of the kernel weights. To account for training dynamics, the learning rate was annealed. A final SoftMax score threshold of 0.5 from the average of raw logits from each pixel was used for 2-class classification (high risk vs. low risk). The software code for the present study was written in Python, version 3.6, using the TensorFlow module, version 1.13.1. Experiments and CNN training were performed using a Linux workstation with the Nvidia GTX 1070 pascal GPU with 8 GB of chip memory, i7 CPU, and 32 GB of RAM.

Statistical Analysis

All statistical analyses included unordered polytomous regression models and used SAS, version 9.4 (SAS Institute, Cary, NC), and a *P* value of < .05 was considered statistically significant.

Results

From January 2007 to December 2015, ~2933 patients with an ICD-9/10 code for AH, LCIS, or DCIS were initially identified through the EHRs. Of these patients, 541 (18.4%) had met all the inclusion criteria and were included in our final analysis. Of the 2452 patients excluded from the original dataset, 1238 (50.4%) had had evidence of invasive breast cancer either before or concurrently with the diagnosis of AH, LCIS, or DCIS and were excluded. An additional 69 patients (2.8%) were excluded owing to a history of bilateral mastectomy or ER⁻/PR⁻ DCIS. Also, 108 (4.4%) were excluded because whether they had had LCIS or DCIS could not be clarified from their medical records. In addition, 250 patients (10.2%) had not undergone a baseline mammogram before beginning chemoprevention therapy and 787 (32.0%) did not have follow-up screening mammograms available and were excluded.

The baseline characteristics of the study population are listed in Table 1. The mean age of the included patients was 60 years (range, 27–90 years), and more than two thirds were postmenopausal. The average age at menopause was 49 years, and the average age at menarche was ~13 years. The included patients were racially and ethnically diverse, with 36.7% non-Hispanic white, 11.6% non-Hispanic black, 36.4% Hispanic or Latina, 5.7% Asian, and 9.4% other. Of the 541 patients, 206 (38%) had DCIS, 215 (39.7%) had AH, and 120 (22.3%) had LCIS.

Among the 541 included subjects, 184 (34.0%) had a history of using SERMs or AIs. Approximately 72% of the women undergoing chemoprevention treatment used a SERM, 19% used an AI, and 9% used multiple agents. Of the women receiving chemoprevention treatment, most (n = 89; 48.4%) had DCIS, 42 (22.8%) had LCIS, and 53 (28.8%) had AH.

We assessed the long-term changes in the CNN-based risk score compared with the baseline risk score (Table 2). The average duration between the baseline and follow-up mammograms was 48 ± 12 months. The average duration between the start of treatment and the follow-up mammogram was 37.5 ± 8 months. More patients in the chemoprevention treatment

group had a decrease in their CNN-based risk score (33.7%) compared with the patients in the no treatment group (22.9%; $P < .01$). Case examples of the dynamic changes in the CNN-based risk score are presented in Figures 1 and 2 between 1 patient who had received chemoprevention treatment (Figure 1) and 1 patient who had not received treatment (Figure 2). In addition, fewer patients in the chemoprevention treatment group had had an increase in their CNN-based risk score (11.4%) compared with patients in the no treatment group (20.2%; $P < .01$). On multivariate analysis, an increase in the CNN-based risk score was negatively associated with chemoprevention treatment ($P = .02$). Thus, those in the chemoprevention treatment group were 1.29 times more likely to have a decrease in their CNN-based risk score compared with those in the no treatment group.

Discussion

In the present study, we have demonstrated that our CNN-based breast cancer risk score is modifiable with known chemoprevention agents. The group of patients who underwent chemoprevention treatment had a significant decrease in breast cancer risk compared with the group who had not undergone treatment. The modifiability of our risk model means that it has the potential to be used as an assessment tool to measure the effectiveness of known chemoprevention agents and testing novel chemoprevention strategies.

To test the modifiability of our CNN-based algorithm, we used patients with a diagnosis of LCIS, DCIS, or AH, including atypical lobular hyperplasia and atypical ductal hyperplasia because randomized controlled trials have demonstrated that SERMs, such as tamoxifen and raloxifene, and AIs, such as anastrozole and exemestane, taken for 5 years will reduce the breast cancer incidence among high-risk women by 50% to 65%, especially among women with AH and LCIS.[10] The National Surgical Adjuvant Breast and Bowel Project P1 trial found that tamoxifen reduced the risk of invasive breast cancer by 49% and the risk of noninvasive breast cancer (DCIS and LCIS) by 50%. In women with a history of AH, the risk was reduced by 86%.[11]

According to our hypothesis, the follow-up CNN-based risk assessment compared with the baseline assessment showed that significantly more patients in the treatment group had decrease in breast cancer risk compared with that in the control group. In addition, significantly fewer women in the treatment group had increase in breast cancer risk compared with the treatment group. This is consistent with our understanding of the therapeutic role of a chemoprevention agent to both decrease and maintain the risk of breast cancer. Despite the known efficacy of chemoprevention therapies to reduce such risk, the compliance with these treatments among women with a diagnosis of AH, LCIS, or DCIS has been limited.[12] For these women, the compliance has been estimated to be < 15%.[13] The reasons for such low compliance include the lack of a routine breast cancer risk assessment, the lack of knowledge among patients and providers about chemoprevention treatment, and concerns about side effects. Similarly, in our cohort, only approximately one third of the patients had undergone treatment and most of the patients had not. The potential utility of our CNN-based risk model could be to better assess the risk at an individual level and better define the need for treatment and to measure the efficacy in patients undergoing treatment.

A recent editorial by Bahl,[14] titled "Harnessing the Power of Deep Learning to Assess Breast Cancer Risk," has highlighted the DL algorithm developed by Yala et al,[6] Dembrower et al,[8] and Ha et al.[5] She stated that these studies show that mammographic images contain indicators of risk not captured by the use of BD alone, highlighting the power of DL. Image-based DL models are based on the rich information contained within a mammographic image, and not on the subjectivity and variability inherent in human assessments of imaging features such as BD. She concluded that such models could potentially replace existing risk prediction models and that continued work is needed to strengthen and evaluate the risk models to support personalized screening and prevention strategies and, ultimately, reduce the burden of breast cancer.[14]

The present study was limited by the relatively small number of cases from a single institution and the retrospective performance. A larger number of cases from multiple institutions are needed to further validate and fine tune our algorithm for potential use in a clinical setting. The main objective of the present study was to show that our DL model can be modifiable. However, future work is planned to fully evaluate the effectiveness of tamoxifen and AIs and the potential differences between these 2 types of chemoprevention agents. Finally, a direct comparison of our DL model with BD could be of value in understanding the relationships among these variables and is planned for a future investigation.

Conclusion

Our CNN-based breast cancer risk score is modifiable with potential use in a clinical setting, not only to assess an individual's risk of breast cancer, but also to evaluate the efficacy of known chemoprevention agents and novel chemoprevention strategies.

Acknowledgments

References

1. Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer. Breast Cancer2018; 25:259–67. [PubMed: 29651637]

2. Mullooly M, Ehteshami Bejnordi B, Pfeiffer RM, et al.Application of convolutional neural networks to breast biopsies to delineate tissue correlates of mammographic breast density. NPJ Breast Cancer2019; 5:43. [PubMed: 31754628]

3. Heine JJ, Cao K, Thomas JA. Effective radiation attenuation calibration for breast density: compression thickness influences and correction. Biomed Eng Online2010; 9:73. [PubMed: 21080916]

4. Ha R, Mutasa S, Sant EPV, et al.Accuracy of distinguishing atypical ductal hyperplasia from ductal carcinoma in situ with convolutional neural network-based machine learning approach using mammographic image data. AJR Am J Roentgenol2019:1–6.

5. Ha R, Chang P, Karcich J, et al.Convolutional neural network based breast cancer risk stratification using a mammographic dataset. Acad Radiol2019; 26:544–9. [PubMed: 30072292]

6. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. Radiology2019; 292:60–6. [PubMed: 31063083]

7. Cuzick J, Warwick J, Pinney E, et al.Tamoxifen and breast density in women at increased risk of breast cancer. J Natl Cancer Inst2004; 96:621–8. [PubMed: 15100340]

8. Dembrower K, Liu Y, Azizpour H, et al.Comparison of a deep learning risk score and standard mammographic density score for breast cancer risk prediction. Radiology2020; 294:265–72. [PubMed: 31845842]

9. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification, Available at: https://arxiv.org/abs/1502.01852.Accessed: February 6, 2015.

10. Reimers L, Crew KD. Tamoxifen versus raloxifene versus exemestane for chemoprevention. Curr Breast Cancer Rep2012; 4:207–15. [PubMed: 23956815]

11. Roetzheim RG, Lee JH, Fulp W, et al.Acceptance and adherence to chemoprevention among women at increased risk of breast cancer. Breast2015; 24:51–6. [PubMed: 25491191]

12. Ropka ME, Keim J, Philbrick JT. Patient decisions about breast cancer chemoprevention: a systematic review and meta-analysis. J Clin Oncol2010; 28:3090–5. [PubMed: 20458026]

13. Trivedi MS, Coe AM, Vanegas A, Kukafka R, Crew KD. Chemoprevention uptake among women with atypical hyperplasia and lobular and ductal carcinoma in situ. Cancer Prev Res (Phila)2017; 10:434–41. [PubMed: 28611039]

14. Bahl MHarnessing the power of deep learning to assess breast cancer risk. Radiology2020; 294:273–4. [PubMed: 31846401]

Clinical Practice Points

- Recent studies have provided strong evidence that mammographic features identified by DL can provide additional insight into individual's risk of breast cancer.

- We evaluated our CNN-based risk model to determine whether it is modifiable by testing it on patients who had received treatment with known risk-reducing chemoprevention agents (tamoxifen and AIs).

- The results from the present study have indicated that our CNN-based risk model is modifiable, with potential added utility for assessing the efficacy of chemoprevention treatment in patients and testing new chemoprevention strategies.
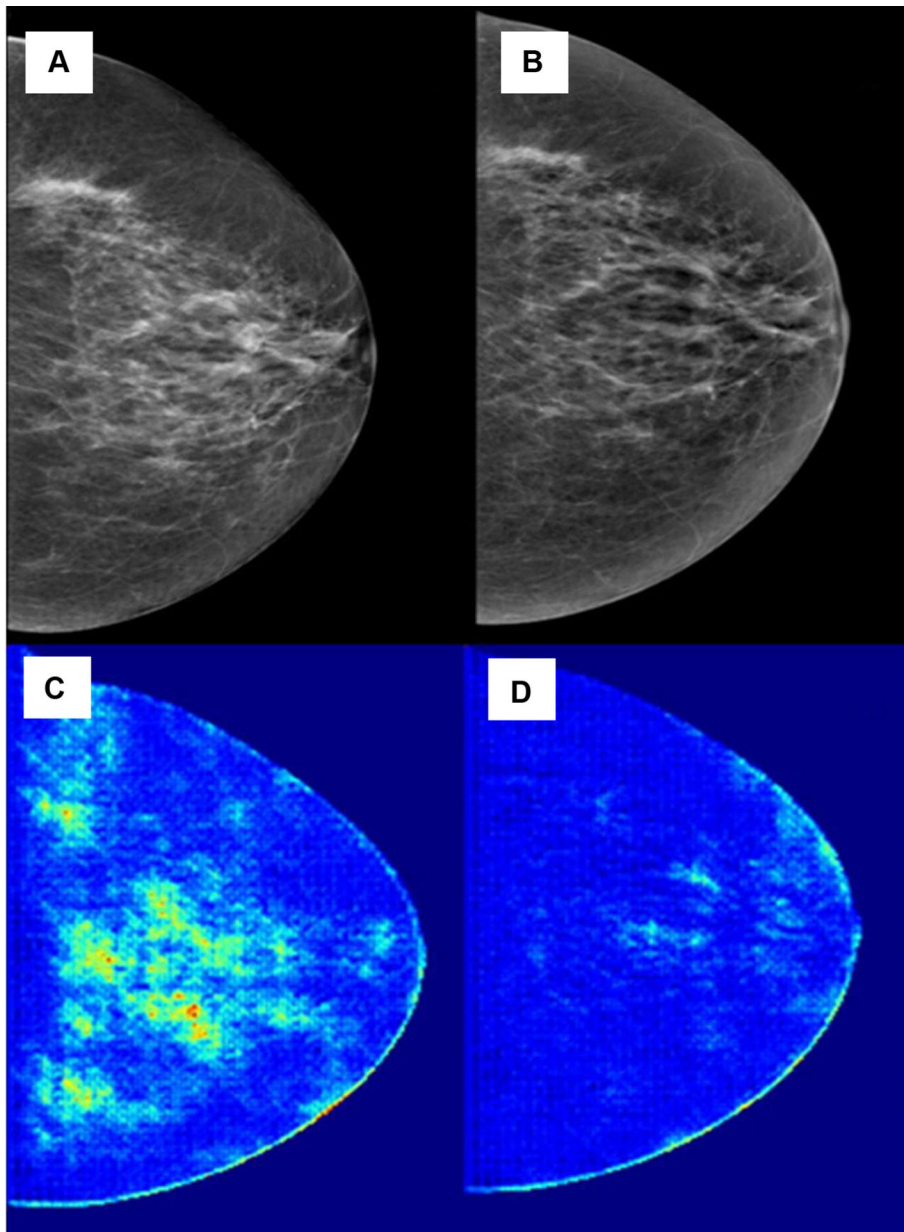
Figure 1.
Imaging Studies of a 62-Year-old Woman With a History of Right Breast Ductal Carcinoma in Situ Who had Undergone Tamoxifen Adjuvant Therapy. Contralateral Left Breast Mammogram (A) before Treatment and (B) 36 months after Treatment. Convolutional Neural Network Model Analysis Yielded Changes in Risk Category From High Risk before Treatment (C) to Low Risk at 36 months after Treatment (D). Pixel-wise Heat Maps Showing Significant Areas of High-Risk Pattern (Color Coded *green*, *yellow*, and *red*) before (C) and after (D) Treatment, Showing a Low-Risk Pixel-wise Pattern Color Predominantly Coded in *blue*
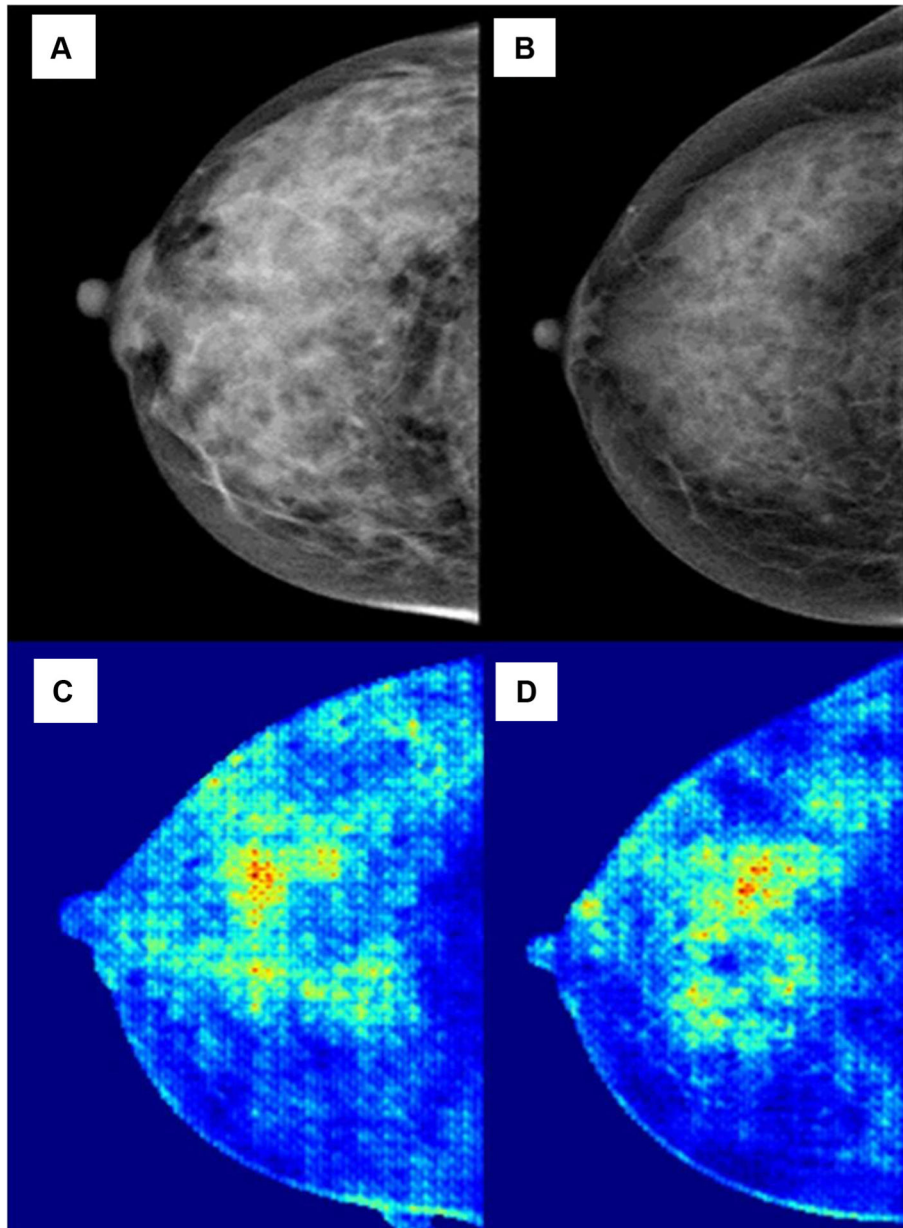
Figure 2.
Imaging Studies of a 52-Year-old Woman With a History of Left Breast Ductal Carcinoma in Situ after Mastectomy Who had Not Undergone Adjuvant Hormonal Therapy. Contralateral Right Breast Mammograms (A) at Baseline and (B) 54 months Later. Convolutional Neural Network Model Analysis Yielded No Change in Risk Category, which Remained in the High-Risk Category at Both Time Points. Pixel-wise Heat Maps Showing Significant Areas of High Risk (Color Coded *green*, *yellow*, and *red*) at Baseline (C), With No Significant Change 54 months Later (D)

Table 1

Demographic Characteristics Stratified by Chemoprevention Treatment (n = 541)

| Variable | Chemoprevention Treatment | | Total (n = 541) |
	Yes (n = 184)	No (n = 357)	
Age at diagnosis			
<45 y	9 (4.4)	24 (6.70)	33 (5.9)
45–54 y	49 (26.8)	110 (30.8)	159 (29.4)
55–64 y	59 (32.2)	111 (31.1)	170 (31.5)
65–74 y	49 (26.8)	77 (21.6)	126 (23.3)
75 y	18 (9.8)	35 (9.8)	53 (9.9)
Race/ethnicity			
Non-Hispanic white	71 (38.3)	129 (36.2)	200 (36.9)
Non-Hispanic black	22 (12.0)	41 (11.5)	63 (11.7)
Hispanic or Latina	67 (36.6)	130 (36.4)	197 (36.5)
Asian	18 (9.8)	13 (3.6)	31 (5.7)
Other	6 (3.3)	44 (12.3)	50 (9.2)
Body mass index			
Underweight (< 18.5 kg/m^2)	3 (1.1)	8 (2.2)	10 (1.9)
Normal weight (18.5–24.9 kg/m^2)	55 (30.1)	101 (28.3)	156 (28.9)
Overweight (25–29.9 kg/m^2)	63 (34.4)	116 (32.5)	179 (33.2)
Obese (30 kg/m^2)	63 (34.4)	102 (28.6)	165 (30.6)
Unknown	0 (0.0)	30 (8.4)	30 (5.4)
Alcohol use			
Yes	84 (45.7)	145 (40.6)	229 (42.3)
No	95 (51.6)	185 (51.8)	280 (51.8)
Unknown	5 (2.7)	27 (7.6)	32 (5.9)
Smoking			
Never	134 (72.8)	250 (70.0)	384 (71.0)
Previously	36 (19.6)	62 (17.4)	98 (18.1)
Current	12 (6.5)	19 (5.3)	31 (5.7)
Unknown	2 (1.1)	26 (7.3)	28 (5.2)
Menopausal status			
Yes	137 (74.5)	236 (66.1)	373 (69.0)
No	47 (25.5)	121 (33.9)	168 (31.0)
HRT use			
Yes	17 (9.1)	4 (1.1)	21 (3.7)
No	167 (90.9)	353 (98.9)	520 (96.3)
Breast lesion type			
AH	53 (28.8)	162 (45.4)	215 (39.7)
DCIS	89 (48.4)	117 (32.8)	206 (38.1)
LCIS	42 (22.8)	78 (21.8)	120 (22.2)

Data presented as n (%).

Abbreviations: AH = atypical hyperplasia; DCIS = ductal carcinoma in situ; HRT = hormone replacement therapy; LCIS = lobular carcinoma in situ.

Table 2

CNN-based Risk Score Change Compared With Baseline Risk Score Stratified by Chemoprevention Use[a]

CNN Risk Score Change Category	Chemoprevention	
	Yes (n = 184)	No (n = 357)
Decreased	62 (33.7)	82 (22.9)
Increased	21 (11.4)	72 (20.2)
Stayed high	58 (31.5)	104 (29.1)
Stayed low	43 (23.4)	99 (27.7)

Abbreviation: CNN = convolutional neural network.

[a]$P < .01$.