Radiology: Imaging Cancer

ORIGINAL RESEARCH

Machine Learning Prediction of Lymph Node Metastasis in Breast Cancer: Performance of a Multi-institutional MRI-based 4D Convolutional Neural Network

Dogan S. Polat, MD • Son Nguyen, PhD • Paniz Karbasi, PhD[1] • Keith Hulsey, PhD • Murat Can Cobanoglu, PhD[2] • Liqiang Wang, PhD • Albert Montillo, PhD • Basak E. Dogan, MD

From the Department of Diagnostic Radiology (D.S.P., K.H., A.M., B.E.D.), Lyda Hill Department of Bioinformatics (S.N., P.K., M.C.C., L.W., A.M.), and Biomedical Engineering Department (A.M.), University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-8585. Received July 18, 2023; revision requested September 27; revision received December 31; accepted February 29, 2024. **Address correspondence to** D.S.P. (email: *Dogan.polat@utsouthwestern.edu*).

[1]Current address: NVIDIA, Santa Clara, Calif (not affiliated with NVIDIA at the time of authorship).

[2]Current address: Exact Sciences, Madison, Wis (not affiliated with Exact Sciences at the time of authorship).

S.N. and A.M. supported by the National Institutes of Health (NIH) National Institute of General Medical Sciences (grant no. R01GM144486). A.M. supported by NIH National Institute of Aging (grant no. R01AG059288), NIH National Cancer Institute (grant no. U01 CA207091), the King Foundation, and the Lyda Hill Foundation.

Conflicts of interest are listed at the end of this article.

Radiology: Imaging Cancer 2024; 6(3):e230107 • https://doi.org/10.1148/rycan.230107 • Content codes: BR MR OI

Purpose: To develop a custom deep convolutional neural network (CNN) for noninvasive prediction of breast cancer nodal metastasis.

Materials and Methods: This retrospective study included patients with newly diagnosed primary invasive breast cancer with known pathologic (pN) and clinical nodal (cN) status who underwent dynamic contrast-enhanced (DCE) breast MRI at the authors' institution between July 2013 and July 2016. Clinicopathologic data (age, estrogen receptor and human epidermal growth factor 2 status, Ki-67 index, and tumor grade) and cN and pN status were collected. A four-dimensional (4D) CNN model integrating temporal information from dynamic image sets was developed. The convolutional layers learned prognostic image features, which were combined with clinicopathologic measures to predict cN0 versus cN+ and pN0 versus pN+ disease. Performance was assessed with the area under the receiver operating characteristic curve (AUC), with fivefold nested cross-validation.

Results: Data from 350 female patients (mean age, 51.7 years ± 11.9 [SD]) were analyzed. AUC, sensitivity, and specificity values of the 4D hybrid model were 0.87 (95% CI: 0.83, 0.91), 89% (95% CI: 79%, 93%), and 76% (95% CI: 68%, 88%) for differentiating pN0 versus pN+ and 0.79 (95% CI: 0.76, 0.82), 80% (95% CI: 77%, 84%), and 62% (95% CI: 58%, 67%), respectively, for differentiating cN0 versus cN+.

Conclusion: The proposed deep learning model using tumor DCE MR images demonstrated high sensitivity in identifying breast cancer lymph node metastasis and shows promise for potential use as a clinical decision support tool.

Supplemental material is available for this article.

Breast cancer is the leading cause of cancer death in female individuals, responsible for approximately 40 000 deaths annually in the United States. Most deaths are due to metastatic disease, for which the first site is usually an ipsilateral axillary lymph node (1,2). Nodal status is one of the most important factors determining disease prognosis and guiding treatment decisions such as neoadjuvant chemotherapy, radiation therapy, and axillary lymph node dissection (3,4). There are two strategies to identify axillary lymph node metastasis: *(a)* a combination of clinical examination, imaging, and imaging-guided biopsy, which determines clinical node (cN) stage; or *(b)* surgical staging with sentinel lymph node biopsy (SLNB) or axillary dissection to determine the pathologic node (pN) stage (3). Preoperative US with guided needle biopsy helps identify a median of 50% of patients with axillary nodal metastases prior to SLNB (5,6). Variations in scan technique and criteria for biopsy and the inability to detect and sample small metastasis limits the reproducibility of US (7–12). Selected patients undergo breast MRI, which allows axillary assessment but has a mean sensitivity of only 60% (range, 33.3%–97%) and negative predictive value of 80% (range, 1.9%–99.5%) (13) in detecting pN stage. Hence, patients with benign findings from axillary imaging or biopsy routinely undergo surgical SLNB. SLNB has higher sensitivity compared with imaging, ranging from 86% to 92% (14), but it is a diagnostic procedure that involves morbidity associated with anesthesia, surgery, and radiation exposure because of radiopharmaceuticals used to identify the sentinel node (4,15,16). Permanent adverse effects of SLNB are reported as subclinical lymphedema in 24.4% of patients (17), clinical lymphedema (mean, 5.6% [range, 0%–11%]) (18), chronic axillary pain in 16%, sensory disorders including paresthesia (2%–22%), and limited arm motion (0%–9%) (19). Clinical predictions of node-positive status are primarily focused on nonsentinel node positivity and do not allow omission of SLNB (20,21). However, these methods do not address the overtreatment of a large proportion of patients (52%–82%) who

Abbreviations

AUC = area under the receiver operating characteristic curve, cN = clinical node, CNN = convolutional neural network, DCE = dynamic contrast-enhanced, ER = estrogen receptor, 4D = four-dimensional, HEBO = heteroscedastic evolutionary Bayesian optimization, HER2 = human epidermal growth factor receptor 2, pN = pathologic node, SLNB = sentinel lymph node biopsy, 3D = three-dimensional

Summary

A highly sensitive MRI-based four-dimensional convolutional neural network model showed high performance in predicting axillary lymph node metastasis in patients with breast cancer and may help reduce the complications of unnecessary interventional and surgical procedures.

Key Points

- An integrated clinical and breast MRI–based four-dimensional (4D) convolutional neural network model for predicting pathologic axillary node status showed high diagnostic performance, with area under the receiver operating characteristic curve (AUC), sensitivity, and specificity values of 0.87, 89%, and 76%, respectively.
- A machine learning model using clinicopathologic features alone demonstrated lower predictive performance of pathologic node status (AUC, sensitivity, specificity, and false-negative rate values of 0.63, 75%, 52%, and 65%, respectively)
- The 4D hybrid model showed similar diagnostic performance (P = .68) on two independent data sets from institutions with variable patient and cancer profiles; when trained on data from the safety-net hospital, the model achieved an AUC of 0.84 (95% CI: 0.75, 0.91), while when trained on university hospital data, the model achieved an AUC of 0.81 (95% CI: 0.73, 0.87).

Keywords

MR Imaging, Breast, Breast Cancer, Breast MRI, Machine Learning, Metastasis, Prognostic Prediction

undergo surgical diagnosis of the axilla (22,23). Therefore, a robust imaging-based predictive model that assesses the presence of lymph node metastasis with high sensitivity would help select patients who could potentially avoid SLNB and facilitate treatment decisions for patients with breast cancer, without the need for an invasive procedure. Identifying cN stage is important as cN+ status can be the determining factor in whether the patient will be treated with neoadjuvant therapy versus upfront surgery. Studies this decade in patients with early-stage breast cancer suggest that deep learning models trained with either US or MR images can yield acceptable diagnostic performance in predicting ipsilateral nodal metastasis (24,25). However, it is unclear whether these models developed based on lower-stage category T1 or T2 breast cancers would be applicable to all breast cancers. Furthermore, these models are trained with single-institutional and single–MRI unit data sets, precluding their uniform application (26–29). These models also suffer from assumptions inherent to black box optimization hyperparameter tuning tasks, which preclude their robust application to imaging data sets from other institutions.

In this study, we aimed to develop a deep learning algorithm to predict cN and pN stages regardless of tumor size by using breast MR image data sets from two different MRI units and institutions with racially diverse population data sets. Heteroscedastic evolutionary Bayesian optimization (HEBO) is advantageous over traditional Bayesian optimization and hyperband in hyperparameter optimization because of its greater adaptability to complex and varied problems, akin to how a radiologist tackles diverse cases. It efficiently explores hyperparameter spaces, making it quicker to find the optimal settings, much like efficiently diagnosing with limited images. Additionally, HEBO excels with limited data and handles uncertainty well, offering more personalized and flexible optimization, which is essential in real-world applications where data can be noisy or incomplete. This makes HEBO a more versatile and reliable choice in various scenarios, much like a radiologist's tailored approach to each patient. We aimed to determine whether our predictive model based on dynamic tumor MR images and selected clinicopathologic variables, which also integrates patch-based learning and hyperparameter optimization techniques (HEBO) that offset small-volume data, would yield high diagnostic performance in noninvasively diagnosing axillary metastasis, exceeding that of the currently available diagnostic imaging technologies used in clinical standard of care.

Materials and Methods

Patient Cohort

Written informed consent was waived for this institutional review board–approved, Health Insurance Portability and Accountability Act–compliant retrospective analysis. Consecutive patients with primary invasive breast cancer who underwent dynamic contrast-enhanced (DCE) breast MRI as part of their initial extent of disease evaluation between July 2013 and June 2016, performed in a university hospital and a safety-net hospital within the same health care network, were retrospectively reviewed. Inclusion and exclusion criteria for patients are summarized in Figure 1. Recurrent cancers and inflammatory (cT4b) breast cancers replacing entire breast tissue were excluded. Age, race, tumor size, clinical regional nodal metastasis (cN1–3), tumor stage, and neoadjuvant chemotherapy history were collected from electronic health records. Grade (Nottingham histologic), estrogen receptor (ER) and human epidermal growth factor receptor 2 (HER2) status (positive vs negative), and Ki-67 index were obtained from standardized College of American Pathologists pathology reports using previously published criteria (Fig S1).

DCE MRI Technique and Image Segmentation

All breast MRI examinations were performed with patients lying prone in 1.5-T scanners (Optima MR450w, GE HealthCare; or Intera, Philips Healthcare) using a dedicated eight-channel (Invivo Sentinelle; Siemens) or seven-channel (dSTream; Philips Healthcare) coil. A single precontrast and four serial bilateral axial dynamic image sets were obtained before and immediately after rapid intravenous bolus infusion of 0.1 mmol of gadopentetate dimeglumine contrast medium (Magnevist; Bayer Healthcare Pharmaceuticals) per kilogram of body weight at a rate of 3 mL/sec with a power injector

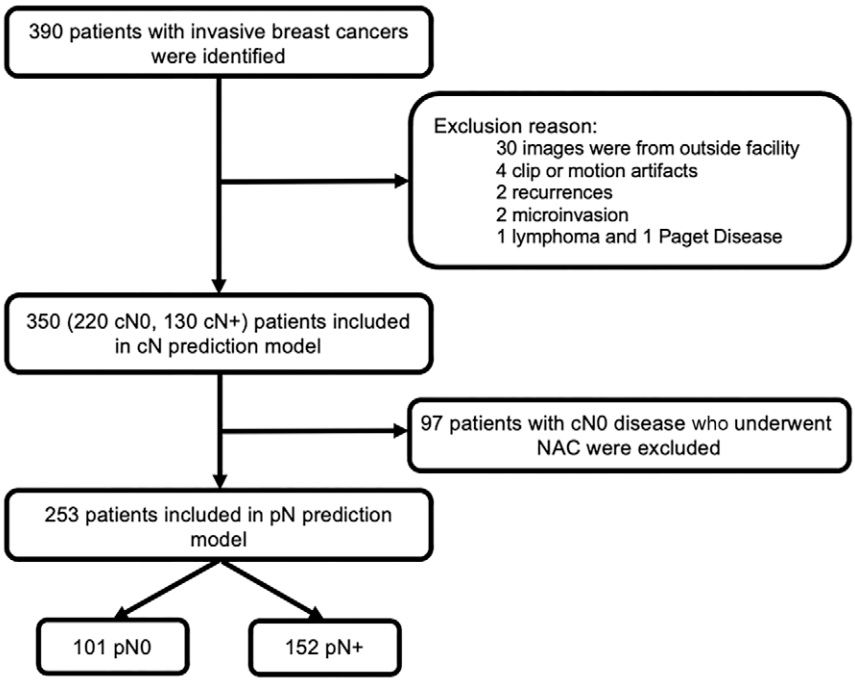

Figure 1: Consort flow diagram. cN = clinical nodal stage, NAC = neoadjuvant chemotherapy, pN = pathologic nodal stage.

(Spectris Solaris MR Injector; Medrad), with an average dynamic temporal resolution of 90 sec/phase. The primary tumor was delineated using Horos with an OsiriX plugin (Pixmeo) using images from the second postcontrast acquisition by a fellowship-trained breast radiologist (B.E.D.) with 15 years of experience in interpreting breast MRI studies. For multifocal or multicentric disease, the index lesion used to clinically stage the patient was delineated. As this exquisite approach to finely delineate tumors requires time and expertise, we derived crude bounding boxes using existing segmentation. Our model uses those bounding boxes, which can be drawn in less than a minute.

Staging the Axilla

The cN stage was determined using clinical and imaging assessment and imaging-guided core-needle biopsy in accordance with the American Joint Committee on Cancer staging system (30). The cN information from oncology treatment planning notes was collected. Patients staged as cN negative based on either clinical examination or a combination of MRI, axillary US, and US-guided needle biopsy findings were labeled as cN0, while patients with positive diagnosis confirmed with clinical examination and imaging-guided needle biopsy (cN1–3) were labeled cN+. The pN status was determined by first excluding all patients with cN0 status who underwent neoadjuvant chemotherapy prior to surgical axillary staging. Patients who had preoperative biopsy-verified axillary metastasis, as well as patients with cN0 status whose metastases were identified at SLNB, were labeled pN+. Patients with benign SLNB findings comprised the pN0 group.

Data Preprocessing, Model Construction, and Interinstitutional Model Performance Comparison Tests

All images were resampled to consistent 1-mm^3 isotropic resolution and harmonized using histogram equalization. A cuboidal box centered on the delineated tumor and encompassing peritumoral volume was defined to crop each image to consistently sized volumes. We used a cuboidal box of 120 × 120 × 120 voxels. The model utilized both tumor and axillary pixels. For the axillary pixels, we used the same cuboidal bounding box on the ipsilateral axilla to derive pixel information. Figure 2 provides a comprehensive overview of the data preprocessing steps. Please see Appendix S1 for further details on image intensity harmonization and voxel patching, which were used to facilitate deriving the pixels that maximally contributed to the model.

We hypothesized that the primary tumor and axilla image pixels would contain information associated with axilla metastasis status. To test this, we developed multiple convolutional neural network (CNN)–based models (31) that learn to recognize patterns in images. In the DCE MRI sequence, various temporal phases are designated as time 0 through time 5. Time 0 is identified as the phase prior to contrast enhancement, while time 1 to time 5 represent successive contrast-enhanced phases following the initial noncontrast sequence. Ablation experiments were conducted by testing a model using only clinical features, a three-dimensional (3D) model using only time 2 to time 0 volumes, an image-only model that utilizes all DCE time points (four-dimensional [4D] model), and last, a 4D hybrid model integrating difference volumes and clinicopathologic data. Diagnostic performance outcomes of each model were compared. Figure 3 presents a detailed flowchart illustrating the integration of data inputs across various models. Please see Appendix S1 for further model development and construction details.

Separate two-category classifiers were trained to predict cN0 versus cN+ and pN0 versus pN+ targets, and patient data were partitioned using nested fivefold cross-validation. Therefore, all data were partitioned into five groups, each including about 70 patients. Each fold had three training groups, one validation group, and one testing group, and one iteration was performed while rotating the test data set. Reported results are the average of the five folds. Please see Appendix S1 for further information on CNN model development.

Model performance was measured on the test data, not used for training or model selection. To gain further insight into the most important voxels, gradient class activation mapping (32) was used to generate voxel saliency maps.

To test the generalizability of our model, the 4D hybrid model for pN prediction was also trained from scratch using data from one institution at a time—university hospital versus safety-net

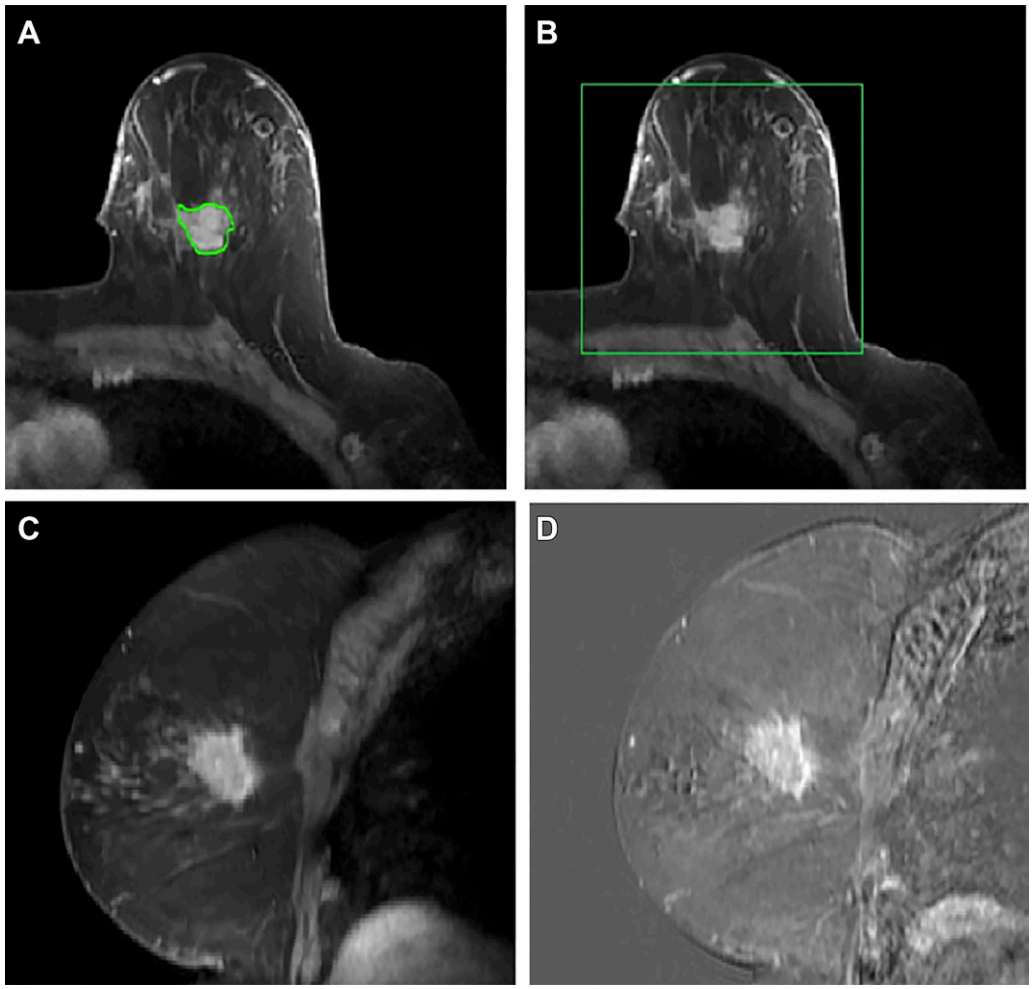

Figure 2: Preprocessing a volumetric dynamic contrast-enhanced MR image. **(A)** The primary tumor is radiologist delineated at time 3 in each section (green contour), though the proposed approach requires only a box around the tumor. **(B)** The MR image is cropped to a cuboidal volume around the tumor. **(C)** Sagittal view shows breast at time 1. **(D)** The tumor is enhanced by computing difference images, shown here: time 3–time 1.

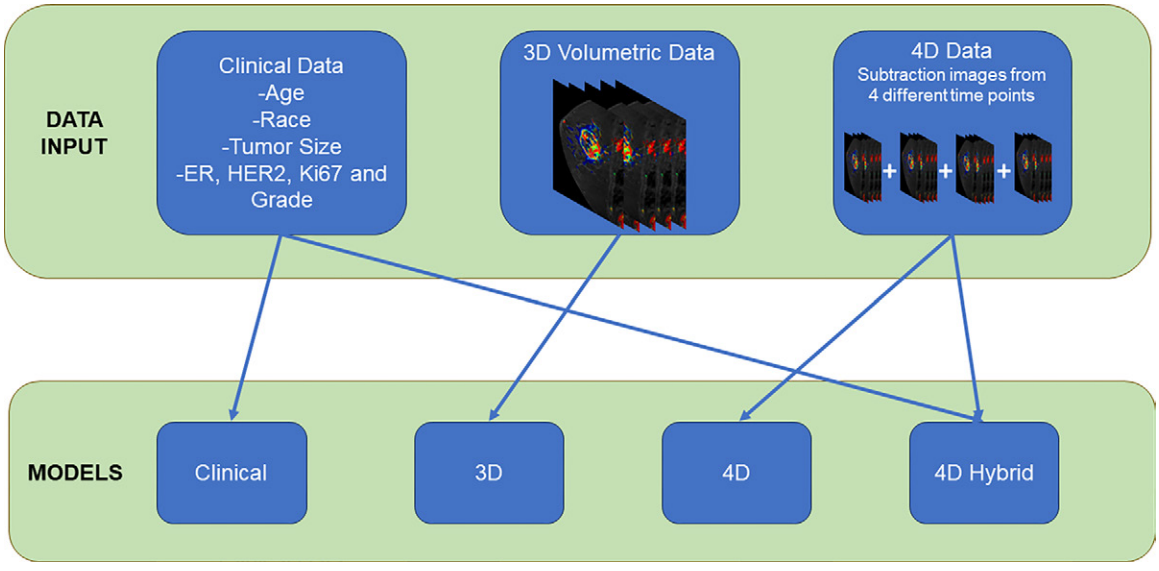

Figure 3: Flowchart depicts data input into the construction of the clinical, 3D volumetric imaging–, and 4D imaging–based models, as well as the hybrid 4D + clinical model. Of note, data used for model development were not used in the test data set. ER = estrogen receptor, 4D = four-dimensional, HER2 = human epidermal growth factor receptor 2, 3D = three-dimensional.

Table 1: Patients with Newly Diagnosed Invasive Breast Cancer Included for Development of Clinical and Pathologic Ipsilateral Nodal Metastasis Status Prediction Models

Characteristic	cN Status Cohort (*n* = 350)	pN Status Cohort (*n* = 253)
Institution		
University hospital	132 (37.7)	95 (37.5)
Safety-net hospital	218 (62.3)	158 (62.5)
Patient age (y)*	51.7 ± 11.9	52.7 ± 11.9
Grade		
1	66 (18.9)	57 (22.5)
2	140 (40.0)	106 (41.9)
3	144 (41.1)	90 (35.6)
Clinical T category		
T1	105 (30.0)	88 (34.8)
T2	153 (43.7)	100 (39.5)
T3	67 (19.1)	46 (18.2)
T4	25 (7.1)	19 (7.5)
Nodal stage		
N0	220 (62.9)	101 (39.9)
N+[†]	130 (37.1)	152 (60.1)
Median Ki-67 index (%)	31.0	30.0
Estrogen receptor		
Negative	95 (27.1)	54 (21.3)
Positive	255 (72.9)	199 (78.7)
HER2[‡]		
Negative	263 (75.4)	197 (78.2)
Positive	86 (24.6)	55 (21.8)
Patient race		
Asian	21 (6.0)	16 (6.3)
Black	79 (22.6)	57 (22.5)
Hispanic	123 (35.1)	88 (34.8)
Non-Hispanic White	127 (36.3)	92 (36.4)
Neoadjuvant chemotherapy[§]		
No	154 (45.0)	142 (56.1)
Yes	193 (55.0)	111 (43.9)

Note.—Unless otherwise indicated, values are numbers, with percentages in parentheses. All percentages are column based. cN = clinical node, HER2 = human epidermal growth factor receptor 2, pN = pathologic node.
* Values are means ± SDs.
[†] Patients with clinical N0 disease who underwent neoadjuvant therapy were excluded from pN prediction.
[‡] Two patients' results were equivocal after fluorescence in situ hybridization.
[§] Neoadjuvant chemotherapy information is missing for three patients.

hospital—and the model trained on one institution's data was tested on the other institution's for diagnostic performance.

Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics for Windows, version 27. As an initial analysis, ER status, HER2 status, tumor grade, tumor stage, and race were compared between cN0 versus cN+ as well as pN0 versus pN+ groups using Fisher exact test. Age was compared between groups using independent samples *t* tests and Ki-67 index using the Mann-Whitney *U* test. Normality for these variables was evaluated using Shapiro-Wilk test, Q-Q plot, and skewness-kurtosis, and a *P* value less than .05 was the significance criterion.

Model performance was measured via the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity. AUCs were compared using DeLong test (33). When selecting the sensitivity and specificity cutoff points, we prioritized sensitivity.

Results

Patient Characteristics

Of 390 patients with invasive breast cancer who underwent breast MRI, 350 female patients (mean age, 51.7 years ± 11.9 [SD]; range: 21–90 years) (Fig 1) with intact newly diagnosed breast cancer met our criteria for cN analysis. Nodal and cancer stage and clinicopathologic features are summarized in Table 1.

Of 350 patients, 97 (27.7%) who had cN0 disease, as determined with imaging and needle biopsy, who underwent preoperative neoadjuvant chemotherapy were excluded from pN prediction and sensitivity analysis. Of the remaining 253 patients, 152 had pathologic evidence of metastasis, of which 118 were diagnosed with imaging plus imaging-guided needle biopsy, yielding a true-positive rate of 77.6%. In the remaining 34 patients with negative imaging or imaging-guided needle biopsy findings, SLNB showed evidence of metastasis, yielding a false-negative rate of 22.4% (34 of 152).

Association of Clinicopathologic Factors with cN and pN Status

Correlation of clinicopathologic factors showed a significant correlation of T category with increasing prevalence of cN+ and pN+ status. Imaging sensitivity in detecting metastasis significantly increased with increasing tumor size (Table 2). The increase in sensitivity was especially significant between T1 and T2 cancers (27.8% vs 81.7%). T2 cancers were associated with significantly higher odds of cN+ disease (92% vs 7.5%, odds ratio: 13.6 [95% CI: 5.2, 35.4]; $P < .001$] and false-negative imaging findings (72.2% vs 18.3%) compared with T1 cancers (Table S3). Patients with cN+ disease were significantly younger (mean age, 40.0 vs 52.7 years; $P = .04$) than patients with cN0 disease. While the cN+ group had a higher rate of grade 3 ($P < .001$) and HER2-positive tumors (55.8% vs 44.2%, odds ratio: 2.8 [95% CI: 1.7, 4.6]; $P < .001$), they also had higher mean Ki-67 index (46.9 vs 37.4, $P < .001$). There was no evidence of

a significant association of race and ER positivity with cN status (Table 3).

We found similar associations with pN status. While patients with pN+ disease were younger (mean age, 50.0 vs 56.8 years; $P < .001$) than patients with pN0 disease, they also showed higher rates of T2–T4 tumors ($P < .001$). When only T1 and T2 tumors were compared, pN+ had higher rates of T2 cancers. HER2 positivity (83.6% vs 16.4%, odds ratio: 7.6 [95% CI: 3.6, 16.1]; $P < .001$) and mean Ki-67 index (43.2 vs 25.3, $P < .001$) were also higher in the pN+ group (Table 3). pN positivity correlated with increasing tumor grade (38.6% for grade 1, 55.7% for grade 2, and 78.9% for grade 3; $P < .001$). There was no evidence of a significant association between race and ER positivity with pN status.

Table 2: Clinical and Pathologic Nodal Stage Distribution Stratified by Clinical Tumor Stage

Tumor Stage and cN Status	pN Status		
	pN0	pN+	Total
T1			
cN0	70 (100)	13 (72.2)	83 (94.3)
cN+	0 (0)	5 (27.8)	5 (5.7)
Total	70	18	88
T2			
cN0	29 (100)	13 (18.3)	42 (42)
cN+	0 (0)	58 (81.7)	58 (58)
Total	29	71	100
T3			
cN0	2 (100)	7 (15.9)	9 (19.6)
cN+	0 (0)	37 (84.1)	37 (80.4)
Total	2	44	46
T4			
cN0	…	1 (5.3)	1 (5.3)
cN+	…	18 (94.7)	18 (94.7)
Total	…	19	19
T5			
cN0	101 (100)	34 (22.4)	135 (53.4)
cN+	0 (0)	118 (77.6)	118 (46.6
Total	101	152	253

Note.—Data are reported as numbers, with percentages in parentheses. All percentages are column based. cN = clinical node, pN = pathologic node.

Ablation Tests

Our machine learning model using clinicopathologic measures alone (ie, using age, race, ER status, HER2 status, Ki-67 index, and tumor grade) predicted cN status with an AUC, sensitivity, specificity, and false-negative rate of 0.55, 35%, 77%, and 65% and predicted pN status with values of 0.63, 75%, 52%, and 25%, respectively.

The imaging-based 3D model had AUC, sensitivity, specificity, and false-negative rate values of 0.67, 67%, 66%, and 33% for cN and 0.69, 77%, 53%, and 23% for pN, respectively. The 4D model that incorporated temporal data had values of 0.70, 81%, 41%, and 19% for cN and 0.71, 77%, 51%, and 23% for pN, respectively.

Performance of the Proposed Imaging Plus Clinicopathologic Hybrid 4D CNN Model

For the prediction of cN0 versus cN+ status, the 4D hybrid model yielded higher performance than the 3D model along the primary AUC performance measure with its AUC of 0.79 (95% CI: 0.76, 0.82) ($P = .004$) (Table 4). Additionally, the model yielded a high sensitivity of 80% (95% CI: 77%, 84%), specificity of 62% (95% CI: 58%, 67%), and a low false-negative rate of 20% (95% CI: 16%, 23%) (Table 4).

For the prediction of pN0 versus pN+ status, the 4D hybrid model yielded higher performance than the 3D model along the primary AUC performance measure with its AUC of 0.87 (95% CI: 0.83, 0.91) ($P < .001$) (Table 4). Additionally, the model produced a high sensitivity of 89% (95% CI: 79%, 93%), specificity of 76% (95% CI: 68%, 88%), and a low false-negative rate of 11% (7%, 21%) (Table 4). Table 5 summarizes the performance of the hybrid 4D model on the test data set at varying specificity cutoff points. At 71% specificity, the model yielded a sensitivity of 91% and a false-negative rate of 9%. At 51% specificity, sensitivity was 95% and the false-negative rate was 5%. Training and validation performances for the 4D hybrid model are shown in Table S4. The 4D hybrid model using tumor pixels alone had an AUC of 0.85, with sensitivity, specificity, and false-negative rates of 90% (95% CI: 85.5%, 94.7%), 64% (95% CI: 56.2%, 73%), and 9% (95% CI: 5.3%, 14.5%), respectively. Therefore, the use of axillary pixels improved model performance, though not significantly ($P = .12$) (Fig 4). Comparing the top-performing models that predict cN status to the model that predicts pN status, we find that for every specificity, the model predicting pN status attains higher sensitivity (Fig 5).

Of 17 patients with pN+ disease missed by the 4D hybrid model, 12 instances of pN+ disease (70.5%) were also missed by the radiologist, being identified as a single metastatic node at SLNB in patients with T1 ER-positive, HER2-negative cancers. The remaining five (30.5%) were identified with axillary imaging in patients with ER-positive ($n = 3$) or negative ($n = 2$) disease (cT1 = three and cT2 = two patients).

Generalizability of the Model: Interinstitutional Comparison

To test the generalizability of our model, the 4D hybrid model for pN prediction was also trained from scratch using data from one institution at a time: a university hospital versus a safety-net hospital. The AUC, sensitivity, and specificity values of the safety-net hospital–trained model tested on the university hospital were 0.84 (95% CI: 0.75, 0.91), 73% (95% CI: 58%, 86%), and 91% (95% CI: 78%, 98%); for the

Table 3: Correlation of Clinicopathologic Factors with Ipsilateral Clinical and Pathologic Nodal Status

	cN Status (n = 350)				pN Status (n = 253)			
Descriptive Parameter	Negative (n = 220)	Positive (n = 130)	Odds Ratio*	P Value	Negative (n = 101)	Positive (n = 152)	Odds Ratio*	P Value
Mean age (y)	52.7	50	…	.04	56.8	50	…	<.001
Race				.21				.05
Asian	17 (81)	4 (19)	…		9 (56.3)	7 (43.8)	…	
Black	50 (63.3)	29 (36.7)	…		28 (49.1)	29 (50.9)	…	
Hispanic	71 (57.7)	52 (42.3)	…		26 (29.5)	62 (70.5)	…	
Non-Hispanic White	82 (64.6)	45 (35.4)	…		38 (41.3)	54 (58.7)	…	
Clinical T category				<.001				<.001
T1	100 (95.2)	5 (4.8)	…		70 (79.5)	18 (20.5)	…	
T2	91 (59.5)	62 (40.5)	…		29 (29)	71 (71)	…	
T3	27 (40.3)	40 (59.7)	…		2 (4.3)	44 (95.7)	…	
T4	2 (8)	23 (92)	…		0 (0)	19 (100)	…	
ER			0.6 (0.4, 1)	.06			0.3 (0.1, 0.5)	.18
Negative	52 (54.7)	43 (45.3)	…		13 (24.1)	41 (75.9)	…	
Positive	168 (65.9)	87 (34.1)	…		88 (44.2)	111 (55.8)	…	
HER2†			2.8 (1.7, 4.6)	<.001			7.6 (3.6, 16.1)	<.001
Negative	181 (68.8)	82 (31.2)	…		91 (46.2)	106 (53.8)	…	
Positive	38 (44.2)	48 (55.8)	…		9 (16.4)	46 (83.6)	…	
Mean Ki-67 index	37.4	46.9	…	<.001	25.3	43.2	…	<.001
Grade				<.001				<.001
1	57 (86.4)	9 (13.6)	…		35 (61.4)	22 (38.6)	…	
2	88 (62.9)	52 (37.1)	…		47 (44.3)	59 (55.7)	…	
3	75 (52.1)	69 (47.9)	…		19 (21.1)	71 (78.9)	…	

Note.—Data are reported as numbers, with percentages in parentheses, unless otherwise specified. All percentages are row based. Fisher two-sided exact test was used to calculate P value for categorical variables. Independent samples t test was used for comparison between groups. Kruskal-Wallis test was used to compare Ki-67 index between groups. cN = clinical node, ER = estrogen receptor, HER2 = human epidermal growth factor receptor 2, pN = pathologic node.
* Values in parentheses are 95% CIs.
† Two patients' results were equivocal after fluorescence in situ hybridization and were excluded from the analysis.

Table 4: Performances of Clinical and Image-based Deep Learning Models in Predicting Ipsilateral Clinical and Pathologic Node Status

End Point	Input Data and CNN Model	AUC	Sensitivity (%)	Specificity (%)	False-Negative Rate (%)
cN0 vs cN+	Clinical	0.55 (0.54, 0.57)	35 (32, 38)	77 (75, 79)	65 (62, 68)
	3D	0.67 (0.60, 0.75)	67 (58, 77)	66 (58, 72)	33 (23, 42)
	4D	0.70 (0.67, 0.74)	81 (77, 86)	41 (36, 44)	19 (14, 23)
	4D hybrid	0.79 (0.76, 0.82)	80 (77, 84)	62 (58, 67)	20 (16, 23)
pN0 vs pN+	Clinical	0.63 (0.61, 0.65)	75 (73, 77)	52 (49, 55)	25 (23, 27)
	3D	0.69 (0.63, 0.74)	77 (71, 83)	53 (45, 61)	23 (17, 29)
	4D	0.71 (0.65, 0.76)	77 (71, 82)	51 (43, 59)	23 (18, 29)
	4D hybrid	0.87 (0.83, 0.91)	89 (79, 93)	76 (68, 88)	11 (7, 21)

Note.—Values in parentheses are 95% CIs. Outputs of 3D and 4D convolutional layers were combined with clinical data to generate hybrid models trained to differentiate cN0 versus cN+ and pN0 versus pN+ disease. Clinical data included age, race, estrogen receptor status, human epidermal growth factor receptor 2 status, Ki-67 index, tumor grade, and tumor stage. Tumor volumetric structure was created for the 3D model. The 4D CNN model included temporal data (three acquisition time points after contrast media injection). AUC = area under the receiver operating characteristic curve, cN = clinical nodal stage, CNN = convolutional neural network, 4D = four-dimensional, pN = pathologic nodal stage, 3D = three-dimensional.

Table 5: Hybrid Deep Learning and Clinical Model Performance to Predict Breast Cancer Nodal Metastatic Status at Various Specificity Cutoff Points

	70% Specificity		75% Specificity		80% Specificity	
Model	True Positive	True Negative	True Positive	True Negative	True Positive	True Negative
pN	137/149 (91)	70/99 (71)	132/149 (89)	75/99 (76)	118/149 (79)	79/99 (80)
cN	95/130 (73)	154/220 (70)	85/130 (65)	165/220 (75)	77/130 (59)	176/220 (80)

Note.—Data are reported as proportions of patients, with percentages in parentheses. cN = clinical nodal stage, pN = pathologic nodal stage.

university hospital–trained model tested on the safety-net hospital, the values were 0.81 (95% CI: 0.73, 0.87), 0.83% (95% CI: 0.6%, 0.97%), and 0.7% (95% CI: 0.49%, 0.9%), respectively.

Saliency Mapping

Our saliency map shows that the model learns important features from the primary tumor and peritumoral voxels to predict absence (Fig 6A) or high probability of metastasis (Fig 6B). More distal voxels are less important. As the algorithm did not explicitly provide the nonlinear tumor boundary, this saliency map result suggests the model learned to approximate the tumor boundary on its own.

Discussion

We developed a noninvasive imaging-based CNN model that helps determine axillary lymph node status in patients with breast cancer. Our MRI-based hybrid model showed AUC, sensitivity, and specificity values of 0.87, 89%, and 76%, respectively, which represents an improvement over the 77.6% radiologist sensitivity in our data set. One of the benefits of our model is its scalability. At a slightly lower specificity of 71%, our model yielded 91% sensitivity and a 9% false-negative rate, which align with the 90% sensitivity and less than 10% false-negative rate benchmarks used to define success for most SLNB studies, while maintaining higher diagnostic performances than axillary US, morphologic evaluation of DCE MRI, or other prognostics nomograms (34–36). The model using tumor pixels alone had comparable performance with our model (using both breast and axillary pixels), with a sensitivity of 90%, specificity of 64%, and false-negative rate of 9%. Hence, integration of our model

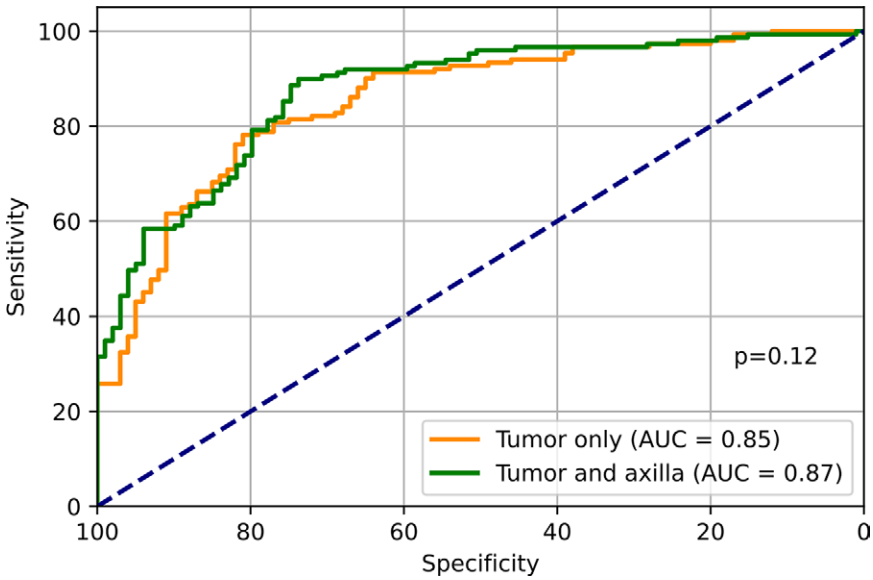

Figure 4: Diagnostic performance of four-dimensional hybrid model to predict ipsilateral metastatic versus benign clinical and pathologic lymph node status (positive vs negative). AUC = area under the receiver operating characteristic curve.

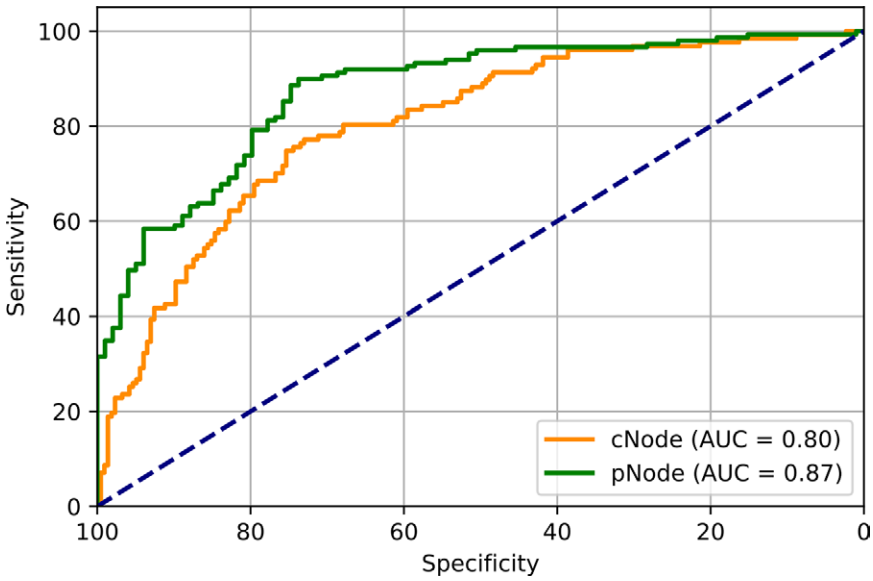

Figure 5: Diagnostic performance of two four-dimensional hybrid models using tumor plus axillary pixels versus tumor pixels alone in predicting pathologic node status (positive vs negative). *P* value was determined by DeLong test. AUC = area under the receiver operating characteristic curve, cNode = clinical node, pNode = pathologic node.

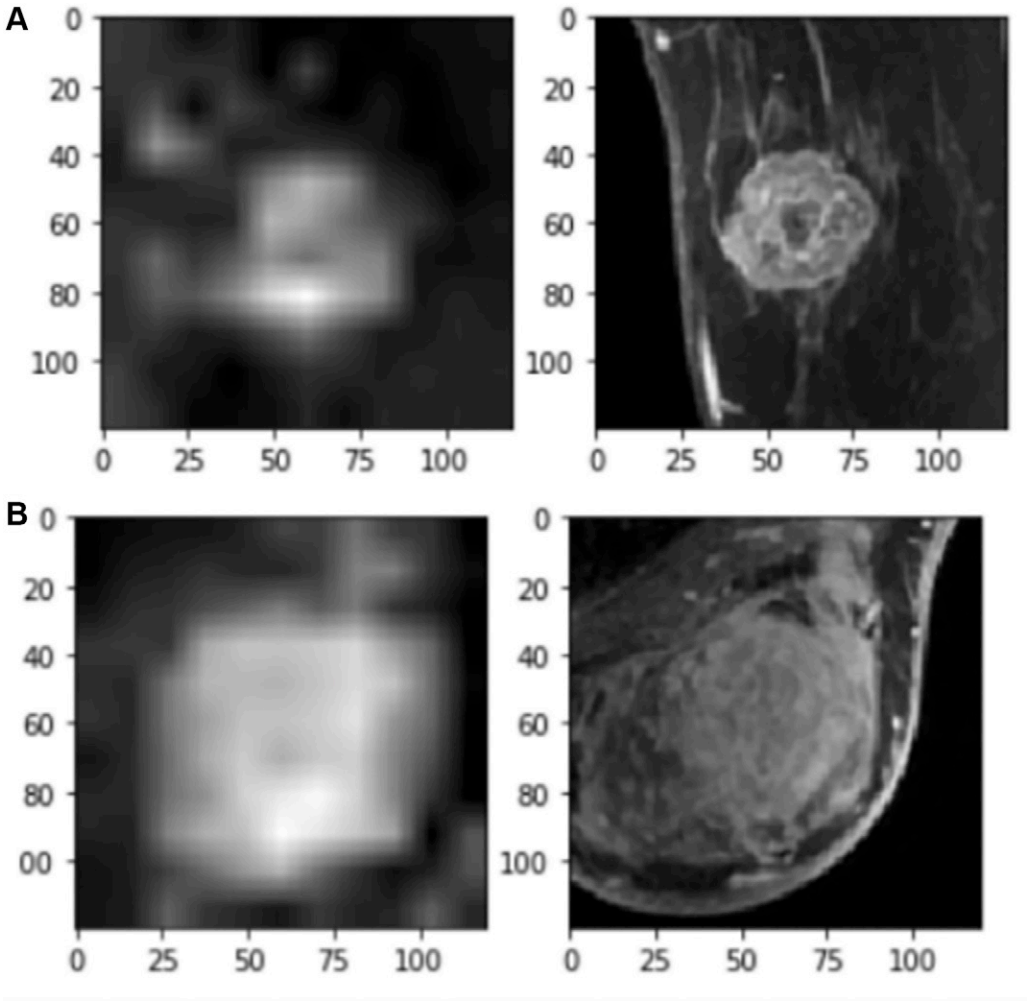

Figure 6: Saliency mapping. Important voxels are revealed through saliency mapping. Saliency mapping with gradient-weighted class activation mapping (left column) demonstrates that the primary active tumor voxels at MRI (right column) are those most valuable for prediction of both **(A)** nonmetastatic cancer and **(B)** axillary metastatic cancer.

into clinical practice can help triage a significant proportion of patients with low probability of metastasis with high sensitivity, mitigating additional diagnostic or invasive procedures and associated risk. Our model would have helped avoid more than half (51%) of benign sentinel node biopsies while correctly detecting 95% of patients with axillary metastasis. This is an important benefit, as some patients undergoing SLNB develop permanent complications, including lymphedema, seroma, wound infection, and pain, despite having a low probability of a positive result, which may substantially impact their management (37).

The diagnostic performance of our model represents an improvement over that of Arefan et al (26), which yielded sensitivity, specificity, and AUC of 82%, 78%, and 0.82, respectively. In a similar study that used DCE MR images from both tumor and axillary nodes (24,26) to develop a predictive clinical-radiomics algorithm for axillary metastasis detection, Yu et al (24) achieved an AUC of 0.9. However, the primary cancer pixel data alone yielded much lower diagnostic performance (AUC: 0.6). Segmentation and delineation of axillary nodes, as performed in their study, is a time-consuming process that is neither practical nor applicable in daily clinical practice. In contrast, our model yields a similar diagnostic performance but requires only a rough bounding box around the tumor without the need for precise tumor delineation. Adding the axillary image zone to our 4D hybrid model using tumor pixels alone improved its performance from 0.85 to 0.87, though not significantly (P = .12), with improved specificity (76% vs 64%) and no significant change in sensitivity (89% vs 90%), indicating that our model learns primarily from tumor images. This is an important advantage because the entire axilla may not consistently be included within the imaging field at standard breast MRI. Furthermore, this increases the day-to-day clinical applicability of our model by eliminating the need to segment out the axillary images, a process that involves extensive manual delineation.

In 2020, we showed that a 4D CNN can achieve high diagnostic accuracy to predict clinical lymph node stage (cN) across primary breast cancers (38). That preliminary study was performed using data obtained in a single hospital, with one MRI scanner and one MRI protocol. In our current work, we incorporate data from multiple hospitals, scanners, and protocols. The increased sample diversity likely improves our

model's generalizability to unseen data in the clinic. Additionally, in our current work, we introduced multiple architectural improvements to our algorithm, including a vision transformer–inspired patch-based approach (39) and improved hyperparameter optimization with HEBO (40) rather than traditional Bayesian optimization and hyperband (41). This improved model performance on cN prognostics and enabled us to develop the high-performing pN prognostic method that is unique to our current work.

In a recent study, Wang et al (42) used diffusion-weighted imaging and T1- and T2-weighted images to develop a CNN model, which yielded exceptional diagnostic performance (AUC: 0.996). In their study, the authors included "precursor lesions" that do not have metastatic potential, and unlike our study, they did not eliminate patients with cN0 disease who underwent neoadjuvant therapy. This is important, as existing axillary metastasis in these patients may have been eradicated over the course of treatment, making reference standard pathology unreliable. Additionally, their results are based on a single fold and thus could be inflated due to a fortuitous partition. While race information was not provided, with data originating from a single MRI unit and institution, lack of diversity may be pervasive. Additionally, no information on data preprocessing was provided.

Compared with previous models, our model has several advantages: *(a)* We partition the data into three parts (training, validation, and testing) through nested cross-validation, and the results we report pertain to the held-out test partition, which was not used during training and model selection. Reporting validation performance tends to overestimate the performance a model will achieve, while our more rigorous approach of reporting test performance provides a more realistic estimate of real-world clinical performance. *(b)* Our results are reported using nested fivefold cross-validation. As such, all data enter the test partition once. Compared with the single hold-out approaches used in the literature, our approach is not subject to a favorable or unfavorable partitioning result. *(c)* The deep learning approach we employ is unbiased by manual feature selection and instead directly learns optimal features for the imaging data, which is likely why our model does better on the tumor-only pixels than the model developed by Yu et al (24). Our model also compares well to published results using US, where original research (25) has revealed that a model trained on US images can yield AUC, sensitivity, and specificity values of 0.89, 85%, and 73%, respectively. However, the footprint of a US breast transducer is only approximately 4 cm, making the training of such models based on US images problematic for large tumors. Our results suggest that a machine learning model using DCE MRI to inform axilla involvement may be a safe and efficient tool for evaluating breast cancer, obviating further axillary imaging or diagnostic/surgical intervention.

We observed that using a sequence of 3D difference images rather than a single 3D difference image improved the AUC of the cN prediction (cN0 vs cN+) from 0.67 to 0.70 and improved pN prediction by a similar amount (from 0.69 to 0.71). Hence, the majority of the information learned by the model is contained in the spatial pattern of intensity differences (time 2 – time 0) and further contained in the combination of the clinical and imaging data, which bolsters performance to 0.79 and 0.87 AUC for cN and pN predictions, respectively.

Our data set had a well-balanced representation of patient race, with 36.3% of our patients being non-Hispanic White, 35% Hispanic, and 26% Black, which is an important consideration while developing CNN models applicable to patients of varying races and ethnicities. While predicting on more ethnically and potentially genetically heterogeneous data sets is more challenging, it is well representative of real-world performance in a heterogeneous population.

Our study had some limitations. The main limitation toward clinical adoption of our model is the current requirement for a manual coarse bounding box around the tumor area by the radiologist. Development and implementation of a fully automatic machine learning–based bounding box tool can allow time-efficient analysis of larger data sets. Another potential limitation was that currently, our data source consists of two affiliated hospitals using two different vendors. While overall model performance on the university hospital versus safety-net hospital were similar, the model trained on the safety-net hospital and tested on the university hospital had higher AUC and specificity compared with the model trained on the university hospital and tested on safety-net hospital data (AUC 0.84 vs 0.81, sensitivity 73% vs 83%, specificity 91% vs 71%). This illustrates the importance of incorporating imaging data from more institutions and different types of MRI units to further increase the variability of our data and enhance the generalizability of our results. In addition, the utilized scans originated from two 1.5-T MRI units, which are used in routine breast MRI in our institution; this may limit the applicability of our data to patients scanned at 3 T. As expected, the prevalence of pN2 and pN3 tumors was much lower than that of N0 and N1, as we excluded patients who underwent neoadjuvant chemotherapy. Hence, we used a binary classification, N0 versus N+ (includes N1, N2, and N3), to prevent any bias and skewing. Additionally, while we propose our best-performing model for further analysis, an independent prospective external validation would be welcome. Last, routine use of breast MRI in patients with breast cancer is still under debate. Patients with dense breasts, hormone receptor–negative disease, or unfavorable tumor subtypes eligible for neoadjuvant therapy may benefit more from the proposed model due to the higher significance and potential impact of axillary metastasis on their care.

In this study, we developed and validated a hybrid clinical and 4D MRI-based model that provides individualized prediction of axillary metastasis without the need for dedicated axillary imaging or invasive procedures. Our model is a safe and time-efficient tool, achieving noteworthy results compared with the existing methods. In the future, we look forward to additional optimization of the image analysis process and the inclusion of more variable data to further enhance the utility of the proposed models.

Author contributions: Guarantors of integrity of entire study, **D.S.P., S.N., A.M., B.E.D.**; 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, **D.S.P., S.N., P.K., M.C.C., L.W., A.M., B.E.D.**; clinical studies, **D.S.P., S.N., P.K., M.C.C., L.W., A.M., B.E.D.**; statistical analysis, **D.S.P., S.N., P.K., M.C.C., A.M.**; and manuscript editing, **D.S.P., S.N., P.K., M.C.C., A.M., B.E.D.**

Disclosures of conflicts of interest: D.S.P. No relevant relationships. **S.N.** No relevant relationships. **P.K.** No relevant relationships. **K.H.** No relevant relationships. **M.C.C.** No relevant relationships. **L.W.** No relevant relationships. **A.M.** National Institutes of Health/National Institute of General Medical Sciences grant number R01GM144486; grants from the Lyda Hill Foundation and the King Foundation. **B.E.D.** Grant support from Seno Medical and Med Cognetics, author was the study principal investigator, and author's institution received research support for clinical trial; royalties from Elsevier for authoring a book chapter; payment or honoraria for services as Grand Rounds speaker for MGH and Brigham Women's Hospital Departmental Grand Rounds (2022) and Mount Sinai Department of Radiology Grand Rounds (2022) and as grant reviewer for the New York State Department of Health Peter T. Rowley Breast Cancer Scientific Research Grant.

References

1. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. CA Cancer J Clin 2019;69(6):438–451.
2. Dillekås H, Rogers MS, Straume O. Are 90% of deaths from cancer caused by metastases? Cancer Med 2019;8(12):5574–5576.
3. Orsaria P, Varvaras D, Vanni G, et al. Nodal status assessment in breast cancer: strategies of clinical grounds and quality of life implications. Int J Breast Cancer 2014;2014:469803.
4. Gipponi M, Bassetti C, Canavese G, et al. Sentinel lymph node as a new marker for therapeutic planning in breast cancer patients. J Surg Oncol 2004;85(3):102–111.
5. Houssami N, Diepstraten SC, Cody HS 3rd, Turner RM, Sever AR. Clinical utility of ultrasound-needle biopsy for preoperative staging of the axilla in invasive breast cancer. Anticancer Res 2014;34(3):1087–1097.
6. Diepstraten SC, Sever AR, Buckens CF, et al. Value of preoperative ultrasound-guided axillary lymph node biopsy for preventing completion axillary lymph node dissection in breast cancer: a systematic review and meta-analysis. Ann Surg Oncol 2014;21(1):51–59.
7. Abe H, Schacht D, Sennett CA, Newstead GM, Schmidt RA. Utility of preoperative ultrasound for predicting pN2 or higher stage axillary lymph node involvement in patients with newly diagnosed breast cancer. AJR Am J Roentgenol 2013;200(3):696–702.
8. Bedrosian I, Mick R, Orel SG, et al. Changes in the surgical management of patients with breast carcinoma based on preoperative magnetic resonance imaging. Cancer 2003;98(3):468–473.
9. Iwamoto N, Aruga T, Horiguchi S, et al. Ultrasound-guided fine-needle aspiration of axillary lymph nodes in breast cancer: Diagnostic accuracy and role in surgical management. Diagn Cytopathol 2019;47(8):788–792.
10. Krishnamurthy S, Sneige N, Bedi DG, et al. Role of ultrasound-guided fine-needle aspiration of indeterminate and suspicious axillary lymph nodes in the initial staging of breast carcinoma. Cancer 2002;95(5):982–988.
11. Podkrajsek M, Music MM, Kadivec M, et al. Role of ultrasound in the preoperative staging of patients with breast cancer. Eur Radiol 2005;15(5):1044–1050.
12. Rajesh YS, Ellenbogen S, Banerjee B. Preoperative axillary ultrasound scan: its accuracy in assessing the axillary nodal status in carcinoma breast. Breast 2002;11(1):49–52.
13. Kuijs VJL, Moossdorff M, Schipper RJ, et al. The role of MRI in axillary lymph node imaging in breast cancer patients: a systematic review. Insights Imaging 2015;6(2):203–215.
14. Schuman S, Walker G, Avisar E. Processing sentinel nodes in breast cancer: when and how many? Arch Surg 2011;146(4):389–393.
15. Kataria K, Srivastava A, Qaiser D. What is a false negative sentinel node biopsy: definition, reasons and ways to minimize it? Indian J Surg 2016;78(5):396–401.
16. Pesek S, Ashikaga T, Krag LE, Krag D. The false-negative rate of sentinel node biopsy in patients with breast cancer: a meta-analysis. World J Surg 2012;36(9):2239–2251.
17. Bucci LK, Brunelle CL, Bernstein MC, et al. Subclinical lymphedema after treatment for breast cancer: risk of progression and considerations for early intervention. Ann Surg Oncol 2021;28(13):8624–8633.
18. Cortina CS, Yen TWF, Bergom C, et al. Breast cancer-related lymphedema rates after modern axillary treatments: How accurate are our estimates? Surgery 2022;171(3):682–686.
19. Pilger TL, Francisco DF, Candido Dos Reis FJ. Effect of sentinel lymph node biopsy on upper limb function in women with early breast cancer: A systematic review of clinical trials. Eur J Surg Oncol 2021;47(7):1497–1506.
20. Specht MC, Kattan MW, Gonen M, Fey J, Van Zee KJ. Predicting non-sentinel node status after positive sentinel lymph biopsy for breast cancer: clinicians versus nomogram. Ann Surg Oncol 2005;12(8):654–659.
21. van la Parra RF, Ernst MF, Bevilacqua JL, et al. Validation of a nomogram to predict the risk of nonsentinel lymph node metastases in breast cancer patients with a positive sentinel node biopsy: validation of the MSKCC breast nomogram. Ann Surg Oncol 2009;16(5):1128–1135.
22. Zhang J, Wang T, Yan C, Huang M, Fan Z, Ling R. Clinical practice status of sentinel lymph node biopsy for early-stage breast cancer patients in china: a multicenter study. Clin Epidemiol 2020;12:917–924.
23. Setton J, Cody H, Tan L, et al. Radiation field design and regional control in sentinel lymph node-positive breast cancer patients with omission of axillary dissection. Cancer 2012;118(8):1994–2003.
24. Yu Y, Tan Y, Xie C, et al. Development and validation of a preoperative magnetic resonance imaging radiomics-based signature to predict axillary lymph node metastasis and disease-free survival in patients with early-stage breast cancer. JAMA Netw Open 2020;3(12):e2028086.
25. Zhou LQ, Wu XL, Huang SY, et al. Lymph node metastasis prediction from primary breast cancer US images using deep learning. Radiology 2020;294(1):19–28.
26. Arefan D, Chai R, Sun M, Zuley ML, Wu S. Machine learning prediction of axillary lymph node metastasis in breast cancer: 2D versus 3D radiomic features. Med Phys 2020;47(12):6334–6342.
27. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nat Commun 2020;11(1):1236. [Published correction appears in Nat Commun 2021;12(1):4370.]
28. Madekivi V, Boström P, Karlsson A, Aaltonen R, Salminen E. Can a machine-learning model improve the prediction of nodal stage after a positive sentinel lymph node biopsy in breast cancer? Acta Oncol 2020;59(6):689–695.
29. Fanizzi A, Pomarico D, Paradiso A, et al. Predicting of sentinel lymph node status in breast cancer patients with clinically negative nodes: a validation study. Cancers (Basel) 2021;13(2):352.
30. Amin MB, Edge S, Greene F, et al. AJCC Cancer Staging Manual. 8th ed. Springer, 2017.
31. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–444.
32. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017; 618–626.
33. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.
34. Wang Z, Wu LC, Chen JQ. Sentinel lymph node biopsy compared with axillary lymph node dissection in early breast cancer: a meta-analysis. Breast Cancer Res Treat 2011;129(3):675–689.
35. Giuliano AE, Ballman KV, McCall L, et al. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: The ACOSOG Z0011 (Alliance) Randomized Clinical Trial. JAMA 2017;318(10):918–926.
36. Bevilacqua JL, Kattan MW, Fey JV, Cody HS 3rd, Borgen PI, Van Zee KJ. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. J Clin Oncol 2007;25(24):3670–3679.
37. Killelea BK, Long JB, Dang W, et al. Associations between sentinel lymph node biopsy and complications for patients with ductal carcinoma in situ. Ann Surg Oncol 2018;25(6):1521–1529.
38. Nguyen S, Polat D, Karbasi P, et al. Preoperative prediction of lymph node metastasis from clinical DCE MRI of the primary breast tumor using a 4D CNN. In: Martel AL, Abolmaesumi P, Stoyanov D, et al, eds. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. MICCAI 2020. Lecture Notes in Computer Science, vol 12262. Springer, 2020; 326–334.
39. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale. arXiv 2010.11929 [preprint] https://arxiv.org/abs/2010.11929. Posted October 22, 2020. Accessed May 17, 2023.
40. Cowen-Rivers AI, Lyu W, Tutunov R, et al. An empirical study of assumptions in bayesian optimisation. arXiv 2012.03826v4 [preprint] https://arxiv.org/abs/2012.03826v4. Posted December 7, 2020. Accessed May 17, 2023.
41. Falkner S, Klein A, Hutter F. BOHB: robust and efficient hyperparameter optimization at scale. arXiv 1807.01774 [preprint] https://arxiv.org/abs/1807.01774. Posted July 4, 2018. Accessed July 1, 2018.
42. Wang Z, Sun H, Li J, et al. Preoperative prediction of axillary lymph node metastasis in breast cancer using CNN based on multiparametric MRI. J Magn Reson Imaging 2022;56(3):700–709.