https://doi.org/10.1007/s00330-023-10256-4

EUROPEAN SOCIETY OF RADIOLOGY
European Radiology

BREAST

Radiomics in cone-beam breast CT for the prediction of axillary lymph node metastasis in breast cancer: a multi-center multi-device study

Yueqiang Zhu[1,2], Yue Ma[1], Zhenzhen Zhai[3], Aidi Liu[1], Yafei Wang[1], Yuwei Zhang[1], Haijie Li[1], Mengran Zhao[1], Peng Han[1], Lu Yin[1], Ni He[4], Yaopan Wu[4], Ioannis Sechopoulos[2,5,6], Zhaoxiang Ye[1*] and Marco Caballo[2]

Abstract

Objectives To develop a radiomics model in contrast-enhanced cone-beam breast CT (CE-CBBCT) for preoperative prediction of axillary lymph node (ALN) status and metastatic burden of breast cancer.

Methods Two hundred and seventy-four patients who underwent CE-CBBCT examination with two scanners between 2012 and 2021 from two institutions were enrolled. The primary tumor was annotated in each patient image, from which 1781 radiomics features were extracted with PyRadiomics. After feature selection, support vector machine models were developed to predict ALN status and metastatic burden. To avoid overfitting on a specific patient subset, 100 randomly stratified splits were made to assign the patients to either training/fine-tuning or test set. Area under the receiver operating characteristic curve (AUC) of these radiomics models was compared to those obtained when training the models only with clinical features and combined clinical-radiomics descriptors. Ground truth was established by histopathology.

Results One hundred and eighteen patients had ALN metastasis ($N+(\geq 1)$). Of these, 74 had low burden ($N+(1–2)$) and 44 high burden ($N+(\geq 3)$). The remaining 156 patients had none (N0). AUC values across the 100 test repeats in predicting ALN status (N0/$N+(\geq 1)$) were 0.75 ± 0.05 (0.67–0.93, radiomics model), 0.68 ± 0.07 (0.53–0.85, clinical model), and 0.74 ± 0.05 (0.67–0.88, combined model). For metastatic burden prediction ($N+(1–2)$/$N+(\geq 3)$), AUC values were 0.65 ± 0.10 (0.50–0.88, radiomics model), 0.55 ± 0.10 (0.40–0.80, clinical model), and 0.64 ± 0.09 (0.50–0.90, combined model), with all the ranges spanning 0.5. In both cases, the radiomics model was significantly better than the clinical model (both $p < 0.01$) and comparable with the combined model ($p = 0.56$ and 0.64).

Conclusions Radiomics features of primary tumors could have potential in predicting ALN metastasis in CE-CBBCT imaging.

Clinical relevance statement The findings support potential clinical use of radiomics for predicting axillary lymph node metastasis in breast cancer patients and addressing the limited axilla coverage of cone-beam breast CT.

Key Points

- *Contrast-enhanced cone-beam breast CT-based radiomics could have potential to predict N0 vs. N + (≥ 1) and, to a limited extent, N + (1–2) vs. N + (≥ 3) from primary tumor, and this could help address the limited axilla coverage, pending future verifications on larger cohorts.*

*Correspondence:
Zhaoxiang Ye
yezhaoxiang@163.com
Full list of author information is available at the end of the article

- *The average AUC of radiomics and combined models was significantly higher than that of clinical models but showed no significant difference between themselves.*
- *Radiomics features descriptive of tumor texture were found informative on axillary lymph node status, highlighting a higher heterogeneity for tumor with positive axillary lymph node.*

Keywords Breast neoplasms, Lymphatic metastasis, Machine learning, Cone-beam computed tomography, Radiomics

Introduction

Breast cancer is currently the most common malignancy and the leading cancer-related cause of death in females worldwide [1]. Patients with positive axillary lymph node (ALN) status are considered at high risk and need to undergo adjuvant chemotherapy generally according to NCCN guidelines [2]. Moreover, the American College of Surgeons Oncology Group (ACOSOG) Z0011 trial demonstrated that among patients of clinical T1/T2 breast cancer with one or two metastatic sentinel LNs (i.e., low metastatic burden) who accept breast-conserving surgery and systemic therapy, the use of sentinel LN biopsy (SLNB) alone would not lead to inferior survival compared to ALN dissection (ALND) [3]. Therefore, ALN status and metastatic burden are decisive factors for prognosis and therapeutic decision-making for breast cancer patients. SLNB is now a standard procedure and has less severe complications compared to ALND. However, SLNB still has limitations, such as prolonged anesthesia time due to the long intraoperative pathology waiting time, and potential complications to the upper limb (numbness, paresthesia, and lymphedema) [4]. Hence, it is desirable to develop a non-invasive and reliable preoperative approach to identify ALN status and metastatic burden in patients with breast cancer.

Contrast-enhanced cone-beam breast CT (CE-CBBCT), which has shown higher comfort level and diagnostic efficiency than mammography [5, 6] and faster acquisition and comparability in diagnostic performance to CE-MRI [6, 7], is a burgeoning technique whose role is starting to be recognized in clinical practice [8]. It has shown promising results in a variety of clinical settings, such as lesion detection and diagnosis [9–13], molecular subtyping [14, 15], extent-of-disease evaluation [16, 17], and breast density assessment [18, 19]. However, axilla coverage is an inherent limitation of CBBCT due to current table and acquisition geometry designs [7, 20]. Therefore, since the axillary area cannot be imaged in its entirety, the status of ALN cannot be assessed directly through visual assessment.

Radiomics, the process of converting medical images into high-dimensional, mineable, and quantitative imaging features via high-throughput data extraction algorithms, possibly improves the accuracy of diagnostic, predictive, and prognostic models upon simple visual interpretation [21]. Although ALNs cannot be entirely imaged in CBBCT, quantitative image information extracted from the primary tumor through radiomics algorithms might yield information on ALN status, as already suggested by previous studies performed with mammography, ultrasound, and MRI. Yang et al [22] developed a radiomics model based on mammography to preoperatively evaluate ALN status, reporting an area under the receiver operating characteristic (ROC) curve (AUC) of 0.88 in the validation cohort. Yu et al [23] constructed a radiomics model for predicting ALN metastasis based on ultrasound, reaching an AUC of 0.71. Dong et al [24] reported an AUC of 0.81 in predicting ALN metastasis based on T2WI and DWI MRI imaging. CE-CBBCT, thanks to its high resolution and full three-dimensional nature, might result in improved, or at least complementary, characterization capabilities compared to other modalities, especially when radiomics methods are leveraged [25–27].

At present, to the best of our knowledge, there is no study published regarding CBBCT-based radiomics for preoperative prediction of ALN status and metastatic burden in breast cancer. Hence, the purpose of this study was to determine the performance of CBBCT radiomics for these tasks, leveraging quantitative imaging biomarkers extracted from the primary tumor, and therefore evaluate if radiomics could help supplement the limitation of the limited axilla coverage in CBBCT.

Materials and methods

Study population

This multicenter retrospective study, conducted following the Declaration of Helsinki, was based on two previously performed prospective clinical trials (NCT01792999 and NCT03861221) approved by the ethics committee of Tianjin Medical University Cancer Institute and Hospital (TCIH) (E2012036 and bc2016039) and Sun Yat-Sen University Cancer Center (SYSUCC) (A2011-030-01). During these trials, written informed consent including permission to re-use the data for any further retrospective analysis was obtained from every patient at the time of enrollment.

CE scan was performed for patients with BI-RADS 3~5 lesions identified via mammography or ultrasound and/or with dense or very dense breast tissue (ACR density type c or d) for the purpose of differential diagnosis, preoperative staging, and treatment monitoring, unless there were known contraindications to iodinated contrast media. Only non-contrast CBBCT was offered as an alternative to mammography in the screening setting.

A total of 1784 patients who underwent CBBCT from May 2012 to June 2021 were initially collected. One hundred and fifty-eight patients with benign lesions confirmed by pathology/follow-up and 919 screening participants without a lesion were excluded, and 707 patients with histologically confirmed breast cancer remained. The exclusion criteria were as follows: (1) only conducted plain scan, (2) interval between CBBCT examination and subsequent surgery exceeded 1 month, (3) received neoadjuvant therapy or biopsy before CBBCT examination, (4) received preoperative therapy, (5) mass lesion with diameter larger than 5 cm (stage T3), (6) only manifested as non-mass enhancement (NME) on CE-CBBCT images, (7) poor image quality, (8) incomplete clinicopathologic data. The patient recruitment flowchart is shown in Fig. 1. The largest lesion was selected for further analysis if multiple lesions or bilateral disease were found, following previously published studies [28, 29]. Finally, data from a total of 274 patients with 274 breast tumors who met the criteria were retrieved. For this study, patients were categorized into two study cohorts: Study 1—disease-free axilla (N0) vs. any axillary metastasis (N + (≥ 1)); Study 2—low metastatic burden (N + (1–2)) vs. high metastatic burden (N + (≥ 3)). The clinicopathologic data were derived from the medical records. The whole pipeline of our study is shown in Fig. 2.

Pathological assessment of ALN

All patients underwent axillary ultrasound first. If patients showed negative axilla at ultrasound, SLNB was performed by the dye method at surgery. If the SLN was positive, and ACOSOG Z0011 criteria were not met (i.e., stage T3/T4, N + (≥ 3), mastectomy), ALND was conducted. If patients were found to have suspicious positive ALN at ultrasound, an ultrasound-guided fine-needle biopsy (FNB) was performed. If there was a histologically positive ALN on FNB, the patient received ALND. If FNB was negative, the patient received SLNB. ALN status was confirmed by histopathology. Isolated tumor cells (≤ 0.2 mm or deposits ≤ 200 cells) were defined as negative, whereas micrometastasis (0.2–2.0 mm or deposits > 200 cells) and macrometastasis (> 2.0 mm) were both

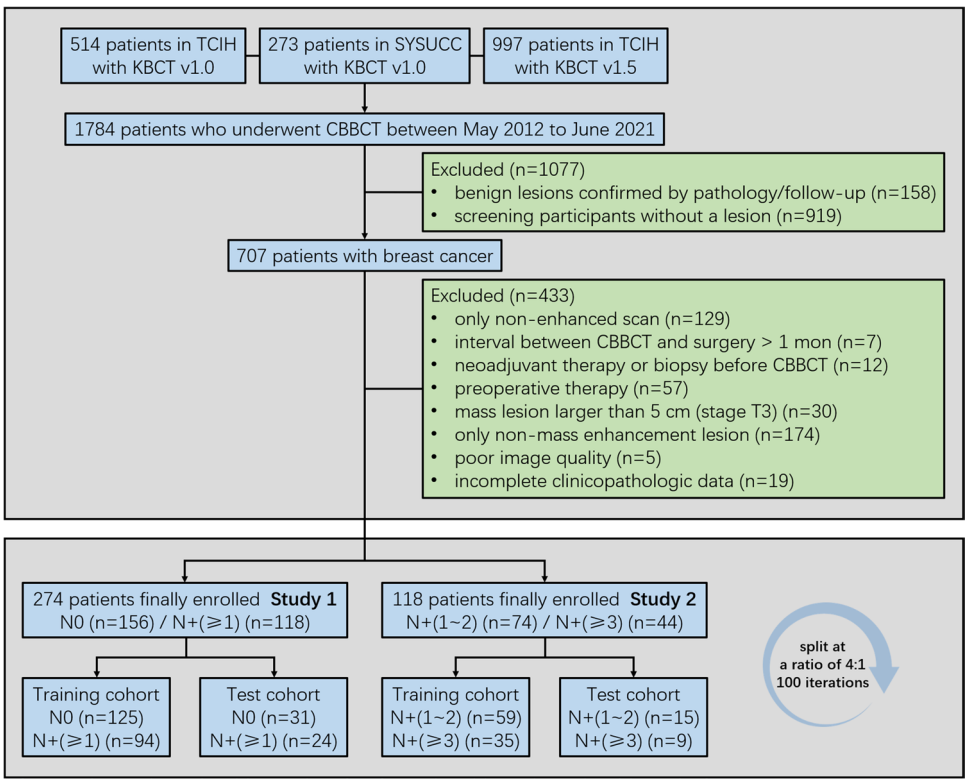

Fig. 1 Flowchart of patient recruitment and study design. *TCIH*, Tianjin Medical University Cancer Institute and Hospital; *SYSUCC*, Sun Yat-Sen University Cancer Center; *KBCT*, Koning breast CT; *CBBCT*, cone-beam breast CT

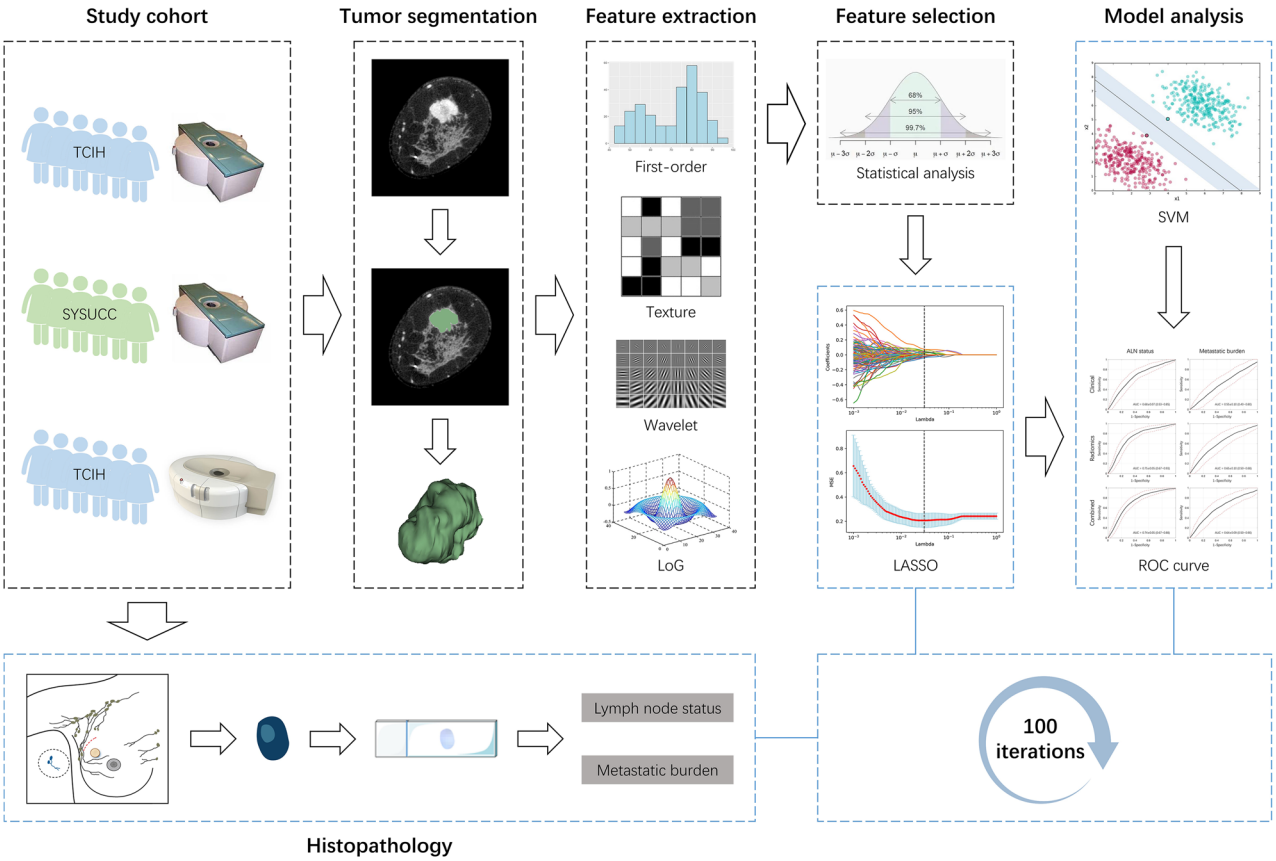

Fig. 2 Schematics of study pipeline. *TCIH*, Tianjin Medical University Cancer Institute and Hospital; *SYSUCC*, Sun Yat-Sen University Cancer Center; *LoG*, Laplace of Gaussian; *LASSO*, least absolute shrinkage selection operator; *SVM*, Support vector machine; *ROC*, receiver operating characteristic

defined as positive [30]. The results were confirmed by two pathologists with more than 10 years of experience.

Imaging system and scan protocol

All patients of the three datasets from two institutions preoperatively received CE-CBBCT examination using a dedicated flat-panel detector breast CT (KBCT-1000, Koning Corporation). Images were obtained from different generation CBBCT scanners (version 1.0: Dataset 1&2; version 1.5: Dataset 3; specific parameters are summarized in Supplementary Material Table S1). The examination was performed with the patient lying prone on the exam table. The imaged breast was suspended through the table opening into the imaging field without compression. After an initial plain scan, 90 ml (Dataset 1&2)/100 ml (Dataset 3) non-ionic contrast media (Iohexol, Omnipaque® 300, GE Healthcare--Dataset 1&2/Iodixanol, Visipaque® 270, GE Healthcare--Dataset 3) was intravenously injected at a rate of 2 ml/s (Dataset 1&2)/2.5 ml/s (Dataset 3) using a power injector, followed by a 30 ml saline solution chaser (Dataset 1&2)/no saline flush followed (Dataset 3) [31, 32]. Both breasts were scanned alternately with the affected breast first after contrast media application. The CE images of the affected breast were obtained exactly at 120 s after the start of contrast injection, and the contralateral breast was imaged at approximately 180 s post-injection, depending on the time of repositioning. The overall examination time was 8–10 min. The mean glandular dose for a complete CE-CBBCT scan was 11.46–14.68 mGy [14].

Tumor segmentation and feature extraction

Initially, the volumes of interest (VOIs) of all primary tumors were semi-automatically delineated in 3D on each slice of the CE-CBBCT images by a breast radiologist with a 10-year CBBCT experience (Y.Z.), blind to any pathologic outcomes on tumor and ALN, using 3D Slicer software (v.4.11.20210226). For 30 randomly selected tumors (from as many patient images), the segmentation was repeated by another breast radiologist with a 5-year CBBCT experience (Y.M.) also blind to outcomes, with the same software. Next, radiomics features were extracted from each tumor using the SlicerRadiomics extension of the 3D Slicer software, the in-house feature

extraction platform developed based on the Python package "PyRadiomics" [33]. A total of 1781 radiomics features were extracted from each VOI using the standard software settings. The features included 107 original features (14 shape, 18 first-order, 75 texture) and 1674 filter transformed features (744 wavelet, 465 Laplacian of Gaussian (LoG) with kernel size of 1–5 voxels, 93 square, 93 square root, 93 logarithm, 93 exponential, and 93 gradient). The involved features and their definitions are listed in the Supplementary Material. The 30 tumors annotated by the two radiologists were used to assess the robustness of radiomics features to variations in tumor annotations from different readers. For this, the intraclass correlation coefficient (ICC (2,1)) was used, with a threshold of 0.9 (i.e., features with an ICC lower than 0.9 were deemed not robust to variations in image annotations).

Feature selection and model development

First, features deemed with low robustness to variations in annotation (ICC < 0.9) were discarded. Second, features that were dependent on contrast protocol and imaging system characteristics were eliminated. For this, the Mann-Whitney *U*-test was applied to test the median values of the features from Dataset 1&2 vs. Dataset 3. Features showing a significant difference ($p < 0.05$) were deemed dependent on variations in image acquisition and therefore discarded. Correction for multiple comparisons was not applied, to be more conservative on the number of features to retain. Third, the remaining features were *z*-score normalized (zero mean, unit variance), and those with extreme variation from a normal distribution (kurtosis > 15) or with too narrow distributions (interquartile range, IQR < 1) across patients were discarded as deemed non-informative [34]. Finally, highly correlated features (Pearson's $r > 0.95$) were eliminated. It should be noted that all the feature selection steps performed so far were performed blind to outcomes or any other patient characteristics, ensuring an unbiased analysis.

After these initial feature selection steps, the patient image dataset was divided into training (including inner validation) and test sets. To avoid generating potentially overconfident results due to evaluating the machine learning models on a single subset of cases, we performed a repeated training/test split for 100 iterations. For each iteration, the images were divided, randomly and on a patient level, into training (including inner validation) (80%) and test (20%), sampling the data so as to maintain the proportion of positive and negative classes in the entire dataset.

The last step of feature selection was performed by cross-validating (five-fold) a logistic regression model with the least absolute shrinkage selection operator (LASSO) regularization on the current training (including inner validation) set. In five-fold cross-validation, the training dataset was divided into five subsets, and in each round of the five-fold cross-validation, four subsets were used to train the model and one for inner validation and feature selection. At each of the 100 iterations, features with the lowest LASSO penalties were retained (to prevent overfitting, up to one-tenth of the number of training (including inner validation) cases).

Finally, a support vector machine (SVM) model with a second-order polynomial kernel was trained with the selected features for the 100 iterations to predict ALN status and metastatic burden. Hyperparameter tuning (i.e., kernel function) was performed, again, on the training (including inner validation) set following a five-fold cross-validation scheme.

The model was then tested (without any further selection or model tuning) for the 100 iterations on the corresponding different subsets of cases devoted to testing, allowing for a robust analysis of the performance. Results were quantified using average performance metrics across the 100 iterations and compared to those obtained when training the models (with the same methodology described above) using only clinical features, and with combined clinical-radiomics features. The frequency of the selected features across all iterations was also reported.

Statistical analysis

The continuous variables were compared using Student's *t*-test or Mann-Whitney *U*-test, as appropriate according to the normality of the feature distributions, and the categorical variables were compared using chi-square test or Fisher's exact test. The performance of all models was assessed by diagnostic statistics. AUC results were compared by quantifying the *p* value from the confidence intervals obtained from the variance across the iterations, following previously published method [35]. All data analyses were performed using Matlab R2021a (Mathworks) and SPSS 26.0 (IBM, International Business Machines Corp.). A two-tailed $p < 0.05$, corrected for multiple comparisons (the Bonferroni correction), was considered statistically significant.

Results

Patient characteristics

A total of 118 of 274 patients were diagnosed with ALN metastasis ($N + (\geq 1)$), and 74 of those had one or two ($N + (1–2)$) and 44 had more than two ($N + (\geq 3)$). The remaining 156 patients were diagnosed with no ALN metastasis (N0). The clinicopathologic characteristics of all patients are summarized in Table 1. With correction for multiple comparisons, size, histologic type, pathologic grade, and Ki-67 index were significantly different

Table 1 Clinicopathologic characteristics of all patients

	ALN status		***p***	**Metastatic burden**		***p***
	N0	***N*+ (≥ 1)**		***N*+ (1~2)**	***N*+ (≥ 3)**	
	(*n* = 156)	**(*n* = 118)**		**(*n* = 74)**	**(*n* = 44)**	
Age	50.84 ± 9.42	49.83 ± 10.33	> 0.05	50.07 ± 10.53	49.43 ± 10.07	> 0.05
Menopausal status			> 0.05			> 0.05
Premenopausal	90 (57.7)	63 (53.4)		36 (48.6)	27 (61.4)	
Postmenopausal	66 (42.3)	55 (46.6)		38 (51.4)	17 (38.6)	
FGT			> 0.05			> 0.05
Almost entirely fat	3 (1.9)	0 (0.0)		0 (0.0)	0 (0.0)	
Scattered	17 (10.9)	19 (16.1)		13 (17.6)	6 (13.6)	
Heterogeneous	105 (67.3)	78 (66.1)		51 (68.9)	27 (61.4)	
Extreme	31 (19.9)	21 (17.8)		10 (13.5)	11 (25.0)	
BPE			> 0.05			> 0.05
Minimal	38 (24.4)	26 (22.0)		16 (21.6)	10 (22.7)	
Mild	59 (37.8)	58 (49.2)		33 (44.6)	25 (56.9)	
Moderate	37 (23.7)	27 (22.9)		20 (27.0)	7 (15.9)	
Marked	22 (14.1)	7 (5.9)		5 (6.8)	2 (4.5)	
Presence			0.21			> 0.05
Mass	126 (80.8)	80 (67.8)		51 (68.9)	29 (65.9)	
Mass with NME	30 (19.2)	38 (32.2)		23 (31.1)	15 (34.1)	
Location			0.45			0.36
Upper outer	58 (37.2)	61 (51.8)		37 (50.0)	24 (54.5)	
Lower outer	30 (19.2)	24 (20.3)		17 (23.0)	7 (15.9)	
Lower inner	15 (9.6)	12 (10.2)		3 (4.1)	9 (20.5)	
Upper inner	51 (32.7)	20 (16.9)		16 (21.5)	4 (9.1)	
Retroareolar	2 (1.3)	1 (0.8)		1 (1.4)	0 (0.0)	
Size	2.05 ± 0.89	2.80 ± 1.44	< 0.001*	2.57 ± 1.44	3.18 ± 1.36	0.38
Focality			> 0.05			> 0.05
Unifocal	135 (86.5)	94 (79.7)		59 (79.7)	35 (79.5)	
Multifocal	17 (10.9)	17 (14.4)		10 (13.5)	7 (16.0)	
Multicentric	4 (2.6)	7 (5.9)		5 (6.8)	2 (4.5)	
Histologic type			0.03*			> 0.05
DCIS	8 (5.1)	0 (0.0)		0 (0.0)	0 (0.0)	
IDC with DCIS	30 (19.2)	19 (16.1)		14 (18.9)	5 (11.4)	
IDC	106 (68.0)	97 (82.3)		59 (79.7)	38 (86.3)	
ILC	1 (0.6)	1 (0.8)		1 (1.4)	0 (0.0)	
Others	11 (7.1)	1 (0.8)		0 (0.0)	1 (2.3)	
Pathologic grade			0.003*			> 0.05
I	13 (8.3)	0 (0.0)		0 (0.0)	0 (0.0)	
II	106 (68.0)	71 (60.2)		45 (60.8)	26 (59.1)	
III	37 (23.7)	47 (39.8)		29 (39.2)	18 (40.9)	
ER			> 0.05			> 0.05
Negative	46 (29.5)	46 (39.0)		30 (40.5)	16 (36.4)	
Positive	110 (70.5)	72 (61.0)		44 (59.5)	28 (63.6)	
PR			0.35			> 0.05
Negative	63 (40.4)	64 (54.2)		39 (52.7)	25 (56.8)	
Positive	93 (59.6)	54 (45.8)		35 (47.3)	19 (43.2)	
HER2			> 0.05			> 0.05
Negative	123 (78.8)	83 (70.3)		51 (68.9)	32 (72.7)	
Positive	33 (21.2)	35 (29.7)		23 (31.1)	12 (27.3)	

Table 1 (continued)

	ALN status		*p*	Metastatic burden		*p*
	N0	***N* + (≥ 1)**		***N* + (1~2)**	***N* + (≥ 3)**	
	(*n* = 156)	**(*n* = 118)**		**(*n* = 74)**	**(*n* = 44)**	
Ki-67	0.29 ± 0.21	0.41 ± 0.23	< 0.001*	0.4 ± 0.22	0.41 ± 0.24	> 0.05
Molecular subtype			0.35			> 0.05
Luminal A	33 (21.2)	10 (8.5)		6 (8.1)	4 (9.2)	
Luminal B	79 (50.6)	62 (52.5)		38 (51.3)	24 (54.5)	
HER2-enriched	20 (12.8)	23 (19.5)		17 (23.0)	6 (13.6)	
Triple-negative	24 (15.4)	23 (19.5)		13 (17.6)	10 (22.7)	

Data in parentheses are percentages. Values shown with asterisk indicate statistical significance

p values were corrected for multiple comparisons by a factor equal to the number of descriptors evaluated ($n = 15$). When the corrected *p* value was larger than 1, it demonstrated as > 0.05

ALN, axillary lymph node; *FGT*, fibroglandular tissue; *BPE*, background parenchymal enhancement; *NME*, non-mass enhancement; *DCIS*, ductal carcinoma in situ; *IDC*, invasive ductal carcinoma; *ILC*, invasive lobular carcinoma; *ER*, estrogen receptor; *PR*, progesterone receptor; *HER2*, human epidermal growth factor receptor 2

between N0 and $N + (\geq 1)$ groups, while no clinicopathologic characteristics had significant statistical difference between $N + (1–2)$ and $N + (\geq 3)$ groups.

Performance of machine learning models

Of the initial 1781 radiomics features, 296 were found not robust to variations in image segmentation, 275 not robust to variations in image acquisition, 579 to have non-informative distributions, and 205 to have high correlation among each other. As a result of these initial selection steps, 426 features were retained. The number of features finally selected by LASSO was never larger than 21 in each iteration in Study 1, and no more than 9 in each iteration in Study 2. Figure 3 shows the selection frequency and the statistical significance of the features from the 100 iterations by LASSO for each model of the two experiments.

Of the clinical features, size, histologic type, pathologic grade, and Ki-67 index were selected in at least 75/100 iterations and statistically significant with the Bonferroni correction when predicting ALN status, indicating that primary tumors with larger size, more invasiveness, higher grade, and higher Ki-67 expression were associated with positive ALNs. Of the radiomics features, wavelet-HLL_glcm_SumSquares (median [IQR] = 2.21 [1.73–2.78] vs. 1.74 [1.51–2.17]) and log-sigma-3-0-mm-3D_glcm_Idn (median [IQR] = 0.93 [0.92–0.95] vs. 0.95 [0.94–0.96]) were selected in at least 75/100 iterations and statistically significant with the Bonferroni correction ($p < 0.001$) when predicting ALN status, indicating that primary tumors with higher heterogeneity seem to be associated with positive ALNs. For metastatic burden prediction, only a single clinical feature (size of the primary tumor) had a high selection frequency (96/100) and was statistically significant ($p = 0.02$) initially, but not after the Bonferroni correction ($p = 0.38$). No radiomics features were found significant in distinguishing low vs. high metastatic burden.

The average AUC values (± one standard deviation) across the 100 test repeats in predicting ALN status (N0/$N + (\geq 1)$) were 0.75 ± 0.05 (radiomics model), 0.68 ± 0.07 (clinical model), and 0.74 ± 0.05 (combined model). For metastatic burden prediction ($N + (1–2)/N + (\geq 3)$), they were 0.65 ± 0.10 (radiomics model), 0.55 ± 0.10 (clinical model), and 0.64 ± 0.09 (combined model) (Fig. 4). Individual performance for each of the 100 training/test splits is reported in Table 2. The average AUC of radiomics and combined models were significantly higher than that of clinical models but showed no significant difference between themselves.

Discussion

In this study, radiomics models were developed and validated for preoperative prediction of ALN status and metastatic burden based on CE-CBBCT images. For both tasks, the AUCs of the radiomics model were significantly

(See figure on next page.)

Fig. 3 Heat map of selection frequency of clinical and radiomics features from 100 iterations by LASSO for each model in predicting ALN status and metastatic burden. The higher the selection frequency, the darker the blue color. Features that had never been selected by LASSO in any model were ignored. Values shown in orange indicate statistical significance. *p* values were corrected for multiple comparisons by a factor equal to the number of descriptors evaluated (Status-Clinical $n = 13$; Status-Radiomics $n = 23$; Status-Combined $n = 26$; Burden-Clinical $n = 15$; Burden-Radiomics $n = 54$; Burden-Combined $n = 60$). *LASSO*, least absolute shrinkage selection operator; *ALN*, axillary lymph node

Fig. 3 (See legend on previous page.)

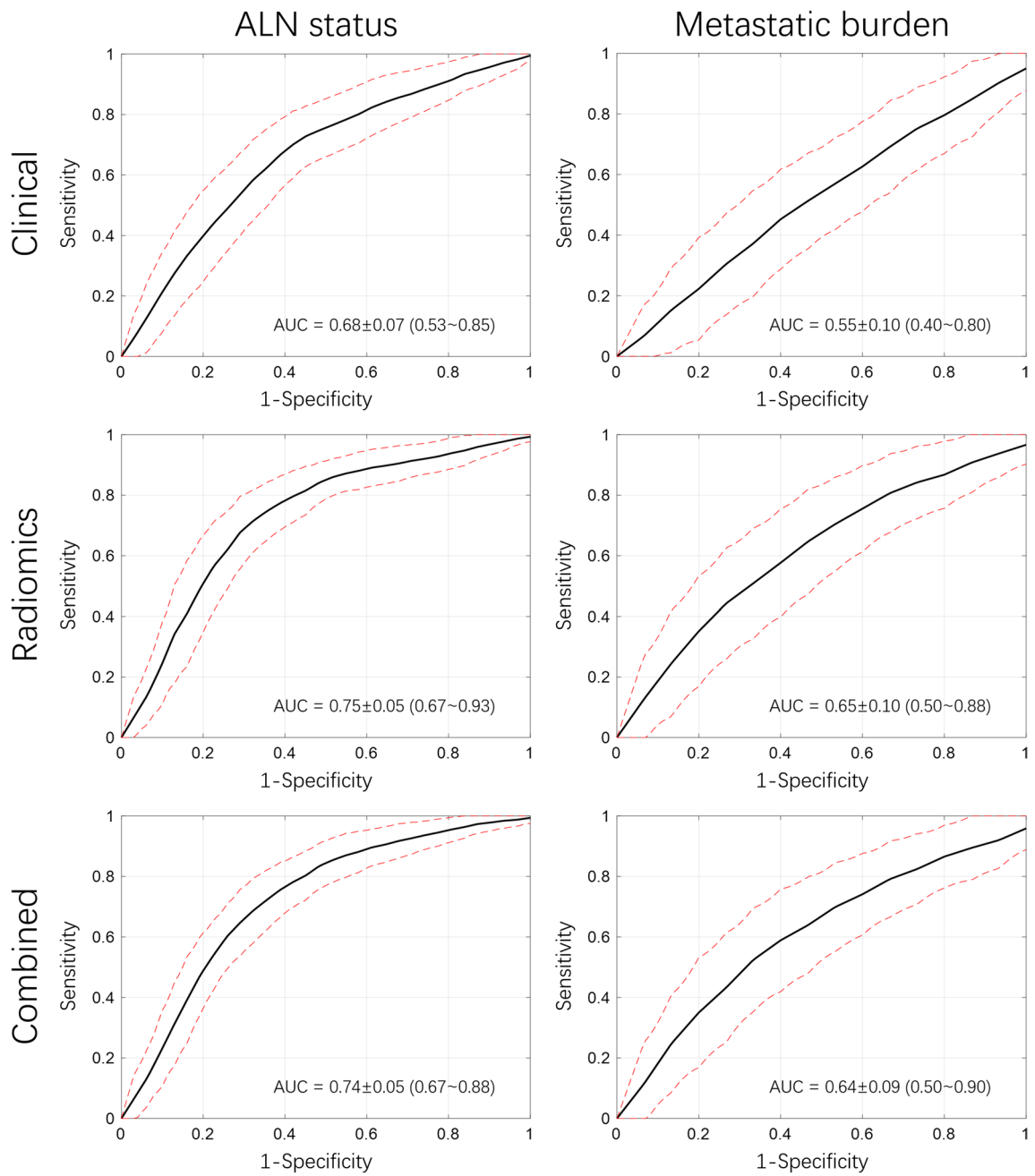

Fig. 4 ROC curves of pooled performance of clinical, radiomics, and combined models in predicting ALN status and metastatic burden. The average AUC of radiomics models was superior to clinical models (both $p < 0.01$), and comparable with combined models ($p = 0.56$ and 0.64, respectively). For each ROC curve, the average AUC (± standard deviation) across 100 test repeats is shown. Numbers in parentheses indicate the AUC ranges. *ROC*, receiver operating characteristic; *ALN*, axillary lymph node; *AUC*, area under the curve

higher than those based only on clinicopathologic indicators. Combining radiomics and clinical features resulted in no relevant performance improvements compared to the radiomics model alone.

Previously, several mammography, ultrasound, and MRI studies have explored the application of radiomics analysis of primary tumors in predicting ALN status and metastatic burden, and achieved good prediction performance, with AUC ranging from 0.64 to 0.89 and from 0.74 to 0.79 [36–38], respectively. Some studies incorporated ultrasound or MRI report of ALN or clinicopathologic characteristics to construct nomogram for further

Table 2 Diagnostic value of clinical, radiomics, and combined models in predicting ALN status and metastatic burden in test cohort

Model		Sensitivity				Specificity				PPV				NPV				Accuracy				AUC			
		AVG	SD	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD	MIN	MAX
Status	Clinical	0.59	0.18	0.13	0.96	0.78	0.12	0.52	1.00	0.69	0.09	0.54	1.00	0.72	0.07	0.58	0.95	0.69	0.05	0.58	0.80	0.68	0.07	0.53	0.85
	Radiomics	0.71	0.12	0.38	0.96	0.77	0.10	0.52	1.00	0.72	0.08	0.58	1.00	0.78	0.06	0.66	0.96	0.74	0.05	0.66	0.89	0.75	0.05	0.67	0.93
	Combined	0.68	0.15	0.25	0.96	0.77	0.10	0.52	1.00	0.71	0.07	0.59	1.00	0.77	0.08	0.63	0.96	0.73	0.04	0.66	0.84	0.74	0.05	0.67	0.88
Burden	Clinical	0.22	0.25	0.00	1.00	0.92	0.09	0.53	1.00	0.53	0.42	0.00	1.00	0.67	0.07	0.60	1.00	0.66	0.07	0.54	0.83	0.55	0.10	0.40	0.80
	Radiomics	0.42	0.25	0.00	0.89	0.89	0.10	0.60	1.00	0.71	0.28	0.00	1.00	0.73	0.08	0.61	0.93	0.72	0.06	0.58	0.88	0.65	0.10	0.50	0.88
	Combined	0.43	0.27	0.00	0.89	0.88	0.12	0.53	1.00	0.71	0.27	0.00	1.00	0.74	0.08	0.61	0.93	0.71	0.06	0.58	0.88	0.64	0.09	0.50	0.90

ALN axillary lymph node, *PPV* positive predictive value, *NPV* negative predictive value, *AUC* area under the curve, *AVG* average, *SD* standard deviation, *MIN* minimum, *MAX* maximum

improvement compared to radiomics alone [22, 23, 29, 39–46], even with some factors that were only available postoperatively, such as lymph-vascular invasion [45] and genomic data [46]. To the best of our knowledge, this is the first study on the prediction of ALN status and metastatic burden by CBBCT-based radiomics, using only preoperative imaging data, clinical information, and pathologic results of the primary breast tumor obtained from biopsy.

Obtaining radiomics results that are robust and generalizable within and across institutions is currently a significant challenge for the clinical translation of radiomics [47]. A recent study has shown that the use of a single random training/test set split may lead to unreliable results in small-sample radiomics studies [48]. With the objective to test the potential generalizability of our radiomics model, and provide more conservative results, we randomly divided the dataset into training (including inner validation) and test sets 100 times, each time using the training (including inner validation) set (in a five-fold cross-validation scheme) for feature selection and model development, and the test set uniquely for result assessment. Overall, satisfactory results were obtained for the prediction of ALN status (considering that all analyses were performed from the primary tumor only), pointing to the value of radiomics compared to simple clinical models. For the prediction of metastatic burden, performance was moderate, indicating that limited information is contained in primary tumors that may be used to differentiate between low- and high-burden ALN metastasis. While average results for ALN status were promising, and in line with previous literature on different modalities, individual AUC values varied considerably across the 100 training/test repeats. This demonstrates that prediction performance may vary considerably according to differences in datasets, and therefore future studies with larger sets are needed before a unique radiomics model can be developed, and to conclusively assess the potential presence of significant CE-CBBCT-based image biomarkers quantifiable through radiomics that are predictive of ALN status. In this study, we adopted a multi-step feature selection process blind to outcomes to discard non-robust, uninformative, and highly correlated features, followed by LASSO applied in a cross-validated fashion. In future work, additional feature selection or dimensionality reduction approaches could be investigated, to evaluate systematically how each method impacts the classification results.

There was no statistical difference in terms of predictive efficacy between the combined and radiomics models in our study, which was not consistent with some previous studies. This perhaps related to the imaging-reported ALN status added into the construction of combined model in some studies, while our study was only based on the primary tumor. According to the "seed and soil" theory [49], ALN metastasis initiation depends on the synergies of tumor cells (seed) and ALN microenvironment (soil), which may explain why including the ultrasound or MRI report of ALN could improve the performance of prediction model. In addition, the preoperative clinicopathologic factors may have little contribution to the combined model of ALN metastasis prediction. Tumor-based clinicopathologic features are the external manifestations of intratumoral heterogeneity [50]. Radiomics features acquired from macroscopic images could comprehensively reflect the innate heterogeneity of tumors. Therefore, the tumor-based clinicopathologic features may be redundant in improving the predicting performance, as shown by our results. Furthermore, another possible explanation may be the limited dataset of our study, in which radiomics features had stronger prediction performance than clinical indicators and thus were more frequently selected during the 100 iterations for developing the combined model. This could lower the impact of potentially relevant clinical descriptors. Perhaps the impact of clinical descriptors could become higher with larger cohorts, where a larger number of features can be selected and used. Of course, additional studies with larger datasets are needed to further explore these findings.

In the literature, relatively high AUCs of some previous studies for ALN status and metastatic burden prediction were obtained. The reasons may be twofold. On the one hand, some modalities (e.g., multiparametric

MRI, PET-CT) allow for the capture of information in a way that CBBCT currently cannot achieve (hemodynamics, diffusion, metabolics) [24, 39, 51]. Further investigation is needed to make it clear whether adding temporal information to CBBCT [52, 53] will contribute to better predictive performance of ALN metastasis. On the other hand, differences in performance might be due to variations in datasets and training/validation/test split, as well as in methodology in general. Even within previously published literature, likely for the same reason, the results reported differ considerably [22, 40, 41]. Therefore, at the moment, all we can conclude is that our findings seem to be in line with previous literature, but direct and objective comparisons are currently impossible to achieve.

The Memorial Sloan-Kettering Cancer Center (MSKCC) nomogram is the most well-known method to predict ALN status with an AUC of 0.754, which has been widely validated in different populations [54]. However, the MSKCC model uses some parameters that can only be obtained after surgery, such as lymphovascular invasion, and does not include any imaging features. Compared with the MSKCC nomogram, our model has similar prediction performance and has the advantage of evaluating ALN metastasis preoperatively and non-invasively. This could be beneficial to facilitate clinical utility and assist physicians in predicting the risk of ALN metastasis (pending more extensive and heterogeneous validations).

Although the frequency of the selected features across the 100 repeats varied considerably, a few clinical and radiomics descriptors were selected most often for the prediction of ALN status. Ki-67 index (selected 99/100 and 67/100 in the clinical and combined models, respectively) is an indicator of tumor aggressiveness and proliferative activity, in which a higher expression level indicates more aggressive growth of breast cancer [55]. Wavelet-HLL_glcm_SumSquares (selected 85/100 and 69/100 in the radiomics and combined models, respectively) and log-sigma-3-0-mm-3D_glcm_Idn (selected 100/100 in both the radiomics and combined models) are high-order features that reflect intratumor heterogeneity. The former is obtained from the gray level cooccurrence matrix of the image, after decomposing the image with high-pass filtering. The latter is also obtained from the image gray level cooccurrence matrix, this time after filtering with LoG (an edge enhancement filter). They both emphasize areas of gray level change, i.e., intratumor heterogeneity, which seems to be better captured by these two high-order statistical descriptors than by other simpler texture features. This could be due to subtle details in the tumor texture that seem to be highlighted only through the transformation of the image obtained with high-order filtering. Although it is impossible to determine exactly which biological characteristics are captured by these two features, we hypothesize that these subtle details might be related to small alterations in microvessel density and in the tumor microenvironment [56, 57]. These alterations, due to angiogenesis and microenvironmental changes, might be too subtle to be captured by simpler texture features, and might instead become highlighted only by more advanced, higher-order descriptors. All in all, the higher the aggressiveness and heterogeneity of the primary cancer, the higher seems to be the possibility of ALN metastasis. However, for metastatic burden prediction, no radiomics features were selected with consistent frequency across the training/test repeats; only the tumor size as a clinical indicator had a high selection frequency, but with no statistical significance.

Our study has some limitations. First, due to CBBCT still being a relatively new breast imaging modality, limited sample size and unbalanced data distribution existed in this study. Although the size of our dataset is in line with many previously reported exploratory radiomics studies [22, 24, 40, 45], larger and more heterogeneous datasets are needed in the future to confirm our findings, especially those pertaining to metastatic burden prediction. Additionally, in this study, we did not explicitly correct for the class imbalance present in our dataset (for both ALN status and metastatic burden prediction). This was due to the moderate imbalance ratio (43% : 57% for ALN status, 37% : 63% for metastatic burden), and to respect the proportion of positive and negative cases as found in the clinical realm [39, 44]. Instead, to maximize the prevention from biased results, we adopted simple machine learning models (less prone to overfitting or skewed datasets), and we used ROC and AUC as figures of merit across the 100 test repeats. Of course, future studies with larger datasets ought to be performed to further confirm our results. Second, selection bias inevitably remained in this retrospective study. For multifocal/multicentric cancers, only the largest mass was considered, and this may ignore the influence of other foci. We performed this choice to be consistent with the AJCC 8th edition staging system, where the T category is based on the size of the largest mass (when multiple masses are present). In other words, for multifocal and multicentric disease, the size of the largest tumor rather than the sum of all foci is used for clinical or pathologic tumor classification (cT or pT) [58, 59]. In future work, and with larger datasets, the inclusion of additional foci for multifocal/multicentric cancers should be investigated, to evaluate whether this brings any benefits in terms of radiomics analyses. Besides, NME was excluded, due to the difficulty of delineating its boundary precisely. Furthermore, only images of patients with BI-RADS 3~5 and/or ACR density type c or d were used in our study, due to the inclusion criteria

of the two prospective clinical trials during which the data were acquired. Therefore, future studies are needed with more heterogeneous datasets. Third, the pathologic result of ALN was confirmed by combining SLNB and ALND as the reference, which is the clinical gold standard [30]. However, it should be noted that SLNB has an overall false-negative rate of 8.61%, ranging from 0 to 27.3% according to different injection materials and sites [60], which could have some impact on the results of this study. Fourth, the images in our study were collected by two different devices but from the same vendor (Koning BCT). The reproducibility with other BCT devices, as well as with other datasets, remains to be investigated. Meanwhile, in future work, when larger and more heterogeneous datasets become available, harmonization strategies may be adopted to reduce the heterogeneity in imaging parameters [61]. This would be especially important when additional breast CT devices are introduced for clinical use, to lower the impact of differences in imaging settings to the extracted radiomics signature. Last, although we followed the radiomics quality score (RQS) [62] when performing our radiomics analysis, where applicable, some criteria could not be met due to inherent characteristics of our study (imaging at multiple time points, prospective study, cost-effectiveness).

Future work may include the evaluation of peritumoral features, which may further improve the prediction performance [28, 63], and, if larger datasets become available, deep learning approaches [29, 64].

Conclusion

Radiomics features of primary tumors in CE-CBBCT images have the potential to predict the extent of ALN involvement prior to surgery. This could be especially important in CBBCT, since it could help address the limited axilla coverage. Furthermore, this might have the potential to aid in clinical decision-making, especially to mitigate false-negative SLNB findings, pending further validation with larger prospective cohorts.

Abbreviations

ACOSOG	American College of Surgeons Oncology Group
ALN	Axillary lymph node
ALND	Axillary lymph node dissection
AUC	Area under the curve
CE-CBBCT	Contrast-enhanced cone-beam breast CT
FNB	Fine-needle biopsy
ICC	Intraclass correlation coefficient
IQR	Interquartile range
LoG	Laplacian of Gaussian
LASSO	Least absolute shrinkage selection operator
NME	Non-mass enhancement
ROC	Receiver operating characteristic
SLNB	Sentinel lymph node biopsy
SVM	Support vector machine
VOI	Volume of interest

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s00330-023-10256-4.

ESM 1 (PDF 182 kb)

Funding

This study was supported by National Key R&D Program of China (No. 2021YFC2500400, 2021YFC2500402, 2017YFC0112600, 2017YFC0112601, 2017YFC0112605), National Natural Science Foundation of China (No. 81571671), Tianjin Science and Technology Major Project (No. 19ZXDBSY00080), Key Project of Tianjin Medical Industry (No. 16KG130), Tianjin Medical University Cancer Institute and Hospital Fund (B2118, B2219), and Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-009A). This study was also supported in part by the National Cancer Institute (NCI) of the National Institutes of Health (NIH) (No. R01CA181171). The content is solely the responsibility of the authors and does not represent the official views of the NCI or the NIH.

Declarations

Guarantor

The scientific guarantor of this publication is Zhaoxiang Ye.

Conflict of Interest

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and Biometry

One of the authors (Marco Caballo) has significant statistical expertise.

Informed Consent

Data for this study was collected from two prospective clinical trials, named "Koning breast CT for breast imaging in China" and "The technical operations and standard clinical application protocol of cone-beam breast CT in diagnostic process of breast cancer", in which written informed consent including permission to re-use the data for any further retrospective analysis was obtained from every patient at the time of enrollment.

Ethical Approval

Institutional Review Board approval was obtained (E2012036, bc2016039, and A2011-030-01).

Study subjects or cohorts overlap

The study subjects were collected from two prospective clinical trials (NCT01792999 and NCT03861221), aiming to assess the performance of new-generation breast imaging modality—CBBCT and explore the clinical application guideline of CBBCT, respectively. There have been several publications that share the same cohort with the current study regarding breast coverage and patient comfort comparison, diagnostic performance analysis, visual and quantitative breast density assessment, molecular subtyping, tumor size evaluation, and BI-RADS atlas exploration. In contrast, here we investigated axillary lymph node status and metastatic burden prediction that has not been evaluated previously.

• Li H, Yin L, Ye Z et al (2015) Comparative study of breast tissue coverage in cone-beam breast CT versus digital mammography. Chin J Radiol 49:488-490 in Chinese
• He N, Wu YP, Kong Y et al (2016) The utility of breast cone-beam computed tomography, ultrasound, and digital mammography for detecting malignant breast tumors: a prospective study with 212 patients. Eur J Radiol 85:392-403
• Yin L, Ye Z (2016) New 3D X-ray modalities in breast imaging: digital breast tomosynthesis and cone-beam breast computed tomography. Chin Med Device Inform 22:17-20 in Chinese
• Liu A, Ye Z, Ma Y, Cao Y (2018) Reliability of breast density estimation based on cone-beam breast CT. Chin J Clin Oncol 45:246-250 in Chinese
• Liu A, Ma Y, Yin L, Han P, Li H, Ye Z (2018) Comparison of the diagnostic efficiency in breast malignancy between cone-beam breast CT and mammography in dense breast. Chin J Oncol 40:604-609 in Chinese

• Liu A, Ma Y, Yin L, Han P, Li H, Ye Z (2018) Diagnostic value of contrast-enhanced cone-beam breast CT in dense breast lesions. Chin Oncol 28:807-812 in Chinese
• Ma Y, Ye Z, Liu A, Yin L, Han P, Li H (2019) The accuracy of tumor size evaluation on invasive breast cancer based on cone-beam breast CT. Chin J Radiol 53:286-291 in Chinese
• Ma Y, Cao Y, Liu A et al (2019) A reliability comparison of cone-beam breast computed tomography and mammography: breast density assessment referring to the fifth edition of the BI-RADS atlas. Acad Radiol 26:752-759
• Li H, Yin L, He N et al (2019) Comparison of comfort between cone-beam breast computed tomography and digital mammography. Eur J Radiol 120:108674
• Zhu Y, Zhang Y, Ma Y et al (2020) Cone-beam breast CT features associated with HER2/neu overexpression in patients with primary breast cancer. Eur Radiol 30:2731-2739
• Ma Y, Liu A, O'Connell AM et al (2021) Contrast-enhanced cone-beam breast CT features of breast cancers: correlation with immunohistochemical receptors and molecular subtypes. Eur Radiol 31:2580-2589
• Wang Y, Ma Y, Zhu Y et al (2021) Value of cone-beam breast CT in differentiating benign from malignant dense breast masses. Chin J Radiol 55:961-967 in Chinese
• Zhang Y, Ma Y, Li Y et al (2021) Comparative study of cone-beam breast CT and breast MRI in diagnosis of BI-RADS 4 lesions on mammography or ultrasound. J Clin Radiol 40:1703-1708 in Chinese
• Zhu Y, O'Connell AM, Ma Y et al (2022) Dedicated breast CT: state of the art-part I. historical evolution and technical aspects. Eur Radiol 32:1579-1589
• Zhu Y, O'Connell AM, Ma Y et al (2022) Dedicated breast CT: state of the art-part II. clinical application and future outlook. Eur Radiol 32:2286-2300
• Liu A, Yin L, Ma Y et al (2022) Quantitative breast density measurement based on three-dimensional images: a study on cone-beam breast computed tomography. Acta Radiol 63:1023-1031
• Ma Y, Liu A, Zhang Y et al (2022) Comparison of background parenchymal enhancement (BPE) on contrast-enhanced cone-beam breast CT (CE-CBBCT) and breast MRI. Eur Radiol 32:5773-5782
• Liu A, Ma Y, Yin L et al (2023) Comparison of malignant calcification identification between breast cone-beam computed tomography and digital mammography. Acta Radiol 64:962-970
• Wang Y, Zhao M, Ma Y et al (2023) Accuracy of preoperative contrast-enhanced cone-beam breast CT in assessment of residual tumor after neoadjuvant chemotherapy: a comparative study with breast MRI. Acad Radiol 30:1805-1815

Methodology

• Retrospective
• diagnostic or prognostic study
• multicenter study

Author details
[1]Department of Radiology, Tianjin Medical University Cancer Institute and Hospital; National Clinical Research Center for Cancer; Tianjin's Clinical Research Center for Cancer; Key Laboratory of Cancer Prevention and Therapy, Tianjin; Key Laboratory of Breast Cancer Prevention and Therapy, Tianjin Medical University, Ministry of Education, Huan-Hu-Xi Road, Ti-Yuan-Bei, Hexi District, Tianjin 300060, China. [2]Department of Medical Imaging, Radboud University Medical Center, PO Box 9101, Nijmegen 6500 HB, The Netherlands. [3]Department of Radiology, The Fifth Affiliated Hospital of Sun Yat-Sen University, Mei-Hua-Dong Road, Xiangzhou District, Zhuhai 519000, China. [4]Department of Medical Imaging and Image-guided Therapy, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Dong-Feng-Dong Road, Yuexiu District, Guangzhou 510060, China. [5]Dutch Expert Center for Screening (LRCB), PO Box 6873, Nijmegen 6503 GJ, The Netherlands. [6]Technical Medicine Centre, University of Twente, PO Box 217, Enschede 7500 AE, The Netherlands.

Received: 4 October 2022 Revised: 9 July 2023 Accepted: 30 July 2023
Published online: 02 October 2023

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A (2022) Cancer statistics, 2022. CA Cancer J Clin 72:7–33
2. Gradishar WJ, Moran MS, Abraham J et al (2022) Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw 20:691–722
3. Giuliano AE, Ballman KV, McCall L et al (2017) 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 318:918–926
4. Zhu Y, Li X, Wang F et al (2018) Intravoxel incoherent motion diffusion-weighted magnetic resonance imaging in characterization of axillary lymph nodes: preliminary animal experience. Magn Reson Imaging 52:46–52
5. Li H, Yin L, He N et al (2019) Comparison of comfort between cone beam breast computed tomography and digital mammography. Eur J Radiol 120:108674
6. Uhlig J, Uhlig A, Biggemann L, Fischer U, Lotz J, Wienbeck S (2019) Diagnostic accuracy of cone-beam breast computed tomography: a systematic review and diagnostic meta-analysis. Eur Radiol 29:1194–1202
7. Zhu Y, O'Connell AM, Ma Y et al (2022) Dedicated breast CT: state of the art-part I. historical evolution and technical aspects. Eur Radiol 32:1579–1589
8. Zhu Y, O'Connell AM, Ma Y et al (2022) Dedicated breast CT: state of the art-part II. clinical application and future outlook. Eur Radiol 32:2286–2300
9. Liu A, Ma Y, Yin L et al (2023) Comparison of malignant calcification identification between breast cone-beam computed tomography and digital mammography. Acta Radiol 64:962–970
10. Wienbeck S, Uhlig J, Luftner-Nagel S et al (2017) The role of cone-beam breast-CT for breast cancer detection relative to breast density. Eur Radiol 27:5185–5195
11. Wienbeck S, Fischer U, Luftner-Nagel S, Lotz J, Uhlig J (2018) Contrast-enhanced cone-beam breast-CT (CBBCT): clinical performance compared to mammography and MRI. Eur Radiol 28:3731–3741
12. Uhlig J, Fischer U, Biggemann L, Lotz J, Wienbeck S (2019) Pre- and post-contrast versus post-contrast cone-beam breast CT: can we reduce radiation exposure while maintaining diagnostic accuracy? Eur Radiol 29:3141–3148
13. Kang W, Zhong W, Su D (2021) The cone-beam breast computed tomography characteristics of breast non-mass enhancement lesions. Acta Radiol 62:1298–1308
14. Zhu Y, Zhang Y, Ma Y et al (2020) Cone-beam breast CT features associated with HER2/neu overexpression in patients with primary breast cancer. Eur Radiol 30:2731–2739
15. Ma Y, Liu A, O'Connell AM et al (2021) Contrast-enhanced cone beam breast CT features of breast cancers: correlation with immunohistochemical receptors and molecular subtypes. Eur Radiol 31:2580–2589
16. Wienbeck S, Uhlig J, Fischer U et al (2019) Breast lesion size assessment in mastectomy specimens: correlation of cone-beam breast-CT, digital breast tomosynthesis and full-field digital mammography with histopathology. Medicine (Baltimore) 98:e17082
17. Wang Y, Zhao M, Ma Y et al (2023) Accuracy of preoperative contrast-enhanced cone beam breast CT in assessment of residual tumor after neoadjuvant chemotherapy: a comparative study with breast MRI. Acad Radiol 30:1805–1815
18. Ma Y, Cao Y, Liu A et al (2019) A reliability comparison of cone-beam breast computed tomography and mammography: breast density assessment referring to the fifth edition of the BI-RADS atlas. Acad Radiol 26:752–759
19. Liu A, Yin L, Ma Y et al (2022) Quantitative breast density measurement based on three-dimensional images: a study on cone-beam breast computed tomography. Acta Radiol 63:1023–1031
20. O'Connell A, Conover DL, Zhang Y et al (2010) Cone-beam CT for breast imaging: Radiation dose, breast coverage, and image quality. AJR Am J Roentgenol 195:496–509
21. Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E (2021) A deep look into radiomics. Radiol Med 126:1296–1311
22. Yang J, Wang T, Yang L et al (2019) Preoperative prediction of axillary lymph node metastasis in breast cancer using mammography-based radiomics method. Sci Rep 9:4429
23. Yu FH, Wang JX, Ye XH, Deng J, Hang J, Yang B (2019) Ultrasound-based radiomics nomogram: a potential biomarker to predict axillary lymph

node metastasis in early-stage invasive breast cancer. Eur J Radiol 119:108658
24. Dong Y, Feng Q, Yang W et al (2018) Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. Eur Radiol 28:582–591
25. Caballo M, Hernandez AM, Lyu SH et al (2021) Computer-aided diagnosis of masses in breast computed tomography imaging: deep learning model with combined handcrafted and convolutional radiomic features. J Med Imaging (Bellingham) 8:024501
26. Caballo M, Pangallo DR, Sanderink W et al (2021) Multi-marker quantitative radiomics for mass characterization in dedicated breast CT imaging. Med Phys 48:313–328
27. Caballo M, Pangallo DR, Mann RM, Sechopoulos I (2020) Deep learning-based segmentation of breast masses in dedicated breast CT imaging: radiomic feature stability between radiologists and artificial intelligence. Comput Biol Med 118:103629
28. Ding J, Chen S, Serrano Sosa M et al (2022) Optimizing the peritumoral region size in radiomics analysis for sentinel lymph node status prediction in breast cancer. Acad Radiol 29:S223–S228
29. Wang D, Hu Y, Zhan C, Zhang Q, Wu Y, Ai T (2022) A nomogram based on radiomics signature and deep-learning signature for preoperative prediction of axillary lymph node metastasis in breast cancer. Front Oncol 12:940655
30. National Health Commission of the People's Republic of China (2022) National guidelines for diagnosis and treatment of breast cancer 2022 in China (English version). Chin J Cancer Res 34:151–175
31. Ma Y, Liu A, Zhang Y et al (2022) Comparison of background parenchymal enhancement (BPE) on contrast-enhanced cone-beam breast CT (CE-CBBCT) and breast MRI. Eur Radiol 32:5773–5782
32. He N, Wu YP, Kong Y et al (2016) The utility of breast cone-beam computed tomography, ultrasound, and digital mammography for detecting malignant breast tumors: a prospective study with 212 patients. Eur J Radiol 85:392–403
33. van Griethuysen JJM, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. Cancer Res 77:e104–e107
34. Chitalia RD, Rowland J, McDonald ES et al (2020) Imaging phenotypes of breast cancer heterogeneity in preoperative breast dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) scans predict 10-year recurrence. Clin Cancer Res 26:862–869
35. Altman DG, Bland JM (2011) How to obtain the p value from a confidence interval. BMJ 343:d2304
36. Zhang J, Li L, Zhe X et al (2022) The diagnostic performance of machine learning-based radiomics of DCE-MRI in predicting axillary lymph node metastasis in breast cancer: a meta-analysis. Front Oncol 12:799209
37. Gong X, Guo Y, Zhu T, Peng X, Xing D, Zhang M (2022) Diagnostic performance of radiomics in predicting axillary lymph node metastasis in breast cancer: a systematic review and meta-analysis. Front Oncol 12:1046005
38. Eldaly AS, Avila FR, Torres-Guzman RA et al (2023) Radiomics and artificial intelligence in predicting axillary lymph node metastasis in breast cancer: a systematic review. Curr Med Imaging 19:564–578
39. Zhang X, Yang Z, Cui W et al (2021) Preoperative prediction of axillary sentinel lymph node burden with multiparametric MRI-based radiomics nomogram in early-stage breast cancer. Eur Radiol 31:5924–5939
40. Tan H, Wu Y, Bao F et al (2020) Mammography-based radiomics nomogram: a potential biomarker to predict axillary lymph node metastasis in breast cancer. Br J Radiol 93:20191019
41. Mao N, Yin P, Li Q et al (2020) Radiomics nomogram of contrast-enhanced spectral mammography for prediction of axillary lymph node metastasis in breast cancer: a multicenter study. Eur Radiol 30:6732–6739
42. Han L, Zhu Y, Liu Z et al (2019) Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. Eur Radiol 29:3820–3829
43. Jiang M, Li CL, Luo XM et al (2022) Radiomics model based on shear-wave elastography in the assessment of axillary lymph node status in early-stage breast cancer. Eur Radiol 32:2313–2325
44. Gao Y, Luo Y, Zhao C et al (2021) Nomogram based on radiomics analysis of primary breast cancer ultrasound images: prediction of axillary lymph node tumor burden in patients. Eur Radiol 31:928–937
45. Liu C, Ding J, Spuhler K et al (2019) Preoperative prediction of sentinel lymph node metastasis in breast cancer by radiomic signatures from dynamic contrast-enhanced MRI. J Magn Reson Imaging 49:131–140
46. Chen H, Lan X, Yu T et al (2022) Development and validation of a radiogenomics model to predict axillary lymph node metastasis in breast cancer integrating MRI with transcriptome data: a multicohort study. Front Oncol 12:1076267
47. Papanikolaou N, Matos C, Koh DM (2020) How to develop a meaningful radiomic signature for clinical use in oncologic patients. Cancer Imaging 20:33
48. An C, Park YW, Ahn SS, Han K, Kim H, Lee SK (2021) Radiomics machine learning study with a small sample size: single random training-test set split may lead to unreliable results. PLoS One 16:e0256152
49. Akhtar M, Haider A, Rashid S, Al-Nabet ADMH (2019) Paget's "seed and soil" theory of cancer metastasis: an idea whose time has come. Adv Anat Pathol 26:69–74
50. McGranahan N, Swanton C (2017) Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 168:613–628
51. Eifer M, Pinian H, Klang E et al (2022) FDG PET/CT radiomics as a tool to differentiate between reactive axillary lymphadenopathy following COVID-19 vaccination and metastatic breast cancer axillary lymphadenopathy: a pilot study. Eur Radiol 32:5921–5929
52. Caballo M, Mann R, Sechopoulos I (2018) Patient-based 4D digital breast phantom for perfusion contrast-enhanced breast CT imaging. Med Phys 45:4448–4460
53. Caballo M, Michielsen K, Fedon C, Sechopoulos I (2019) Towards 4D dedicated breast CT perfusion imaging of cancer: development and validation of computer simulated images. Phys Med Biol 64:245004
54. Fong W, Tan L, Tan C et al (2022) Predicting the risk of axillary lymph node metastasis in early breast cancer patients based on ultrasonographic-clinicopathologic features and the use of nomograms: a prospective single-center observational study. Eur Radiol 32:8200–8212
55. Pathmanathan N, Balleine RL (2013) Ki67 and proliferation in breast cancer. J Clin Pathol 66:512–516
56. Majidpoor J, Mortezaee K (2021) Angiogenesis as a hallmark of solid tumors-clinical perspectives. Cell Oncol (Dordr) 44:715–737
57. Bhat AA, Yousuf P, Wani NA et al (2021) Tumor microenvironment: an evil nexus promoting aggressive head and neck squamous cell carcinoma and avenue for targeted therapy. Signal Transduct Target Ther 6:12
58. Kalli S, Semine A, Cohen S, Naber SP, Makim SS, Bahl M (2018) American Joint Committee on Cancer's Staging System for Breast Cancer, eighth edition: what the radiologist needs to know. Radiographics 38:1921–1933
59. Teichgraeber DC, Guirguis MS, Whitman GJ (2021) Breast cancer staging: updates in the AJCC cancer staging manual, 8th edition, and current challenges for radiologists, from the AJR special series on cancer staging. AJR Am J Roentgenol 217:278–290
60. Pesek S, Ashikaga T, Krag LE, Krag D (2012) The false-negative rate of sentinel node biopsy in patients with breast cancer: a meta-analysis. World J Surg 36:2239–2251
61. Singh A, Horng H, Chitalia R et al (2022) Resampling and harmonization for mitigation of heterogeneity in image parameters of baseline scans. Sci Rep 12:21505
62. Lambin P, Leijenaar RTH, Deist TM et al (2017) Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 14:749–762
63. Sun Q, Lin X, Zhao Y et al (2020) Deep learning vs. radiomics for predicting axillary lymph node metastasis of breast cancer using ultrasound images: don't forget the peritumoral region. Front. Oncol 10:53
64. Zheng X, Yao Z, Huang Y et al (2020) Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nat Commun 11:1236

Publisher's note