

RESEARCH

Open Access



Predicting the efficacy of neoadjuvant chemotherapy in breast cancer patients based on ultrasound longitudinal temporal depth network fusion model

Xiaodan Feng^{1†}, Yan Shi^{2†}, Meng Wu¹, Guanghe Cui¹, Yao Du¹, Jie Yang¹, Yuyuan Xu¹, Wenjuan Wang¹ and Feifei Liu^{1*}

Abstract

Objective The aim of this study was to develop and validate a deep learning radiomics (DLR) model based on longitudinal ultrasound data and clinical features to predict pathologic complete response (pCR) after neoadjuvant chemotherapy (NAC) in breast cancer patients.

Methods Between January 2018 and June 2023, 312 patients with histologically confirmed breast cancer were enrolled and randomly assigned to a training cohort ($n=219$) and a test cohort ($n=93$) in a 7:3 ratio. Next, pre-NAC and post-treatment 2-cycle ultrasound images were collected, and radiomics and deep learning features were extracted from NAC pre-treatment (Pre), post-treatment 2 cycle (Post), and Delta (pre-NAC—NAC 2 cycle) images. In the training cohort, to filter features, the intraclass correlation coefficient test, the Boruta algorithm, and the least absolute shrinkage and selection operator (LASSO) logistic regression were used. Single-modality models (Pre, Post, and Delta) were constructed based on five machine-learning classifiers. Finally, based on the classifier with the optimal predictive performance, the DLR model was constructed by combining Pre, Post, and Delta ultrasound features and was subsequently combined with clinical features to develop a combined model (Integrated). The discriminative power, predictive performance, and clinical utility of the models were further evaluated in the test cohort. Furthermore, patients were assigned into three subgroups, including the HR+/HER2-, HER2+, and TNBC subgroups, according to molecular typing to validate the predictability of the model across the different subgroups.

Results After feature screening, 16, 13, and 10 features were selected to construct the Pre model, Post model, and Delta model based on the five machine learning classifiers, respectively. The three single-modality models based on the XGBoost classifier displayed optimal predictive performance. Meanwhile, the DLR model (AUC of 0.827) was superior to the single-modality model (Pre, Post, and Delta AUCs of 0.726, 0.776, and 0.710, respectively) in terms of prediction performance. Moreover, multivariate logistic regression analysis identified Her-2 status and histological grade as independent risk factors for NAC response in breast cancer. In both the training and test cohorts, the

[†]Xiaodan Feng and Yan Shi contributed equally to this work.

*Correspondence:

Feifei Liu
niu.feif521@163.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Integrated model, which included Pre, Post, and Delta ultrasound features and clinical features, exhibited the highest predictive ability, with AUC values of 0.924 and 0.875, respectively. Likewise, the Integrated model displayed the highest predictive performance across the different subgroups.

Conclusion The Integrated model, which incorporated pre-NAC treatment and early treatment ultrasound data and clinical features, accurately predicted pCR after NAC in breast cancer patients and provided valuable insights for personalized treatment strategies, allowing for timely adjustment of chemotherapy regimens.

Keywords Breast cancer, Neoadjuvant chemotherapy, Deep learning, Pathologic complete response

Introduction

As is well documented, breast cancer poses a serious threat to the lives and health of women, as it is the leading cause of both cancer incidence and mortality for women [1]. Before radical surgery or radiotherapy, NAC is administered as a systemic chemotherapy, with the primary goal of reducing tumor burden, lowering tumor stage, and increasing the rate of breast conservation to improve overall patient outcomes and quality of life [2]. NAC, however, has a significantly variable response among breast cancer patients, with merely 19–30% of patients achieving pCR and 5–20% of patients experiencing disease progression after chemotherapy [3, 4]. Patients with poor response to NAC not only have to endure various toxic effects and high treatment costs but also face the risk of disease progression and loss of the opportunity for radical surgery. Therefore, there is a pressing need to timely and accurately predict pathologic response to NAC in breast cancer patients to enhance surgical risk stratification and guide individualized treatment.

At present, the gold standard for assessing the efficacy of NAC in the clinical setting is the Miller-Payne grading system [5], which involves comparing pre-chemotherapy biopsy specimens with post-chemotherapy surgical pathology specimens. Nevertheless, this method is invasive, time-delayed, and unable to accurately predict the response of breast cancer patients to chemotherapy before or early chemotherapy. Additionally, the “Response Evaluation Criteria in Solid Tumors” measures the effectiveness of NAC by using ultrasound, magnetic resonance imaging (MRI), and mammography to measure tumor size changes [6]. However, this approach is susceptible to the operator’s subjective judgment and clinical experience and is typically evaluated following the completion of the chemotherapy cycle [7]. Hence, a non-invasive method is urgently needed to determine the pathological responsiveness of NAC in breast cancer patients at an early stage, so that patients with poor response can be identified early and subsequently adjust their treatment strategy to optimize patient outcomes.

Radiomics has emerged as a novel tool for non-invasive analysis of tumors by mining a large number of high-throughput “semantic” features from medical images to

establish predictive models that can quantitatively assess and diagnose highly complex and heterogeneous malignant tumors [8]. With the excellent ability of deep network hierarchical structure to analyze medical images, deep learning offers outstanding performance in applications such as auxiliary diagnosis, risk assessment, and prognosis prediction for various diseases [9–12]. Previous studies [13, 14] have demonstrated that deep learning radiomics models (DLR) based on ultrasound images can effectively assess the response of breast tumors or axillary lymph nodes to NAC. However, most of the aforementioned studies exclusively generated models based on single-point ultrasound images or a single machine-learning algorithm. Biological behavior is a dynamic ecosystem, and tumor heterogeneity may not be fully captured at a single time point [15, 16]. In addition, Delta deep learning features and radiomics features may provide more complementary information for the prediction of NAC response. Therefore, this study aimed to develop a comprehensive model that combined ultrasound radiomics and deep learning features of pre-NAC, post-treatment 2 cycle, and Delta ultrasound images to evaluate the pathological responsiveness of breast cancer patients to NAC. This strategy has the potential to enable early assessment of NAC response and assist in guiding treatment plans.

Methods

Study population

The Ethics Committee of Binzhou Medical University Hospital accepted this retrospective investigation, and informed consent was waived.

A total of 312 breast cancer patients undergoing NAC between January 2018 and December 2023 were enrolled in the present study. The following were the inclusion criteria: (i) Primary breast cancer confirmed by pathological biopsy prior to treatment; (ii) Complete clinical and pathological information available before and after NAC treatment; (iii) Underwent complete, standardized neoadjuvant therapy and did not receive other therapies prior to neoadjuvant therapy. The following were the exclusion standards: (i) Ipsilateral multifocal, bilateral multiple lesions, or distant metastases during NAC; (ii) Failure to complete a full cycle of neoadjuvant therapy; (iii) Incomplete or unavailable clinicopathological information or

ultrasound images. The enrollment process for the study population is illustrated in Fig. 1.

In order to improve the generalizability of the model, the dataset was randomly divided into a training cohort containing 219 patients (91 pCR, 128 non-pCR) and a test cohort comprising 93 patients (40 pCR, 53 non-pCR). Clinically relevant information, including age, NAC cycle, NAC protocol, and molecular typing, was collected from medical records.

Treatment plan and pathologic evaluation

All patients received 6 or 8 cycles of NAC. The regimens consisted of TEC (docetaxel, doxorubicin, cyclophosphamide) or CEF (cyclophosphamide, epirubicin, fluorouracil). Moreover, trastuzumab was administered to patients who confirmed positive for Her-2.

By using immunohistochemistry (IHC) assays, the tumor type and receptor status were verified. These assays included the expression of Ki-67, human epidermal growth factor receptor-2 (Her-2), progesterone receptor (PR), and estrogen receptor (ER). HR positive was identified as $\geq 1\%$ of nuclear staining of ER or PR based on the IHC index assessment criteria [13]. 20% was chosen as the cutoff index for Ki-67, with $<20\%$ denoting low expression and $\geq 20\%$ denoting high expression. An IHC grade of 3+ indicated her-2 positive. When IHC revealed HER2 expression to be graded 2+, the HER2 gene amplification was confirmed by fluorescence in situ hybridization (FISH) [17]. Based on their receptor status,

all patients were divided into three subtypes: (i) HR+/HER2-; (ii) HER2+; (iii) TNBC (triple negative breast cancer). Based on the pathological evaluation of surgical specimens, the pathological response to NAC was assessed using the Miller-Payne grading system [5], with grades 1–4 were classified as npCR and grade 5 was classified as pCR.

Ultrasound images acquisition

All ultrasonography was performed within 1 week prior to NAC treatment and after the completion of the NAC 2 cycle. Sonographers with over 10 years of experience in breast tumor imaging used GE LOGIQ E9, Aplio i900, and Esaote Mylab Twice color doppler ultrasound diagnostic instruments, each equipped with linear array high-frequency probe, with probe frequencies ranging between 5–13 MHz, 5–18 MHz, and 4–13 MHz, respectively (Supplementary Table 1). The maximum long-axis cross-sectional images of the patient's tumor before treatment were collected and stored in DICOM format. All images remove direct or indirect personally identifiable information, such as name, personal phone number, email address, marital status, personal ID, etc., to protect patient privacy.

Tumor region of interest segmentation and radiomics analysis

Two specialized sonographers used 3D Slicer software (version 4.10.1, www.slicer.org) to manually delineate

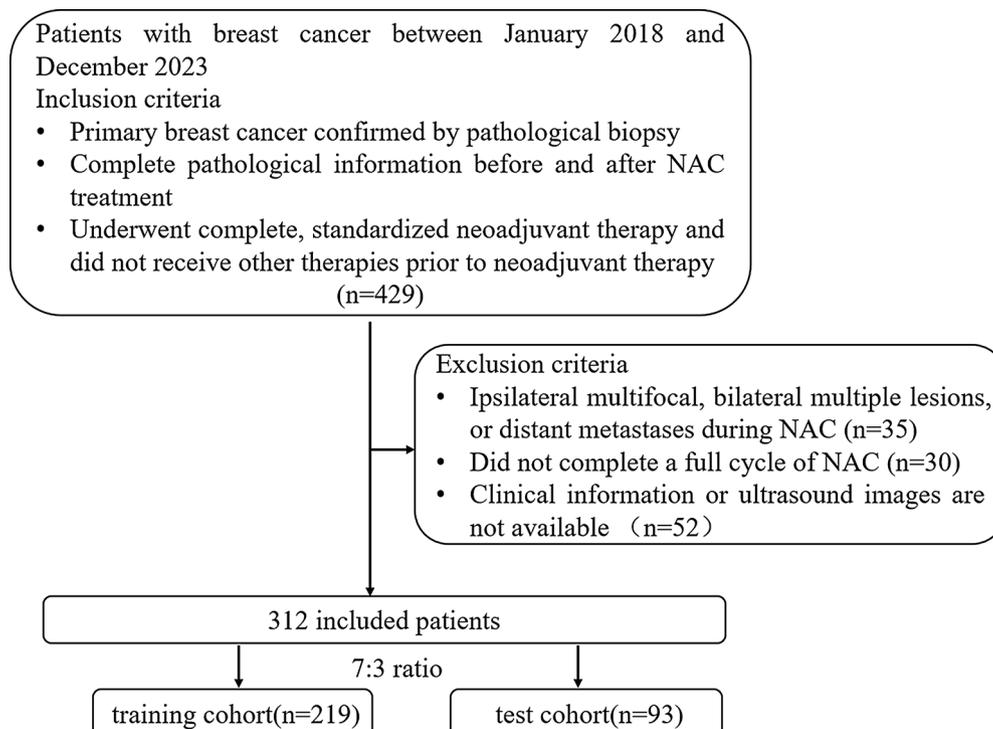


Fig. 1 Flowchart of patient enrollment in the study

regions of interest (ROI) along the lesion boundary on grayscale images, avoiding areas of necrosis, calcification, or hemorrhage. Disagreements were resolved by a senior sonographer (with 20 years of experience). By computing the intraclass correlation coefficient (ICC) for the two volumes of interest segmented by the two sonographers for the same lesion, the repeatability of radiomics features was investigated. The training cohort images were randomly cropped and flipped horizontally to increase data diversity (Supplementary Information 3). All images were resized to 224×224 pixels. The intensity distribution of the ultrasound images varied greatly since they were obtained across various image acquisition devices. All pixel values were normalized to the interval $[0, 1]$ and normalized using the mean and standard deviation of the ImageNet dataset. In addition, Gaussian filtering and histogram averaging methods were used to make the pixel distribution of each intensity level in the image more uniform, so as to minimize the difference between ultrasound devices. Next, the Combat method was used to reduce the batch effect caused by the images acquired by different machines, and the principal component analysis showed that the Combat method did correct the batch effect of the machine, as detailed in Supplementary Fig. 1. Subsequently, the Pyradiomics platform was utilized to extract hand-crafted radiomics features. From the pre-NAC ultrasound images, a total of 1032 radiomics features were extracted; additionally, 1032 radiomics features were extracted based on the post-NAC 2-cycle ultrasound images. In addition, in order to capture longitudinal changes in tumor characteristics, the relative net change between the pre-NAC radiomic eigenvalues and the post-NAC radiomic eigenvalues was calculated as the Delta radiomics feature. A total of 3096 radiomics features were extracted for each patient, including pre-NAC, post-NAC 2 cycle, and Delta radiomics feature sets.

Deep learning analysis

In the present study, the VGG19 pre-trained model was selected for transfer learning, which included 16 convolutional layers, 5 maximum pooling layers, 3 fully connected layers, and a Softmax layer. During transfer learning, the last fully connected layer of the pre-trained model was substituted with an output layer that adapted to our binary classification task. The remaining layers of the model were frozen, and only the last layer was trained to accelerate training and mitigate overfitting. A stochastic gradient descent (SGD) optimizer and a cross-entropy loss function were used to optimize the model. The initial value of the learning rate was set to 0.003, as detailed in Supplementary Information 2. Rectangular ROIs with bounding boxes were manually extracted around the entire tumor area and its surrounding tissues using 3D Slicer software. All images were resized to 224×224

pixels for integration into the deep convolutional neural network. After training the deep learning model, features from the fully connected layer were extracted as deep learning features. As a result, 4096 deep-learning features were extracted based on the VGG19 network structure. Delta deep learning features were obtained by calculating the relative net change between the pre-NAC and post-NAC deep learning features.

Screening of deep learning features and radiomics features

To ensure that the features most relevant for predicting pCR were retained, the following approaches were employed. To begin, preliminary feature selection was carried out according to the intra-class correlation coefficient, and the six ICC thresholds (0.7–0.95) were compared to determine the feature set that yielded optimal predictive performance. Next, the importance score for each feature was calculated using the Boruta algorithm, wherein “shadow features”, which are randomized copies of feature values, were generated. These shadow features served as benchmarks to compare the importance of real features. The importance score of each real feature was compared to that of the shadow feature. Features significantly more important than shadow features were marked as “Confirmed”. Features with a similar importance score to shadow features had a marginal impact on predictions and were marked as “Rejected”. In the third step, the Least Absolute Shrinkage and Selection Operator (LASSO) regression was utilized to filter the features, with the parameter λ controlling the number of selected features. During model training, a 10-fold cross-validation method was used to select the optimal parameter λ to identify the optimal number of features and concomitantly avoid overfitting.

Model development and validation

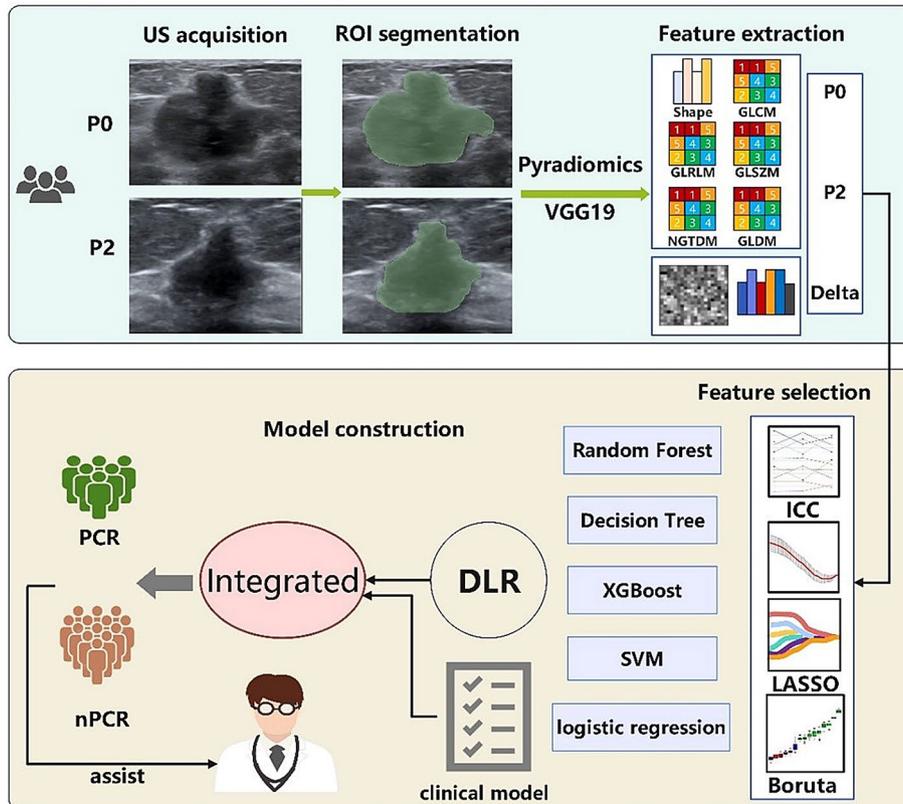
The model was constructed using the training cohort, and multiple cross-validations were carried out. Following this, the test cohort was used to evaluate the predictive performance of the model. Five machine learning algorithms were investigated, namely random forest, decision tree, XGBoost, support vector machine, and logistic regression analysis, as detailed in Supplementary Information 1, 4, 5. For each algorithm, three single-modality models (Pre, Post, Delta) were constructed using the screened radiomics and deep learning features. The best-performing machine learning algorithm was used to integrate the features of different time points and generate a DLR model. Based on the clinicopathological features collected in this study, univariate and multivariate logistic regression analyses were conducted to identify independent predictors and construct a clinical model. A combined model (Integrated) was constructed by integrating independent clinical predictors

and characteristics of NAC from different time points. In order to ensure the robustness of the model, the whole construction process was repeated 1000 times using the bootstrap method. The outline of model establishment is portrayed in Fig. 2.

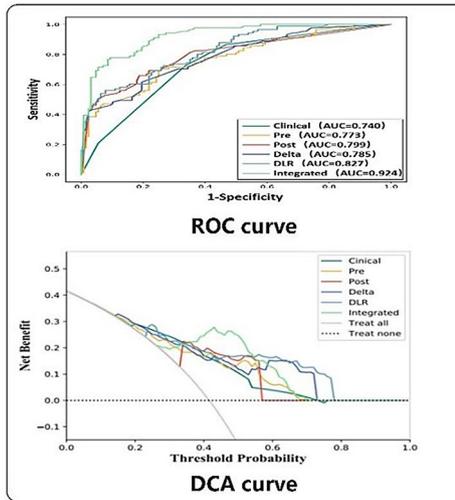
The model was assessed using the decision curve analysis (DCA) and receiver operating characteristic (ROC) curve. The model's sensitivity, specificity, accuracy, and

negative predictive value, and positive predictive value were determined by maximizing the Youden index and determining the ideal cut-off value. In addition, the F1 score, Matthews correlation coefficient, accuracy, and recall indicators were used to evaluate the prediction performance of the model. The Delong test was used to compare the area under the curve (AUC) values of different models.

a. Development of Integrated



b. Performance evaluation



c. Interpretability of the model

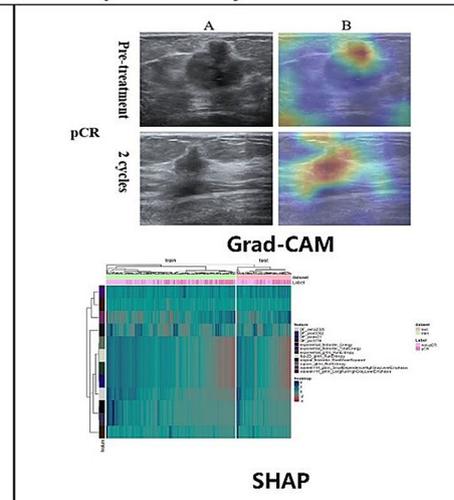


Fig. 2 Workflow of the study. (a) Development of Integrated; (b) Evaluating the predictive performance of Integrated; (c) Interpretability of the model

Visualization and interpretability of the model

In this study, interpretable algorithms such as SHAP and Grad-CAM were used to visualize the decision-making process of the model and to explore the most relevant features and tissue structures in ultrasound images associated with pCR after NAC to provide evidence to support the future clinical applicability of the model. The SHAP algorithm, based on cooperative game theory, measures the importance of features by calculating their contribution values (Shapley value), and Grad-CAM visually interprets the areas that the deep convolutional neural network model focuses on when making predictions to form a “heatmap” and displays the most relevant parts of the image to the model’s decision, thereby addressing “black box” challenges associated with artificial intelligence.

Statistical analysis

Python 3.6.12 and SPSS 26.0 were used for statistical analyses. The t-test was used to compare normally distributed continuous variables, which were reported as mean \pm standard. The Chi-square test or the Fisher exact test were used to compare categorical variables, which were reported as frequencies and percentages. A two-sided $p < 0.05$ was considered statistically significant. The code used in this research has been uploaded to GitHub, see link (https://github.com/shi4180/BreastCancer_NAC_Predictor) for details.

Results

Clinical features

A total of 131 patients achieved pCR, whereas the remaining 181 patients were classified as non-pCR. The pCR rates in the training cohort and the test cohort were 41.6% and 43.0%, respectively. Among the three molecular subtypes, the pCR rate was 29.6% (56/189) for the HR+/HER2- subtype, 61.4% (43/70) for the HER2+ subtype, and 60.4% (32/53) for the TNBC subtype. No significant differences were noted in clinicopathological characteristics between the training cohort and the test cohort ($P > 0.05$) (Supplementary Table 2). There were notable variations between the pCR and non-pCR groups in terms of molecular types, ER, PR, HER2, and Ki-67 expression ($P < 0.05$). The difference in histological grade was statistically significant only in the training cohort ($P < 0.05$) (Table 1).

Model construction

A single radiomics model was developed based on features from pre-NAC and post-NAC 2-cycle time points screened using different ICC thresholds. Six thresholds were tested, and the trained radiomics model displayed superior performance when features with $ICC \geq 0.80$ were used, as displayed in Fig. 3.

After feature screening, 16, 13, and 10 features were selected to construct the Pre model, Post model, and Delta model based on the five machine learning

Table 1 Clinical characteristics of patients in different cohorts

Characteristics	Training cohort (N=219)			Test cohort (N=93)		
	pCR (n=91)	Non-pCR (n=128)	P value	pCR (n=40)	Non-pCR (n=53)	P value
Age, mean \pm SD, years	48.4 \pm 10.13	50.4 \pm 11.12	0.200	50.0 \pm 10.50	49.5 \pm 10.20	0.807
Grade(%)						
I	0(0.00%)	4(3.12%)	0.007	1(2.50%)	0(0.00%)	0.085
II	59(64.8%)	100(78.1%)		24(60.0%)	41(77.4%)	
III	32(35.2%)	24(18.8%)		15(37.5%)	12(22.6%)	
Moduler(%)						
luminalA	2(2.20%)	24(18.8%)	<0.001	3(7.50%)	8(15.1%)	0.036
luminalB	36(39.6%)	70(54.7%)		15(37.5%)	31(58.5%)	
Her2overexpression	32(35.2%)	18(14.1%)		11(27.5%)	9(17.0%)	
Triple-negative	21(23.1%)	16(12.5%)		11(27.5%)	5(9.43%)	
ER status(%)						
Positive	32(35.2%)	89(69.5%)	<0.001	16(40.0%)	38(71.7%)	0.004
Negative	59(64.8%)	39(30.5%)		24(60.0%)	15(28.3%)	
PR status(%)						
Positive	26(28.6%)	76(59.4%)	<0.001	10(25.0%)	33(62.3%)	0.001
Negative	65(71.4%)	52(40.6%)		30(75.0%)	20(37.7%)	
HER-2 status(%)						
Positive	66(72.5%)	44(34.4%)	<0.001	25(62.5%)	15(28.3%)	0.002
Negative	25(27.5%)	84(65.6%)		15(37.5%)	38(71.7%)	
Ki-67 status(%)						
<20%	34(37.4%)	80(62.5%)	<0.001	9(22.5%)	37(69.8%)	<0.001
\geq 20%	57(62.6%)	48(37.5%)		31(77.5%)	16(30.2%)	

pCR, pathologic complete response; ER, estrogen receptor; PR, progesterone receptor; Her-2, human epidermal growth factor receptor-2

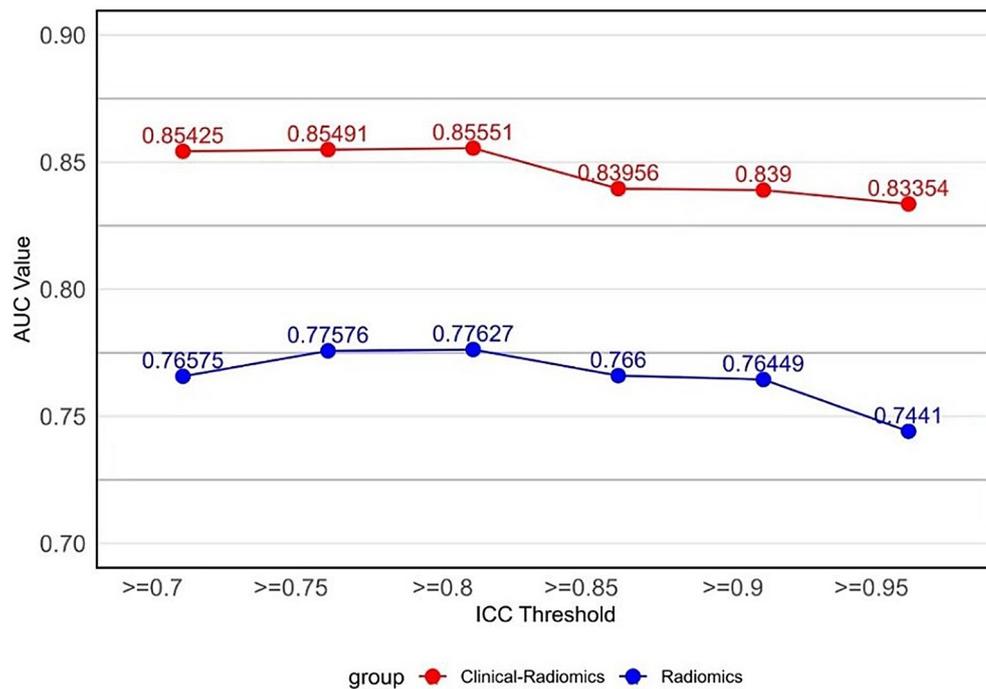


Fig. 3 Performance of the XGBoost model based on different radiomics feature sets screened by different ICC thresholds. ICC = intra-class correlation coefficient; AUC: area under the ROC curve

Table 2 Performance of three single-modality models constructed using five machine learning algorithms for predicting the efficacy of NAC in the training cohort and the test cohort

Model	Features	Training cohort				Test cohort			
		Mean AUC (95% CI)	ACC	SEN	SPE	Mean AUC (95% CI)	ACC	SEN	SPE
XGBoost	Pre	0.773(0.735–0.855)	0.708	0.714	0.703	0.726(0.660–0.732)	0.645	0.775	0.847
	Post	0.799(0.752–0.880)	0.753	0.676	0.891	0.776(0.664–0.881)	0.752	0.700	0.792
	Delta	0.785(0.727–0.846)	0.763	0.703	0.805	0.710(0.667–0.767)	0.697	0.650	0.798
RF	Pre	0.768(0.733–0.845)	0.685	0.660	0.773	0.721(0.690–0.737)	0.699	0.750	0.660
	Post	0.814(0.753–0.863)	0.735	0.626	0.812	0.713(0.650–0.779)	0.688	0.750	0.642
	Delta	0.759(0.722–0.843)	0.772	0.659	0.852	0.707(0.657–0.817)	0.690	0.625	0.830
SVM	Pre	0.744(0.732–0.824)	0.644	0.668	0.784	0.708(0.678–0.720)	0.624	0.688	0.817
	Post	0.705(0.677–0.752)	0.717	0.643	0.883	0.784(0.737–0.808)	0.710	0.750	0.792
	Delta	0.832(0.806–0.890)	0.749	0.648	0.820	0.724(0.669–0.750)	0.677	0.626	0.792
logistics	Pre	0.698(0.617–0.755)	0.653	0.605	0.775	0.679(0.606–0.703)	0.602	0.675	0.736
	Post	0.797(0.772–0.872)	0.753	0.659	0.820	0.771(0.735–0.827)	0.731	0.700	0.755
	Delta	0.783(0.749–0.862)	0.703	0.627	0.828	0.700(0.670–0.733)	0.691	0.641	0.749
DT	Pre	0.696(0.678–0.829)	0.676	0.835	0.797	0.640(0.500–0.730)	0.634	0.875	0.736
	Post	0.740(0.656–0.781)	0.753	0.659	0.820	0.727(0.607–0.727)	0.731	0.650	0.755
	Delta	0.695(0.657–0.755)	0.671	0.825	0.655	0.614(0.579–0.704)	0.602	0.700	0.728

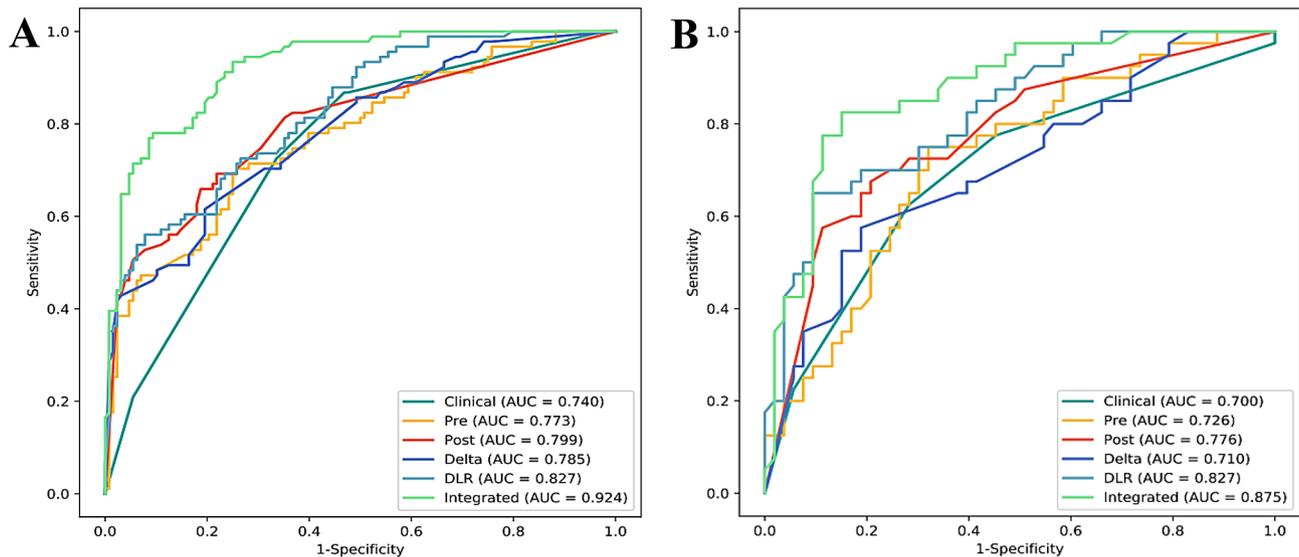
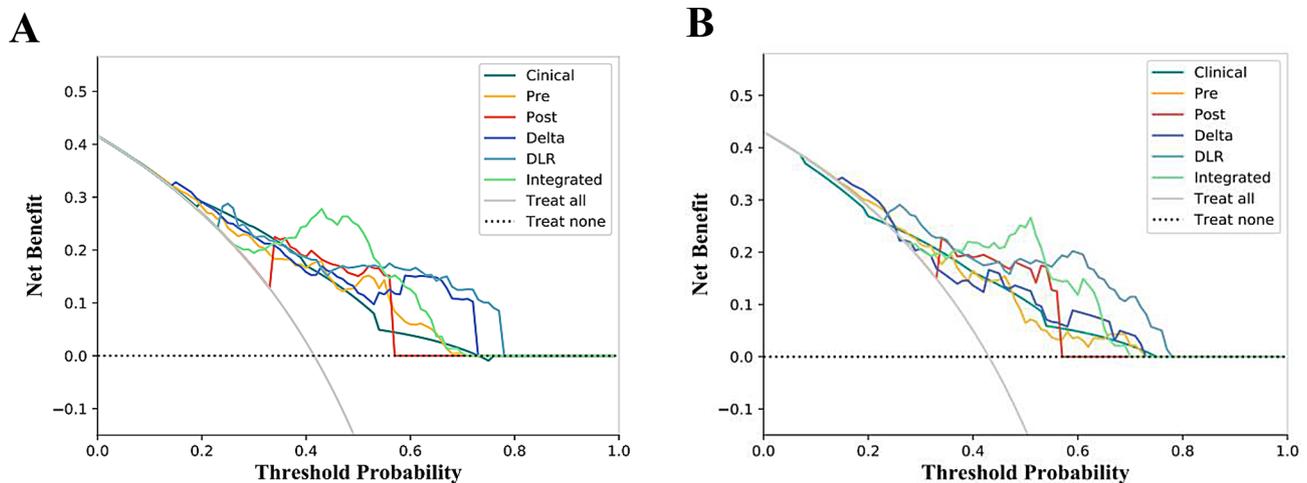
AUC, the area under curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; 95% CI, 95% confidence interval; XGBoost, eXtreme Gradient Boosting; RF, random forest; SVM, support vector machine; logistics, multivariate logistic regression analysis; DT, decision tree

classifiers, respectively (Supplementary Table 3). In the current study, five robust machine learning algorithms were developed to construct deep learning radiomics models, and their performance was compared to determine the optimal model. In the training and test cohorts, three single-modality models based on the XGBoost algorithm outperformed the other classifiers, as detailed

in Table 2. Therefore, based on the XGBoost algorithm, the DLR model was constructed by combining pre, post, and delta deep learning features and radiomics features. After univariate and multivariate logistic regression analysis, histological grade and Her-2 status were identified as independent predictors for the efficacy of NAC. Following this, the DLR model was combined with independent

Table 3 Performance of the DLR and Integrated models constructed using the XGBoost machine learning algorithm in predicting the efficacy of NAC in the training and test cohorts

Cohort	Model	AUC	95%CI	Accuracy	Sensitivity	Specificity
Training cohort	DLR model	0.827	0.817–0.886	0.763	0.665	0.890
	Integrated	0.924	0.898–0.967	0.831	0.786	0.960
Test cohort	DLR model	0.827	0.799–0.836	0.752	0.700	0.792
	Integrated	0.875	0.806–0.877	0.817	0.775	0.849

**Fig. 4** ROC curves for the Clinical, Pre, Post, Delta, DLR, and Integrated models in the training cohort (A) and test cohort (B)**Fig. 5** Decision curves for the Clinical, Pre, Post, Delta, DLR, and Integrated models in the training cohort (A) and test cohort (B)

clinical risk factors to construct the Integrated model. As anticipated, the Integrated model displayed the highest predictive performance (training cohort AUC 0.924, accuracy 0.831, sensitivity 0.786, specificity 0.960, test cohort AUC 0.875, accuracy 0.817, sensitivity 0.775, specificity 0.849) in the training and test cohorts, with its AUC value, accuracy, sensitivity and specificity being better than those of the DLR model (training cohort AUC 0.827, accuracy 0.763, sensitivity 0.665, specificity 0.890,

test cohort AUC 0.827, accuracy 0.752, sensitivity 0.700, specificity 0.792), as shown in Table 3. Figure 4 depicts the AUC of all models across different cohorts. In addition to the above indicators, we also used the F1 score, Matthews correlation coefficient, precision, and recall to evaluate the prediction performance of the model (Supplementary Table 4). Decision curve analysis delineated that the Integrated model provided higher clinical net benefit compared to the other models, as shown in Fig. 5.

Table 4 Performance of the Integrated model in predicting pCR after NAC across various molecular subtypes in the training and test cohorts

Molecular subtype	Model	Training cohort				Test cohort			
		Mean AUC (95% CI)	ACC	SEN	SPE	Mean AUC (95% CI)	ACC	SEN	SPE
HR+/HER2- (n = 112)	Pre	0.764(0.640–0.868)	0.863	0.625	0.792	0.743(0.523–0.796)	0.641	0.714	0.725
	Post	0.833(0.600–0.919)	0.863	0.650	0.877	0.806(0.693–0.878)	0.795	0.708	0.812
	Delta	0.744(0.654–0.783)	0.794	0.625	0.830	0.708(0.590–0.810)	0.709	0.712	0.743
	DLR	0.898(0.847–0.917)	0.890	0.625	0.923	0.846(0.769–0.890)	0.821	0.714	0.843
	Integrated	0.952(0.810–0.955)	0.931	0.752	0.944	0.862(0.753–0.897)	0.897	0.725	0.847
HER2+ (n = 147)	Pre	0.824(0.723–0.836)	0.663	0.681	0.654	0.819(0.739–0.828)	0.625	0.720	0.733
	Post	0.749(0.709–0.798)	0.691	0.682	0.704	0.735(0.687–0.784)	0.675	0.670	0.667
	Delta	0.833(0.790–0.835)	0.709	0.700	0.727	0.811(0.604–0.842)	0.686	0.640	0.766
	DLR	0.866(0.799–0.873)	0.718	0.773	0.809	0.827(0.677–0.845)	0.750	0.680	0.867
	Integrated	0.871(0.774–0.876)	0.790	0.797	0.886	0.879(0.725–0.928)	0.760	0.783	0.877
TNBC (n = 53)	Pre	0.762(0.659–0.814)	0.694	0.796	0.789	0.729(0.680–0.749)	0.714	0.728	0.833
	Post	0.845(0.809–0.876)	0.647	0.801	0.758	0.854(0.760–0.875)	0.716	0.872	0.802
	Delta	0.780(0.720–0.892)	0.722	0.648	0.785	0.719(0.680–0.807)	0.785	0.665	0.855
	DLR	0.827(0.762–0.851)	0.750	0.647	0.842	0.812(0.604–0.895)	0.714	0.750	0.667
	Integrated	0.952(0.829–0.967)	0.843	0.809	0.956	0.896(0.693–0.895)	0.857	0.876	0.863

AUC, the area under curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; 95% CI, 95% confidence interval; HR, hormone receptor; Her-2, human epidermal growth factor receptor-2; TNBC, triple negative breast cancer

Table 5 Results of the Delong test for all models in the training and test cohorts

Models	Clinical	Pre-radiomic	Post-radiomic	Delta-radiomic	DLR	Integrated-model
Clinical	-	0.746	0.292	0.901	0.004*	0.002*
Pre-radiomic	0.454	-	0.343	0.839	0.046*	0.004*
Post-radiomic	0.159	0.494	-	0.295	0.065	0.002*
Delta-radiomic	0.309	0.79	0.664	-	0.036*	0.005*
DLR	0.008*	0.027*	0.141	0.126	-	0.056
Integrated-model	< 0.001*	< 0.001*	< 0.001*	< 0.001*	< 0.001*	-
Training cohort			Test cohort			

Assessing model performance across the three subgroups

Table 4 Summarizes the Integrated model's diagnostic metrics for every subtype in the test and training cohorts. In the training cohort, the AUC of the Integrated model for the HR+/HER2-, HER2+, and TNBC subgroups was 0.952, 0.871, and 0.952, respectively. Indeed, the Integrated model outperformed the DLR model, the pre model, the Post model, and the Delta model, and similar results were obtained in the test cohort. Supplementary Figure 2 shows the ROC curves of all models for each subgroup. Importantly, the accuracy, sensitivity, and specificity of the Integrated model were higher than those of the DLR and the three single-modality models. Meanwhile, the Delong test revealed that the combination of Pre, Post, and Delta deep learning features and radiomics features significantly improved the performance of the model in predicting the pCR of patients after NAC, as detailed in Table 5. Other predictions of the model, such as F1 score, Matthews correlation coefficient, precision, and recall are shown in Supplementary Table 5.

Decision curve analysis was performed to evaluate the clinical benefit of the models across different molecular

subtypes. In the test cohort, the Integrated model exhibited a higher clinical net benefit compared to most other models when the thresholds set between 0.3 and 0.65, 0.25–0.75, and 0.25–0.73 for the HR+/HER2-, HER2+, and TNBC subtypes, respectively. The decision curves of all models for each molecular subtype are displayed in Supplementary Fig. 3.

Interpretability of the model

In this study, the SHAP algorithm was used to quantify the importance of each feature and visually depict the impact of individual features on the predictive results of the model. The feature importance bar chart (Fig. 6A) was generated by averaging the absolute SHAP values for each feature to show the degree of influence of the feature on the final predicted probability. The results unveiled that 8 radiomics features and 4 deep learning features had the greatest impact on the prediction probability of the model. At the same time, the heat map (Fig. 6B) displayed differences in these 12 features between patients in the pCR group and npCR group.

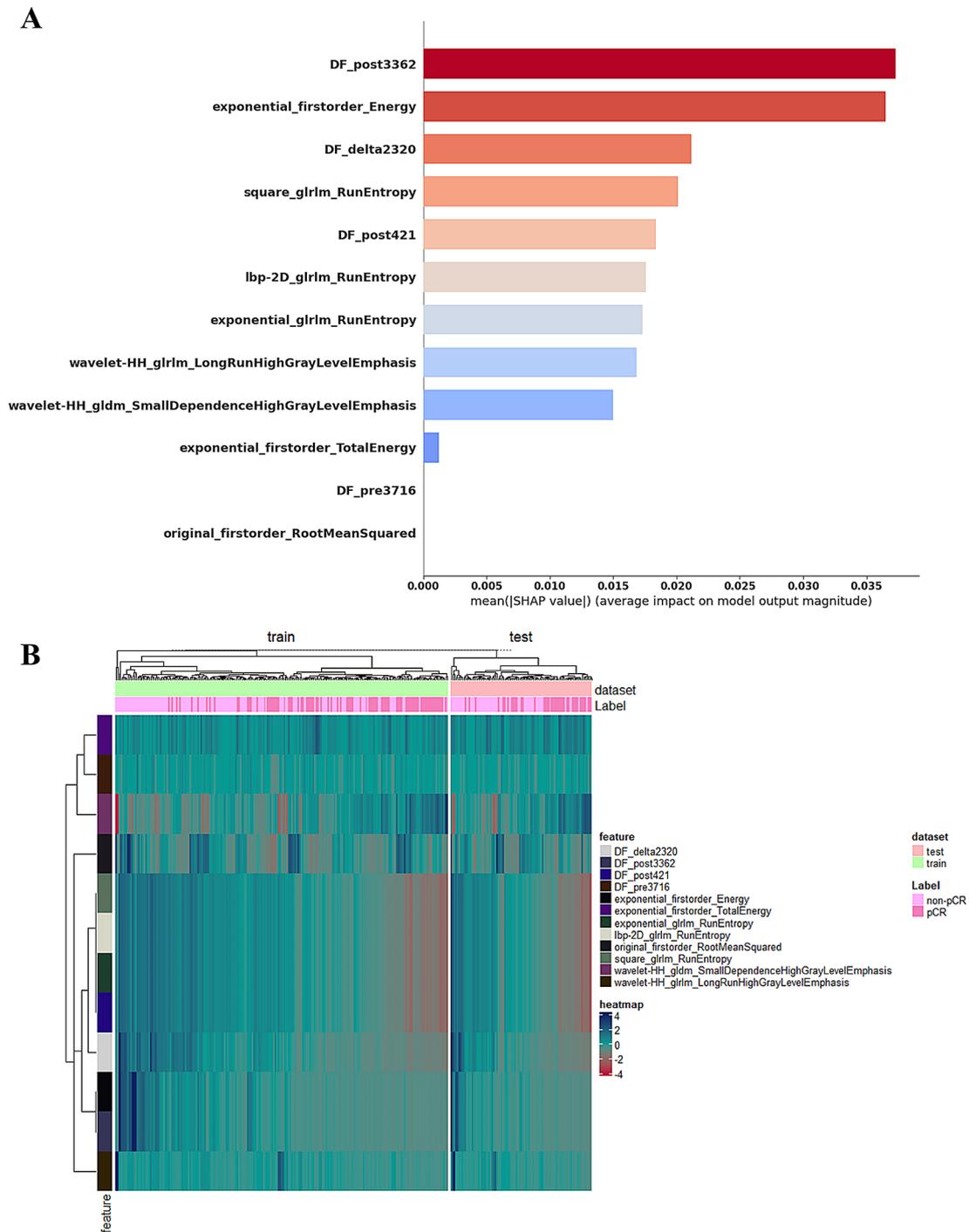


Fig. 6 SHAP Feature Importance Bar Chart (A) and Heat Map (B). SHAP bar chart: SHAP calculates and sorts the features in the random forest model according to their importance, showing the contribution of each feature to the overall prediction. Heat map shows the association of pathologic complete response (pCR) of breast cancer to neoadjuvant chemotherapy with radiomics features and deep learning features in the training and test cohorts

In the final convolutional layer, Grad-CAM determines the weights of each feature map to the image category, computes the weighted sum of each feature map, and then projects the weighted feature image onto the original image as a heat map and visually interprets the prediction results of the model. Grad-CAM showed that the

areas in and near the breast cancer lesions were activated in the patients with correct predictions, confirming that our model effectively identified the target region, and the extracted features correctly reflected relevant information on pCR, consistent with the results of previous studies [14, 18] (Fig. 7A, C). The morphology of the lesions

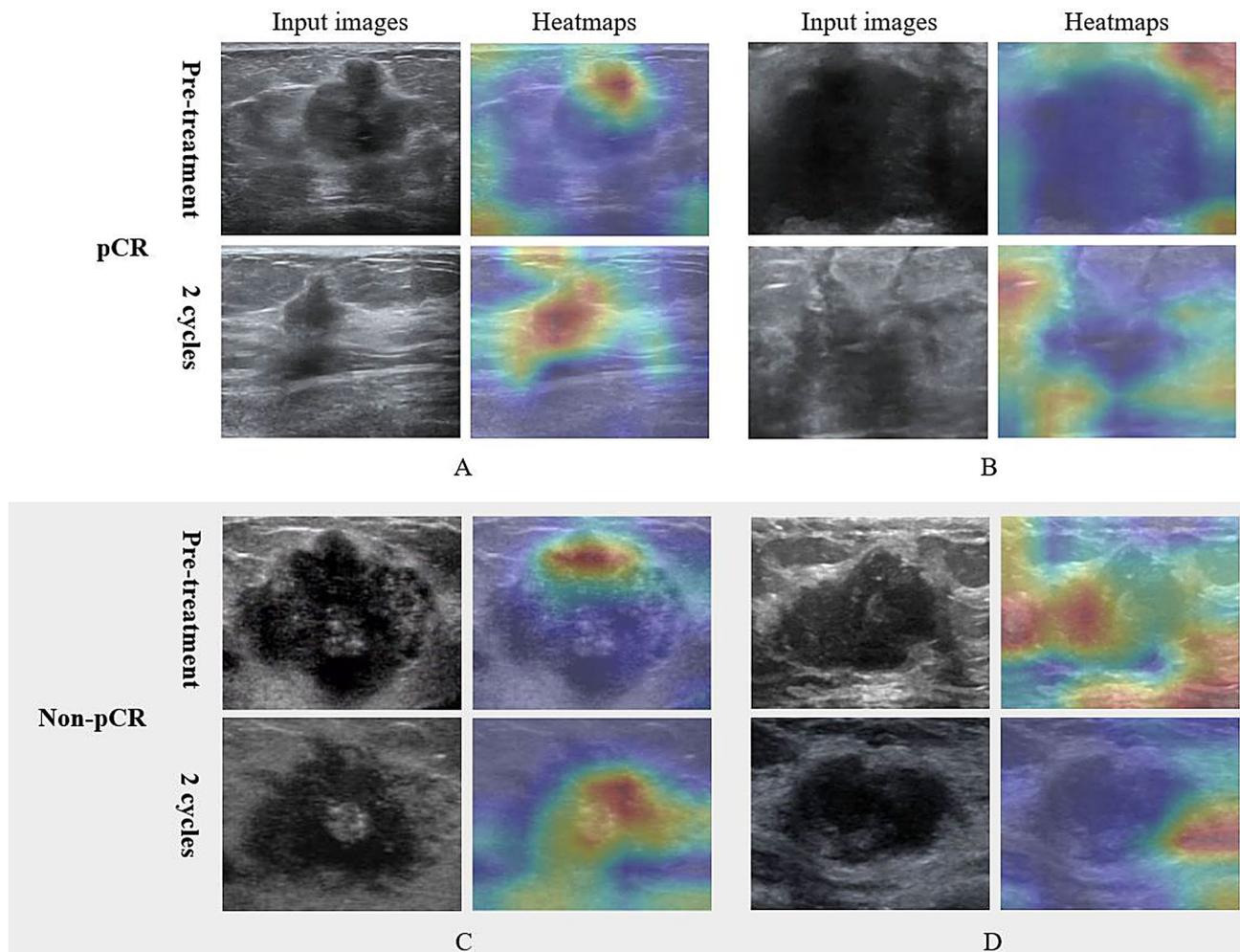


Fig. 7 Grad-CAM visualization provides four predictions for the model: true positives (A), true negatives (C), false positives (D), and false negatives (B). The red and yellow regions represent areas of model activation with the greatest predictive significance for the results, whereas the green and blue regions reflect areas of weaker predictive significance

in the case may be atypical, making it ineffective for the model to capture it. In patients with incorrect predictions, the model failed to accurately identify the surrounding areas within and immediately adjacent to the tumor, but instead identified unrelated areas around the lesion, which may be the reason for the misclassification of Grad-CAM visualization (Fig. 7B, D).

Discussion

Herein, the response to NAC in breast cancer patients was examined using pre-NAC and post-NAC 2-cycle ultrasound data. A combination model that integrated multi-cycle imaging features and clinical data was established. Notably, the results uncovered that the model generated the highest AUC values of 0.924 and 0.875 in the training and test cohorts, respectively. The high performance of the combined model suggests that the model combining multi-cycle ultrasound images and clinical information can effectively predict the tumor response

to NAC and provide valuable insights to guide clinical decision-making.

Of note, images of breast cancer primary lesions contain valuable biological information, and the biological behavior of tumors can be effectively predicted by extracting high-throughput features, converting the images into digital matrices, and then correlating molecular features and clinical prognostic factors. The imaging findings of primary breast cancer prior to treatment are primarily related to tumor characteristics, whereas ultrasound images following NAC treatment directly reflect the response of breast cancer to chemotherapy, such as the presence of hypoxic and fragmented tumor cells, which ultimately culminates in the formation of fibrosis and collagen tissue [19, 20]. Timely identification of these changes may significantly benefit patients. Therefore, two single-modality models, Pre and Post, were constructed based on single-point ultrasound images herein. The results highlighted the high predictive performance

of both models (training cohort AUC 0.773, 0.799, test cohort AUC 0.726, 0.776, respectively).

As is well documented, intra-tumoral heterogeneity drives tumor progression and treatment response [21, 22] and evolves spatially and temporally [23]. Delta radiomics can capture the heterogeneity of change information typically overlooked by single time-point models. While previous studies have demonstrated that delta radiomics features and deep learning features can also provide information for predicting the response of breast masses or axillary lymph nodes to NAC [24–26], the majority of studies were based on MRI images, and studies exploring the use of delta deep learning features based on ultrasound images for predicting response to NAC are scarce. Consequently, the Delta model was constructed based on changes in the internal characteristics of tumors after early treatment. The results signaled that the Delta model could effectively predict pCR in both the training cohort (AUC: 0.785) and the test cohort (AUC: 0.710). Huang et al. [26] used the deep learning features and radiomics features derived from multiparameter MRI sequences before and after treatment to construct a model to predict NAC response and reported AUC values ranging between 0.796 and 0.812, which was higher than the model constructed in this study. This finding may be ascribed to its large sample size and multi-center design, which incorporated three MRI sequences, namely T2-weighted (T2WI) images, dynamic contrast-enhanced (DCE) images, and diffusion-weighted (DWI) images, and acquired more information about the underlying biological behaviors of the tumor and its sensitivity to chemotherapy.

It is worthwhile emphasizing that in this study, the predictive performance of the Post model based on early treatment ultrasound data was higher than that of the Pre and Delta models, highlighting the significant predictive value of ultrasound information obtained during early treatment. In line with the results of Yang et al. [24], these results may be attributed to ultrasound imaging relying on the operational technique of the examiner, resulting in differences between imaging characteristics before and after chemotherapy, such as gray gain scale and image quality. Additionally, the comparison of radiomics properties may be impacted by the presence of necrotic and remaining tumor tissues following treatment. These factors may lead to the unsteady and stochastic in differences between deep learning features and radiomics features before and after chemotherapy, which may result in a decline in the predictive performance of the Delta model.

In this study, there was a significant improvement in model performance when information from different time points during NAC was introduced into the DLR model. The Delong test showed that the AUC value for

predicting pCR using the DLR model was statistically superior to that using the Pre, Delta, and Clinical models. Besides, several clinicopathologic features were associated with NAC response, namely histological grade and molecular typing, as well as ER status, PR status, Her-2 status, and Ki-67 index. After multivariate logistic regression analysis, only histological grade and Her-2 indexes were associated with pCR after NAC. This result may be related to the higher sensitivity of Her-2-positive patients to targeted drugs, which act by inhibiting Her-2 signaling and dimerization [27]. Thus, it is more feasible to achieve pCR after NAC in Her-2-positive patients. A high histological grade reflects the active proliferation of tumor cells and a high degree of malignancy, and chemotherapy drugs play a role in inhibiting tumor cell proliferation and inducing apoptosis. Considering the effects of imaging and clinical information on NAC response, an Integrated model was developed that was capable of accurately predicting pCR after NAC in both the training cohort (AUC 0.924) and the test cohort (AUC 0.875). Gu et al. [28] extracted ultrasound radiomics features from ultrasound images from pre-NAC, NAC 2 cycle, and NAC 4 cycle, constructed two deep learning models, and described that the AUC of the deep learning model based on the ultrasound images of pre-NAC and post-NAC 4 cycles (AUC value of 0.937) was significantly higher than that of the model based on pre-NAC and post-NAC 2 cycle ultrasound images (AUC value 0.812). Wu et al. [29] integrated continuous ultrasound features from different phases of NAC, before, during, and after treatment, with clinicopathological factors to construct a deep learning model for predicting the efficacy of NAC (AUC 0.924). While the Integrated model constructed in this study demonstrated predictive performance comparable to the above-mentioned models, this study was solely based on pre-NAC and post-NAC 2-cycle ultrasound images. Noteworthy, our model can accurately predict pCR after NAC at an early stage and guide the selection and adjustment of patient treatment regimens.

Given that distinct molecular subtypes of breast cancer may lead to varying responses to NAC, patients were further grouped into subgroup analyses in this study. Our results revealed that the Integrated model had the highest predictive performance compared to the other models across the three subtypes and was effective in predicting patient response to NAC, particularly among the HR+/HER2- and TNBC subtypes (AUC 0.952). Previous studies have reported that the HR+/HER2- subtype has a low response rate to NAC, whereas the HER2+ and TNBC subtypes have a higher response rate to NAC [30]. Therefore, for patients with the HR+/HER2- subtype who wish to preserve breast tissue, the Integrated model (training cohort AUC: 0.952, test cohort AUC: 0.862) can assist in identifying patients who can derive

more benefits from NAC at an early stage, thereby sparing them from unnecessary chemotherapy-related toxicities. The predictive model constructed in this study based on ultrasonography had a similar predictive performance to the deep learning model developed by Wu et al. [29], which was based on continuous time-point ultrasound images (HR+/HER2- AUC: 0.899–0.945, HER2+ AUC: 0.876–0.955, TNBC AUC: 0.802–0.932), and outperformed the predictive model based on multiparametric MRI at a single time point developed by Liu et al. (HR+/HER2- AUC: 0.87–0.87, HER2+ AUC: 0.58–0.79, TNBC AUC: 0.79–0.84) [31]. This signifies that multi-stage biological and pathophysiological changes during NAC captured via serial ultrasonography substantially contribute to outcome prediction, which exceeds the information detected by MRI at a single pre-NAC time point.

The most important radiomics features screened in this study include 4 GLRLM, 1 GLDM and 3 firstorder parameters, GLRLM parameters are used to describe the degree of difference in gray value between pixels in the lesion, and it has been proved that the higher the value of some GLRLM features (e.g., *glrlm_LongRunEmphasis*), the rougher and heterogeneous the tumor texture is, and the greater the tumor heterogeneity [32]. The higher the heterogeneity of breast tumors, the more poorly differentiated and malignant, and the lower the efficacy of NAC [33]. The GLDM parameter describes the pairwise queuing of pixels with a set absolute difference in gray scale in a given direction and distance, which is used to highlight the local heterogeneous information, and the *SmallDependenceHighGrayLevelEmphasis* (HGSDE) parameter measures the joint distribution of high gray values and small dependencies in the image, the larger the HGSDE value, the smaller the dependence of the high gray area in the image, the less homogeneous texture of the image, and the more uneven the texture, the greater the tumor heterogeneity. The first-order feature can be based on the global grayscale histogram to reflect the symmetry, uniformity, and local intensity distribution changes of the measured voxels, and can show the degree and extent of intra-tumoral spatial heterogeneity by quantifying the multi-region space of the lesion [34]. Therefore, the above parameters may be the key variables in predicting the efficacy of NAC in breast cancer by suggesting the heterogeneity and roughness of tumor internal tissues.

We also take into account the implementation of the comprehensive model into future clinical practice, and from its performance on the test cohort, all predicted non-pCR patients would be directed to surgery, which would cause 9.7% (9 of 93) overtreatment; all predicted pCR patients might be directed to extended non-invasive biopsy, undertreatment for those patients with the prediction of pCR could be prevented by extended imaging-guided vacuum-assisted biopsy of the tumor bed or

radiation therapy and omitting surgery. However, false negative patients (8/93; 8.6%) with positive biopsy results need to be directed to surgery. Finally, 33.3% (31 of 93) would benefit from this de-escalating concept.

Nevertheless, some limitations of our study merit acknowledgment. Firstly, given that this was a single-center retrospective study with a limited sample size, the possibility of selection bias cannot be excluded. The developed model requires validation from multi-center studies in the future. Secondly, the imbalance in the proportion of molecular subtypes may have compromised the predictive results of the model. Thirdly, this study exclusively extracted the characteristics from the tumor region, and previous studies established that tissues adjacent to tumors could also provide relevant information for the prediction of NAC response. In the future, we will obtain the surrounding tissues of the tumor for comprehensive analysis. Lastly, only ultrasound data were used to develop the model, and the information contained in pathological images and MRI images may improve the performance of our model.

Conclusion

In this study, pre-NAC, post-NAC 2 cycle, and Delta ultrasound data were used to develop the Integrated model, which serves as a non-invasive method to accurately and safely identify breast cancer patients who can achieve pCR after NAC preoperatively. This model provides strong clinical evidence to guide treatment strategies for breast cancer patients.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13058-025-01971-5>.

Supplementary Material 1

Acknowledgements

We are grateful to all the individuals and breast ultrasound experts who participated in this study, and we are grateful for the assistance of breast surgeons at Binzhou Medical University Hospital.

Author contributions

XDF conceptualised the study. FFL and YD provided the data. FFL and GHC contributed funding, resources and methodology. YYX, JY and WJW collected the data. YS conducted the experiments and analysed the data. YS and MW visualised the results. XDF wrote the original draft of the manuscript. All authors checked the data, carefully revised the manuscript, and agreed to submit this publication.

Funding

The study was supported by the Natural Science Foundation of Shandong Province (ZR2023QH231).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

The research ethics committee of the Binzhou Medical University Hospital approved this retrospective study (approval no. LW-60) and waived the requirement for informed consent since it solely utilized pre-existing medical data. This study was conducted in compliance with the ethical standards of the participating institution and adhered to the tenets outlined in the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Ultrasound, Binzhou Medical University Hospital, Binzhou, Shandong 256603, China

²Department of Ultrasonography, Weihai Municipal Hospital, Cheeloo College of Medicine, Shandong University, Weihai, Shandong 264200, China

Received: 19 November 2024 / Accepted: 30 January 2025

Published online: 27 February 2025

References

- Sung H, Ferlay J, Siegel RL, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 Countries[J]. *CA Cancer J Clin*. 2021;71(3):209–49.
- Spring LM, Gupta A, Reynolds KL, et al. Neoadjuvant endocrine therapy for Estrogen receptor-positive breast Cancer: a systematic review and Meta-analysis[J]. *JAMA Oncol*. 2016;2(11):1477–86.
- Romeo V, Accardo G, Perillo T et al. Assessment and Prediction of response to neoadjuvant chemotherapy in breast Cancer: a comparison of imaging modalities and future Perspectives[J]. *Cancers (Basel)*, 2021,13(14).
- Shubeck S, Zhao F, Howard FM, et al. Response to treatment, racial and Ethnic Disparity, and survival in patients with breast Cancer undergoing Neoadjuvant Chemotherapy in the US[J]. *JAMA Netw Open*. 2023;6(3):e235834.
- Ogston KN, Miller ID, Payne S, et al. A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival[J]. *Breast*. 2003;12(5):320–7.
- Wang H, Mao X. Evaluation of the efficacy of neoadjuvant chemotherapy for breast Cancer[J]. *Drug Des Devel Ther*. 2020;14:2423–33.
- De Los SJ, Bernreuter W, Keene K, et al. Accuracy of breast magnetic resonance imaging in predicting pathologic response in patients treated with neoadjuvant chemotherapy[J]. *Clin Breast Cancer*. 2011;11(5):312–9.
- Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *Eur J Cancer*. 2012;48(4):441–6.
- Wang K, Lu X, Zhou H, et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study[J]. *Gut*. 2019;68(4):729–41.
- Vigil N, Barry M, Amini A et al. Dual-intended deep learning model for breast Cancer diagnosis in Ultrasound Imaging[J]. *Cancers (Basel)*, 2022,14(11).
- Jiang M, Zhang D, Tang SC, et al. Deep learning with convolutional neural network in the assessment of breast cancer molecular subtypes based on US images: a multicenter retrospective study[J]. *Eur Radiol*. 2021;31(6):3673–82.
- Guo X, Liu Z, Sun C, et al. Deep learning radiomics of ultrasonography: identifying the risk of axillary non-sentinel lymph node involvement in primary breast cancer[J]. *EBioMedicine*. 2020;60:103018.
- Jiang M, Li CL, Luo XM, et al. Ultrasound-based deep learning radiomics in the assessment of pathological complete response to neoadjuvant chemotherapy in locally advanced breast cancer[J]. *Eur J Cancer*. 2021;147:95–105.
- Yu FH, Miao SM, Li CY, et al. Pretreatment ultrasound-based deep learning radiomics model for the early prediction of pathologic response to neoadjuvant chemotherapy in breast cancer[J]. *Eur Radiol*. 2023;33(8):5634–44.
- Natrajan R, Sailem H, Mardakheh FK, et al. Microenvironmental heterogeneity parallels breast Cancer progression: a histology-genomic integration Analysis[J]. *PLoS Med*. 2016;13(2):e1001961.
- Failmezger H, Muralidhar S, Rullan A, et al. Topological tumor graphs: a graph-based spatial model to Infer Stromal Recruitment for Immunosuppression in Melanoma Histology[J]. *Cancer Res*. 2020;80(5):1199–209.
- Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast Cancer 2013[J]. *Ann Oncol*. 2013;24(9):2206–23.
- Li Y, Fan Y, Xu D, et al. Deep learning radiomic analysis of DCE-MRI combined with clinical characteristics predicts pathological complete response to neoadjuvant chemotherapy in breast cancer[J]. *Front Oncol*. 2022;12:1041142.
- Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response[J]. *Nature*. 2013;501(7467):346–54.
- Pietras K, Ostman A. Hallmarks of cancer: interactions with the tumor stroma[J]. *Exp Cell Res*. 2010;316(8):1324–31.
- O'Connor JP, Rose CJ, Waterton JC, et al. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome[J]. *Clin Cancer Res*. 2015;21(2):249–57.
- Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer Genome landscapes[J]. *Sci*. 2013;339(6127):1546–58.
- Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing[J]. *N Engl J Med*. 2012;366(10):883–92.
- Yang M, Liu H, Dai Q, et al. Treatment response prediction using Ultrasound-based Pre-, Post-early, and Delta Radiomics in Neoadjuvant Chemotherapy in breast Cancer[J]. *Front Oncol*. 2022;12:748008.
- Liu S, Du S, Gao S, et al. A delta-radiomic lymph node model using dynamic contrast enhanced MRI for the early prediction of axillary response after neoadjuvant chemotherapy in breast cancer patients[J]. *BMC Cancer*. 2023;23(1):15.
- Huang Y, Zhu T, Zhang X, et al. Longitudinal MRI-based fusion novel model predicts pathological complete response in breast cancer treated with neoadjuvant chemotherapy: a multicenter, retrospective study[J]. *EClinicalMedicine*. 2023;58:101899.
- Barron AU, Hoskin TL, Day CN, et al. Association of Low nodal positivity rate among patients with ERBB2-Positive or triple-negative breast Cancer and breast pathologic complete response to Neoadjuvant Chemotherapy[J]. *JAMA Surg*. 2018;153(12):1120–6.
- Gu J, Tong T, He C, et al. Deep learning radiomics of ultrasonography can predict response to neoadjuvant chemotherapy in breast cancer at an early stage of treatment: a prospective study[J]. *Eur Radiol*. 2022;32(3):2099–109.
- Wu L, Ye W, Liu Y, et al. An integrated deep learning model for the prediction of pathological complete response to neoadjuvant chemotherapy with serial ultrasonography in breast cancer patients: a multicentre, retrospective study[J]. *Breast Cancer Res*. 2022;24(1):81.
- Earl H, Provenzano E, Abraham J, et al. Neoadjuvant trials in early breast cancer: pathological response at surgery and correlation to longer term outcomes - what does it all mean? [J]. *BMC Med*. 2015;13:234.
- Liu Z, Li Z, Qu J, et al. Radiomics of Multiparametric MRI for Pretreatment Prediction of Pathologic Complete Response to neoadjuvant chemotherapy in breast Cancer: a Multicenter Study[J]. *Clin Cancer Res*. 2019;25(12):3538–47.
- Huang S, Shi K, Zhang Y, et al. Texture analysis of T2-weighted cardiovascular magnetic resonance imaging to discriminate between cardiac amyloidosis and hypertrophic cardiomyopathy[J]. *BMC Cardiovasc Disord*. 2022;22(1):235.
- Shi Z, Huang X, Cheng Z, et al. MRI-based quantification of Intratumoral Heterogeneity for Predicting Treatment response to neoadjuvant chemotherapy in breast Cancer[J]. *Radiology*. 2023;308(1):e222830.
- Gordic S, Wagner M, Zanato R, et al. Prediction of hepatocellular carcinoma response to (90) Yttrium radioembolization using volumetric ADC histogram quantification: preliminary results[J]. *Cancer Imaging*. 2019;19(1):29.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.