

Multichannel Deep Learning Method Based on Multi-Regional DCE-MRI for Predicting Axillary Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer

Jun Liao

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Mengting Xu

People's Hospital of Henan University & Henan Provincial People's Hospital

Ruimin Li

Fudan University Shanghai Cancer Center

Shunian Li

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Ziwei Cao

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Danyang Wang

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Yongli Li

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Meiyun Wang

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Hongna Tan

natan2000@126.com

People's Hospital of Zhengzhou University & Henan Provincial People's Hospital

Research Article

Keywords: Multichannel deep learning, Breast cancer, Axillary lymph node, Pathologic complete response, Neoadjuvant chemotherapy

Posted Date: February 17th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8598485/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Medical Imaging on September 1st, 2026. See the published version at <https://doi.org/10.1186/s12880-026-02736-y>.

1 **Multichannel Deep Learning Method Based on Multi-Regional DCE-MRI for**
2 **Predicting Axillary Pathologic Complete Response to Neoadjuvant**
3 **Chemotherapy in Breast Cancer**

4 **[Abstract] Objectives:** To develop and compare multi-regional 2D and 2.5D deep
5 learning models based on DCE-MR for noninvasive prediction of axillary pathological
6 complete response (pCR) after neoadjuvant chemotherapy (NAC). **Methods:** This
7 retrospective study enrolled 305 patients with invasive breast cancer and ipsilateral
8 ALN metastasis, which were randomly assigned to a training set (n = 214) and a
9 validation set (n = 91). The principal component analysis, U test, Spearman correlation
10 analysis, mRMR and least absolute shrinkage and selection operator were used to select
11 the most significant features. Seven supervised predictive models were constructed
12 based on regions of interest from the primary tumor and ALNs. Model performance
13 was evaluated using the area under the receiver operating characteristic curve (AUC),
14 calibration curve analysis, and decision curve analysis. **Results:** Among the 305
15 patients, axillary pCR accounted for 46.6% (142/305). ER status, tumor MP grade and
16 HER2 status were selected as independent predictors for axillary pCR($P < 0.05$). Both
17 the 2.5D_T and the 2.5D_{T+ALN} models had higher AUC than those of the 2.5D_T and
18 2D_{T+ALN} models in the validation set (AUC: 0.797 vs 0.706, 0.834 vs 0.815). When
19 combining clinicopathological factors, the 2.5D_{T+ALN}+Clinic model achieved the
20 highest performance among all models (AUC=0.861). Model-guided decision-making
21 reduced the estimated rate of unnecessary ALND from 40.7% to 8.8% and increased
22 the overall clinical benefit rate from 59.3% to 74.7%. **Conclusion** The

2.5D_{T+ALN}+Clinic model could accurate, noninvasive prediction of axillary pCR after NAC, which has the potential to support individualized axillary surgical decision-making.

Keywords: Multichannel deep learning; Breast cancer; Axillary lymph node; Pathologic complete response; Neoadjuvant chemotherapy

Introduction

Breast cancer is the most common malignant tumor among women worldwide and remains one of the leading causes of cancer-related mortality in women^[1]. Axillary lymph nodes (ALNs) serve as the primary route of metastasis for breast cancer, and their status is the critical factor in determining the prognosis of patients^[2]. It was reported that the five year survival rate of patients with pathological positive ALNs decreased by approximately 15%^[3]. Therefore, assessing accurately the status of the ALNs is essential for guiding treatment decision and improving patient prognosis. Neoadjuvant chemotherapy (NAC) has been widely adopted for the treatment of ALN-positive breast cancer. In addition to improve the prognosis of patients with positive ALNs, NAC enables tumor downstaging, thereby facilitating less invasive surgical options. Currently, approximately 40%-75% of patients achieved axillary pathological complete response (pCR) following NAC^[4]. However, accurate identifying patients who achieve axillary pCR to avoid unnecessary axillary lymph node dissection (ALND) remains a significant challenge. While ALND is considered the standard primary procedure for assessing axillary pCR, it is associated with complications such as lymphedema, restricted shoulder movement, and sensory abnormalities. To minimize

the morbidity associated with ALND, various minimally invasive surgical alternatives have been developed. Sentinel lymph node biopsy (SLNB) is well-established for patients with initially lymph node-negative disease. However, following NAC, its clinical application is limited due to a high false-negative rate (FNR) of up to 14.2%, resulting from changes in lymphatic drainage and other factors^[5, 6]. Targeted axillary dissection (TAD) has been shown to reduce the FNR to between 4.1% and 10.8% by removing both marked lymph nodes and sentinel lymph nodes ^[7-9], significantly improving diagnostic accuracy. Nevertheless, TAD remains an invasive procedure and cannot fully eliminate the risk of false negatives. Hence, a reliable, noninvasive method that can predict ALN treatment response after NAC is urgently needed, as it would enable individualized axillary management and support more precise decision-making for patient selection and treatment planning.

Dynamic contrast-enhanced MRI (DCE-MRI) provides detailed morphological and hemodynamic information, revealing tumor angiogenesis and heterogeneity. It has become a key imaging tool for evaluating breast cancer treatment response to NAC. However, traditional visual interpretation methods have demonstrated inconsistent performance in predicting axillary lymph node pCR, with their accuracy often limited by tumor heterogeneity and the subjectivity of assessment^[10, 11]. With advances in computer technology, new methods to improve image quality and extract more comprehensive information are being explored. Radiomics models based on primary breast tumors or metastatic ALNs were reported to have certain clinical value in the prediction of axillary pCR after NAC ^[12-14]. However, these methods have limitations,

such as a heavy reliance on manually designed feature selection and suboptimal diagnostic performance.

Deep learning(DL) techniques have demonstrated the capacity to non-invasively extract underlying information from imaging data, processing raw data and generating representations necessary for pattern recognition. These methods have been widely used in tasks such as image segmentation and classification^[15-17], with significant success in evaluating axillary pCR^[18, 19]. However, existing studies typically focus on extracting features from 2D(2D dimensional) or 3D (3D dimensional) images. It is important to note that deep learning models based on 2D images typically analyze specific slices, which may fail to capture the overall tumor characteristics and spatial heterogeneity, potentially omitting critical information. On the other hand, 3D images data require high computational costs and large-scale datasets^[20]. In response to these challenges, the present study employs a novel 2.5D image fusion approach, focusing on the region of interest (ROI) corresponding to the maximum cross-sectional area, along with the surrounding ± 1 and ± 2 slices, resulting in a total of five consecutive slices. By fusing these slices, the 2.5D method maximizes the advantages of 3D data while preserving inter-slice correlation, which is a common limitation of 2D deep learning models. This approach strikes a balance between computational efficiency and information integrity, which is expected to improve the robustness of the model.

To our knowledge, limited studies have explored the clinical value of multi-channel deep learning models in predicting axillary lymph node pCR after NAC in breast cancer patients. Therefore, this study aims to investigate whether 2D and 2.5D

deep learning models based on multi-region MRI data can be used for non-invasive preoperative prediction of axillary pCR in breast cancer patients after NAC.

Materials and Methods

Study Participants

This retrospective study included patients who received NAC and surgery from September 2020 to January 2025, with pathological confirmation of invasive breast cancer and ipsilateral axillary lymph node metastasis. The study was approved by the institutional ethics committee, and patient informed consent was waived. The inclusion criteria were as follows: (1) biopsy-confirmed invasive breast cancer with ipsilateral ALN metastasis; (2) completing standard NAC treatment and following surgery; (3) DCE-MRI scans were available and performed within 2 weeks before NAC; (4) complete clinical and pathological data. The exclusion criteria were as follows: (1) previous history of treatment or prior malignancy; (2) incomplete or non-standard NAC treatment or surgery; (3) lack of clinicopathological data; (4) poor quality of DCE-MRI images; (5) multifocal or bilateral breast tumours or positive with bilateral ALNs. Clinicopathological data were reviewed and collected from the electronic medical record system. The detailed patient enrolment process is shown in Fig. 1.

Pathological Assessment

Breast cancer and ALN metastasis were diagnosed via pathological examination of core needle biopsy samples before NAC. Estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) status and Ki-67 proliferation index were determined by immunohistochemistry and fluorescence in situ

hybridization techniques. Tumors with $\geq 1\%$ nuclear staining were defined as ER/PR positive, and a Ki-67 index cutoff of 20% was used, with $< 20\%$ indicating low expression and $\geq 20\%$ indicating high expression^[21]. HER2 status was defined as negative for an IHC score of 0 or 1+, positive for a score of 3+, and confirmed by FISH for IHC 2+ cases^[22]. Following NAC, all patients underwent ALND in conjunction with either mastectomy or breast-conserving surgery. Axillary lymph nodes pCR was defined as the absence of residual invasive tumor cells in all resected ALNs. The pathological results from ALND served as the gold standard for evaluating the accuracy and effectiveness of the model developed in this study.

Image Acquisition and Preprocessing

DCE-MRI scans were performed within two weeks prior to NAC using 1.5- or 3.0- Tesla MRI scanners, with detailed scanning parameters shown in Table 1. The third phase of the DCE-MRI was selected as the target phase due to its clear contrast between tumor and normal breast tissue, facilitating accurate delineation of the tumor area. Contrast agents, including gadobenate dimeglumine or gadopentetate dimeglumine, were administered at a dose of 0.1 mmol/kg and an injection rate of 2.5 mL/s. Image preprocessing steps included N4 bias field correction to mitigate signal distortion caused by magnetic field inhomogeneity, and image normalization to enhance image uniformity and contrast, minimizing inter-scan variability. All images were resampled to a fixed resolution of $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ for spatial normalization.

Tumor and ALNs Segmentation

Tumor and ALNs regions of interest on DCE-MRI were manually delineated by

two experienced radiologists using ITK-SNAP software (version 3.6, www.itk-snap.org). Areas such as necrosis and cystic changes were excluded to ensure accurate delineation. The delineated lymph node ROI encompassed all enlarged lymph nodes on the same side, capturing features that fully reflect the characteristics of the ALN. Radiologist A, with 5 years of experience in breast MRI interpretation, initially performed the delineation for all patients, whereas Radiologist B, with 7 years of experience, randomly selected 20 patients for repeated delineation. Radiologist A then repeated the delineation for these 20 patients after 2 weeks to verify the consistency of the features. Discrepancies in annotations were resolved through consensus, with a senior radiologist with 20 years of experience providing final interpretation. All radiologists were blinded to clinicopathological data.

Deep Learning Analysis

The OKT-crop_max_roi tool from the OnekeyAI platform was used for image cropping. For each ROI, the slice with the maximum cross-sectional area was selected as the central slice. Adjacent slices at ± 1 and ± 2 positions relative to the central slice were extracted and stacked along the channel dimension, resulting in a five-slice multi-channel input, which was treated as 2.5D data. For comparison, the central slice alone was used to construct the corresponding 2D input.

Deep learning model was trained using the DenseNet121 architecture. The network was pretrained on the ImageNet dataset and subsequently fine-tuned on the DCE-MRI dataset of this study using a transfer learning strategy. The entire dataset was randomly divided into training and validation sets at a ratio of 7:3. Five-fold cross-

validation was performed within the training set to optimize model hyperparameters, while the validation set was held out and not involved in model training or hyperparameter tuning. The overall technical workflow is illustrated in Fig. 2.

All images were resampled to a spatial resolution of 224×224 pixels using bilinear interpolation. Intensity normalization was performed using min–max scaling followed by Z-score normalization to reduce inter-scan variability. Data augmentation strategies, including random cropping and horizontal and vertical flipping, were applied to the training set only to enhance sample diversity and mitigate overfitting. A cosine annealing learning rate schedule was adopted to dynamically adjust the learning rate and improve training stability and generalization. The learning rate was initialized at 0.01, and stochastic gradient descent (SGD) with softmax cross-entropy loss was used for model optimization.

Deep learning features were extracted from the penultimate layer of the DenseNet121 network. Principal component analysis (PCA) was then applied to reduce feature dimensionality. PCA was fitted using the training set only, and the learned transformation was subsequently applied to the validation set to avoid data leakage. This process resulted in a total of 256 deep learning features per sample, derived from the tumor and axillary lymph node ROIs.

Feature Selection

To select the DL features most correlated with the axillary pCR outcome, a multi-step feature selection process was designed and implemented. First, all features were standardized using Z-score to eliminate scale influence, enabling comparison on the

same scale. A univariate analysis was performed using the Mann-Whitney U test, retaining features with significant differences between groups ($P < 0.05$). Collinearity among features was evaluated using Spearman correlation analysis, and any pair of features with an absolute correlation coefficient greater than 0.9 had one feature randomly removed. The Max-Relevance and Min-Redundancy (mRMR) algorithm was then applied to select feature combinations that were highly correlated with the target variable and had minimal internal redundancy. Finally, LASSO regression was used for the final feature selection. Through 10 repeated five-fold cross-validation, the optimal regularization parameter (λ), minimizing the mean squared error (MSE), was determined, and features with non-zero coefficients under this parameter were selected as the final input feature subset. This process allowed for the retention of only those DL features most relevant to model fitting.

Model Construction and Evaluation

Model construction was formulated as a supervised learning task using axillary pCR and non-pCR as outcome labels. Independent clinicopathologic predictors were identified using univariable and multivariable logistic regression analyses. Variables with statistical significance ($P < 0.05$) in multivariable logistic regression were incorporated to construct the clinical model for predicting axillary pCR after NAC. Subsequently, five robust machine learning algorithms—logistic regression, SVM, XGBoost, LightGBM, and multi-layer perceptron (MLP) neural network—were evaluated. The optimal algorithm was selected to develop a series of prediction models, including 2D and 2.5D single-region models, combined tumor–ALN models, and a

fusion model. Specifically, these models comprised a 2D tumor (2D_T) model and a 2D ALN (2D_{ALN}) model; a 2.5D tumor (2.5D_T) model and a 2.5D ALN (2.5D_{ALN}) model; two combined tumor–ALN models (2D_{T+ALN} and 2.5D_{T+ALN}); and a comprehensive fusion model integrating 2.5D tumor and ALN features with clinicopathologic variables (2.5D_{T+ALN}+Clinic). Model predictive performance was evaluated using the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The DeLong test was used to compare AUCs among different models. Decision curves and calibration curves analysis were applied to assess clinical benefit and agreement between predicted and observed outcomes.

Statistical Analysis

All statistical analyses were performed using R software (version 3.5.0) and Python (version 3.6.6). Continuous variables were summarized as means $\pm$ standard deviations, and categorical variables were described as frequencies and percentages. The normal or non-normal distribution of continuous variables was determined by the Kolmogorov–Smirnov test. The homogeneity of continuous variance was tested by the Levene test. Continuous variables were compared using the student t-test or Mann-Whitney U test, while the categorical variables were analyzed using the χ^2 test or Fisher's exact test. Univariate and multivariate logistic regression analyses were performed to identify independent predictors for axillary pCR. The 95%CI of AUC was calculated using the bootstrap method (with 1000 bootstrap intervals). A two-sided $p < 0.05$ was considered statistically significant.

Results

Patient Characteristics

A total of 305 patients with breast cancer and ipsilateral ALN metastasis were enrolled in this study (mean age, 47.6 ± 9.5 years). Patients with axillary pCR accounted for 49.1% (105/214) and 40.7% (37/91) of the training and validation sets, respectively. There were significant differences in ER status, PR status, tumor Miller–Payne (MP) grade, HER2 status and molecular subtypes between patients with and without axillary pCR in the training set (all $P < 0.05$). The detailed clinicopathological characteristics of the patients are summarized in Table 2.

Clinical Model

ER status (odds ratio [OR]: 0.296; 95% CI: 0.148–0.593), tumor MP grade (OR: 2.687; 95% CI: 1.486–4.860) and HER2 status (OR: 5.525; 95% CI: 3.037–10.044) were selected as independent predictors for axillary pCR using logistic regression analysis, which were used to construct the clinical predictive model (Table 3). The clinical model achieved AUCs of 0.838 (95% CI: 0.783–0.892) and 0.698 (95% CI: 0.588–0.809) in the training and validation sets, respectively.

Model Performance

The performance of different models were illustrated in Table 4 and Fig. 3. In our study, the 2.5D_T model achieved AUCs of 0.784 (95% CI: 0.724–0.844) and 0.797 (95% CI: 0.706–0.888) in the training and validation sets, respectively, which outperformed that of the 2D_T model (AUCs: 0.727 and 0.706, respectively). While the 2.5D_{ALN} model demonstrated lower discriminative performance than the 2D_{ALN} model

in both the training and validation sets (AUCs: 0.764 vs 0.801, 0.742 vs 0.757, respectively). For the combined models, the 2.5D_{T+ALN} model showed a numerically higher AUC than 2D_{T+ALN} model in the validation set (0.834 vs 0.815). Given the strong performance of the 2.5D_{T+ALN} model, the independent predictors of axillary pCR were further integrated to construct the 2.5D_{T+ALN}+Clinic model, in which all retained features showed strong associations with axillary pCR (Spearman $r > 0.8$; $P < 0.05$) (Fig. 4). The 2.5D_{T+ALN}+Clinic model demonstrated overwhelmingly higher discriminative performance among all models, with AUC, accuracy, sensitivity, specificity, PPV and NPV were 0.861(95%CI: 0.786~0.937), 81.3%, 75.7%, 85.2%, 77.8% and 83.6% in the validation set, respectively.

Decision curve analysis indicated that the fusion model consistently achieved superior clinical benefit compared with the clinical model and both single-region and multiregional deep learning models across most threshold ranges (Fig. 5). Calibration curve analysis showed that the observed values of ALN status after NAC were well aligned with the predicted values (Fig. 6).

Clinical Benefit Assessment

As shown in Fig. 7, among the 91 patients in the validation set, 37 (40.7%) cases were pathologically confirmed as axillary pCR, and 54 (59.3%) cases as axillary non-pCR. Of the patients who achieved axillary pCR, 29 cases (78.4%) were accurately identified, while 39 cases (72.2%) of those with axillary non-pCR were accurately identified. If ALND had been performed based on the 2.5D_{T+ALN}+Clinic model results for these patients, the rate of unnecessary ALND would have decreased from 40.7% to

8.8%, and the final benefit rate would have increased from 59.3% to 74.7%.

Discussion

Accurate preoperative assessment of axillary pCR is essential for guiding post-NAC surgical decision-making and for minimizing overtreatment. In our study, we compared multi-regional 2D and 2.5D deep learning models based on pretreatment DCE-MRI to predict axillary pCR. Our results indicated that integrating tumor and ALN information using a 2.5D fusion strategy yields improved predictive performance compared with corresponding 2D models. Furthermore, incorporation of clinicopathologic characteristics led to additional performance gains, with the 2.5D_{T+ALN}+Clinic model achieving AUCs of 0.974 and 0.861 in the training and validation sets, respectively. Notably, decision-making guided by this model may reduce unnecessary ALND and improve overall clinical benefit.

In this study, it was observed that ER-negative tumors (OR=0.296), HER2-positive tumors (OR=5.525), and breast pCR (OR=2.687) were more likely to be associated with axillary pCR after NAC. These findings are consistent with the well-established biological paradigm in which ER-positive tumors, which are typically characterized by lower proliferative activity, are relatively less sensitive to cytotoxic chemotherapy^[23, 24]. Furthermore, patients with HER2-positive breast cancer, who are commonly treated with anti-HER2-targeted therapy in combination with chemotherapy, are more likely to achieve axillary ypN0 status^[25], thereby representing a subgroup that may benefit from axillary surgical de-escalation strategies. In addition, tumor MP grade, which serves as a quantitative indicator of the histopathologic response of the primary

tumor to NAC, demonstrated high sensitivity when breast pCR was used as a predictive factor to identify patients who achieved axillary pCR after NAC^[26]. This finding indicates that treatment response in the primary tumor is closely associated with clearance of metastatic ALNs, suggesting a high degree of biological concordance between the two in terms of chemosensitivity, tumor biological behavior, and response to NAC^[27, 28]. Notably, although PR status and molecular subtype were associated with axillary pCR in univariable analyses, they did not retain independent significance after adjustment for ER and HER2 status, suggesting potential information overlap or collinearity among these variables.

Previous deep learning studies focusing on ALNs status have predominantly employed 2D models, which have demonstrated favorable predictive performance but relatively limited specificity^[29], findings that are consistent with those observed in our study. In contrast, 2.5D deep learning models are capable of capturing spatial features across multiple slices, which may facilitate the extraction of more discriminative features and improve model robustness. In a previous study, a deep learning model constructed using consecutive slices of breast tumors achieved an AUC of 0.826 in the validation set for early breast cancer prediction^[30]. However, because this model was not directly compared with a corresponding 2D model, it remains unclear whether the spatial contextual information introduced by consecutive slices independently contributed to the observed performance improvement, and the true incremental benefit of multi-slice modeling over conventional 2D approaches could not be fully determined. In our study, the 2.5D_T model achieved a higher AUC than the corresponding 2D_T

model in the validation set (0.797 vs 0.706), whereas the opposite trend was observed for models based solely on ALNs. This may be attributable to the larger size and better inter-slice continuity of tumor ROIs, which enable 2.5D models to more effectively capture spatial heterogeneity and treatment-induced changes through multi-slice integration. In contrast, ALNs are smaller, less consistently localized across slices, and more prone to signal contamination from surrounding tissues, which may diminish the advantages of 2.5D modeling and lead to slightly reduced performance. Consistent with previous findings reported by Gao et al.^[31], combined models integrating information from both the primary tumor and ALNs demonstrated superior predictive performance compared with single-region models. In this study, the 2.5D_{T+ALN} and 2D_{T+ALN} models achieved AUCs of 0.834 and 0.815 in the validation set, respectively, indicating that simultaneous incorporation of tumor and nodal information can substantially enhance predictive performance. Liu et al.^[32] developed a model for predicting axillary pCR by learning longitudinal representations from ultrasound images of both the primary tumor and ALNs, achieving AUCs ranging from 0.950 to 0.971, which were higher than those obtained in this study. This discrepancy may be attributable to the dynamic nature of neoadjuvant chemotherapy, during which both primary tumors and metastatic lymph nodes undergo continuous morphologic and functional changes. In contrast, this study primarily relied on imaging features extracted at a single pretreatment time point and did not incorporate longitudinal analysis of treatment response.

In our study, incorporation of clinicopathologic characteristics into the best-performing deep learning model resulted in the 2.5D_{T+ALN}+Clinic model, which

demonstrated superior predictive performance compared with all other models. Previous studies have also explored approaches that combine clinicopathologic characteristics with multiregional imaging for predicting axillary pCR. For example, Li et al.^[33] proposed an ensemble prediction system based on multiregional MRI that achieved encouraging performance; however, its accuracy was slightly lower than that observed in our study. This difference may be attributable to the fact that their approach primarily relied on single-slice images and did not account for inter-slice spatial interactions. In recent years, SLNB has been adopted for axillary staging after NAC in patients with clinically node-positive breast cancer. However, assessment of axillary status in current clinical practice still largely relies on invasive approaches. Moreover, several clinical studies have reported overall false-negative rates of 12.6%–14.2% for SLNB^[34, 35], which exceed the clinically acceptable threshold of 10%^[36]. In our study, the proposed 2.5D_{T+ALN}+Clinic model correctly identified 78.4% (29/37) of patients with axillary pCR and 72.2% (39/54) of patients with axillary non-pCR, with a false-negative rate of 8.8%. These findings suggest that this model may serve as an accurate and noninvasive tool for predicting axillary pCR, with the potential to reduce false-negative results, facilitate omission of unnecessary ALND, and thereby decrease surgery-related complications.

Several limitations of this study should be acknowledged. Firstly, as a single-center retrospective study, it may be subject to selection bias, and prospective validation is warranted. In future work, this limitation will be addressed through strengthened collaboration with other institutions. Secondly, although 2.5D data were used for model

training, this approach may not fully capture all intralesional details across slices; therefore, future studies will explore deep learning models based on 3D data. Finally, a standard breast MRI protocol was adopted in this study, which may provide limited axillary coverage. Future investigations will consider the incorporation of multimodal imaging and dedicated axillary imaging sequences to enable a more comprehensive assessment of axillary pCR.

In conclusion, multi-regional and multi-channel deep learning strategies demonstrated superior performance for predicting axillary pCR compared with conventional 2D deep learning strategies. The proposed 2.5D_{T+ALN}+Clinic model, which integrates multi-regional 2.5D deep learning features with clinicopathologic characteristics, demonstrated potential for noninvasive prediction of axillary pCR and for guiding axillary surgical management. Moreover, model-guided decision-making was associated with a reduction in unnecessary ALND and an increase in the overall clinical benefit, supporting the clinical applicability of this approach.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024,74(3):229-63. DOI:10.3322/caac.21834.
2. Mousavi-Kiasary SMS, Bayat M, Abbasvandi F, et al. Tumor characteristics and survival rate of axillary metastatic breast cancer patients: a three decades retrospective cohort study. *Sci Rep.* 2025,15(1):4571. DOI:10.1038/s41598-024-84115-7.

- 375 3. Mangone L, Marinelli F, Bisceglia I, et al. Five-year relative survival by stage of breast and colon
376 cancers in northern Italy. *Front Oncol.* 2022,12:982461. DOI:10.3389/fonc.2022.982461.
- 377 4. Pilewskie M, Morrow M. Axillary Nodal Management Following Neoadjuvant Chemotherapy: A
378 Review. *JAMA Oncol.* 2017,3(4):549-55. DOI:10.1001/jamaoncol.2016.4163.
- 379 5. Aragon-Sanchez S, Oliver-Perez MR, Madariaga A, et al. Accuracy and Limitations of Sentinel
380 Lymph Node Biopsy after Neoadjuvant Chemotherapy in Breast Cancer Patients with Positive Nodes.
381 *Breast J.* 2022,2022:1507881. DOI:10.1155/2022/1507881.
- 382 6. Kuehn T, Bauerfeind I, Fehm T, et al. Sentinel-lymph-node biopsy in patients with breast cancer
383 before and after neoadjuvant chemotherapy (SENTINA): a prospective, multicentre cohort study. *Lancet*
384 *Oncol.* 2013,14(7):609-18. DOI:10.1016/s1470-2045(13)70166-9.
- 385 7. Kuemmel S, Heil J, Rueland A, et al. A Prospective, Multicenter Registry Study to Evaluate the
386 Clinical Feasibility of Targeted Axillary Dissection (TAD) in Node-positive Breast Cancer Patients. *Ann*
387 *Surg.* 2022,276(5):e553-e62. DOI:10.1097/sla.0000000000004572.
- 388 8. Banys-Paluchowski M, Gasparri ML, de Boniface J, et al. Surgical Management of the Axilla in
389 Clinically Node-Positive Breast Cancer Patients Converting to Clinical Node Negativity through
390 Neoadjuvant Chemotherapy: Current Status, Knowledge Gaps, and Rationale for the EUBREAST-03
391 AXSANA Study. *Cancers (Basel).* 2021,13(7). DOI:10.3390/cancers13071565.
- 392 9. de Boniface J, Frisell J, Kühn T, et al. False-negative rate in the extended prospective TATTOO trial
393 evaluating targeted axillary dissection by carbon tattooing in clinically node-positive breast cancer
394 patients receiving neoadjuvant systemic therapy. *Breast Cancer Res Treat.* 2022,193(3):589-95.
395 DOI:10.1007/s10549-022-06588-2.
- 396 10. Gatenby RA, Grove O, Gillies RJ. Quantitative imaging in cancer evolution and ecology. *Radiology.*

397 2013,269(1):8-15. DOI:10.1148/radiol.13122697.

398 11. van Hemert AKE, van Duijnhoven FH, Vrancken Peeters M. This house believes that: MARI/TAD
399 is better than sentinel node biopsy after PST for cN+ patients. *Breast.* 2023,71:89-95.
400 DOI:10.1016/j.breast.2023.06.011.

401 12. Liu S, Du S, Gao S, et al. A delta-radiomic lymph node model using dynamic contrast enhanced
402 MRI for the early prediction of axillary response after neoadjuvant chemotherapy in breast cancer
403 patients. *BMC Cancer.* 2023,23(1):15. DOI:10.1186/s12885-022-10496-5.

404 13. Lin Y, Wang J, Li M, et al. Prediction of breast cancer and axillary positive-node response to
405 neoadjuvant chemotherapy based on multi-parametric magnetic resonance imaging radiomics models.
406 *Breast.* 2024,76:103737. DOI:10.1016/j.breast.2024.103737.

407 14. Li YL, Wang LZ, Shi QL, et al. CT Radiomics for Predicting Pathological Complete Response of
408 Axillary Lymph Nodes in Breast Cancer After Neoadjuvant Chemotherapy: A Prospective Study.
409 *Oncologist.* 2023,28(4):e183-e90. DOI:10.1093/oncolo/oyad010.

410 15. You J, Huang Y, Ouyang L, et al. Automated and reusable deep learning (AutoRDL) framework for
411 predicting response to neoadjuvant chemotherapy and axillary lymph node metastasis in breast cancer
412 using ultrasound images: a retrospective, multicentre study. *EClinicalMedicine.* 2024,69:102499.
413 DOI:10.1016/j.eclinm.2024.102499.

414 16. Sun Y, Guan M, Chen X, et al. Deep learning-based classification and segmentation of interictal
415 epileptiform discharges using multichannel electroencephalography. *Epilepsia.* 2025,66(9):3398-410.
416 DOI:10.1111/epi.18463.

417 17. Rayed ME, Islam SMS, Niha SI, et al. Deep learning for medical image segmentation: State-of-the-
418 art advancements and challenges. *Informatics in Medicine Unlocked.* 2024,47:101504.

DOI:<https://doi.org/10.1016/j.imu.2024.101504>.

18. Huang JX, Lu Y, Tan YT, et al. Elastography-based AI model can predict axillary status after neoadjuvant chemotherapy in breast cancer with nodal involvement: a prospective, multicenter, diagnostic study. *Int J Surg*. 2025;111(1):221-9. DOI:10.1097/js9.0000000000002105.

19. Gu J, Tong T, Xu D, et al. Deep learning radiomics of ultrasonography for comprehensively predicting tumor and axillary lymph node status after neoadjuvant chemotherapy in breast cancer patients: A multicenter study. *Cancer*. 2023;129(3):356-66. DOI:10.1002/cncr.34540.

20. Avesta A, Hossain S, Lin M, et al. Comparing 3D, 2.5D, and 2D Approaches to Brain Image Auto-Segmentation. *Bioengineering (Basel)*. 2023;10(2). DOI:10.3390/bioengineering10020181.

21. Gradishar WJ, Moran MS, Abraham J, et al. NCCN Guidelines® Insights: Breast Cancer, Version 4.2023. *J Natl Compr Canc Netw*. 2023;21(6):594-608. DOI:10.6004/jnccn.2023.0031.

22. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol*. 2013;31(31):3997-4013. DOI:10.1200/jco.2013.50.9984.

23. Huppert LA, Wolf D, Yau C, et al. Pathologic complete response (pCR) rates for patients with HR+/HER2- high-risk, early-stage breast cancer (EBC) by clinical and molecular features in the phase II I-SPY2 clinical trial. *Ann Oncol*. 2025;36(2):172-84. DOI:10.1016/j.annonc.2024.10.018.

24. Quiroga D, Pei C, Stephens JA, et al. Impact of low hormone receptor expression on neoadjuvant chemotherapy response and patterns of care in non-metastatic HER2-negative breast cancer: a US National Cancer Database analysis. *Ther Adv Med Oncol*. 2025;17:17588359251342874. DOI:10.1177/17588359251342874.

25. Cha C, Ahn SG, Kim D, et al. Axillary response according to neoadjuvant single or dual human epidermal growth factor receptor 2 (HER2) blockade in clinically node-positive, HER2-positive breast cancer. *Int J Cancer*. 2021,149(8):1585-92. DOI:10.1002/ijc.33726.
26. Yu CC, Cheung YC, Ueng SH, et al. Factors Associated with Axillary Lymph Node Status in Clinically Node-Negative Breast Cancer Patients Undergoing Neoadjuvant Chemotherapy. *Cancers (Basel)*. 2022,14(18). DOI:10.3390/cancers14184451.
27. Wang Y, Li L, Liu X, et al. Treatment response correlation between primary tumor and axillary lymph nodes after neoadjuvant therapy in breast cancer: a retrospective study based on real-world data. *Gland Surg*. 2021,10(2):656-69. DOI:10.21037/gs-20-686.
28. Guo R, Su Y, Si J, et al. A nomogram for predicting axillary pathologic complete response in hormone receptor-positive breast cancer with cytologically proven axillary lymph node metastases. *Cancer*. 2020,126 Suppl 16:3819-29. DOI:10.1002/ncr.32830.
29. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun*. 2020,11(1):1236. DOI:10.1038/s41467-020-15027-z.
30. Niu H, Li L, Wang X, et al. The value of predicting breast cancer with a DBT 2.5D deep learning model. *Discov Oncol*. 2025,16(1):420. DOI:10.1007/s12672-025-02170-6.
31. Gao J, Zhong X, Li W, et al. Attention-based Deep Learning for the Preoperative Differentiation of Axillary Lymph Node Metastasis in Breast Cancer on DCE-MRI. *J Magn Reson Imaging*. 2023,57(6):1842-53. DOI:10.1002/jmri.28464.
32. Liu Y, Wang Y, Huang J, et al. Deep learning-based prediction of axillary pathological complete response in patients with breast cancer using longitudinal multiregional ultrasound. *EBioMedicine*. 2025,119:105896. DOI:10.1016/j.ebiom.2025.105896.

33. Li Z, Gao J, Zhou H, et al. Multiregional dynamic contrast-enhanced MRI-based integrated system for predicting pathological complete response of axillary lymph node to neoadjuvant chemotherapy in breast cancer: multicentre study. *EBioMedicine*. 2024;107:105311. DOI:10.1016/j.ebiom.2024.105311.
34. Boughey JC, Suman VJ, Mittendorf EA, et al. Sentinel lymph node surgery after neoadjuvant chemotherapy in patients with node-positive breast cancer: the ACOSOG Z1071 (Alliance) clinical trial. *Jama*. 2013;310(14):1455-61. DOI:10.1001/jama.2013.278932.
35. Boileau JF, Poirier B, Basik M, et al. Sentinel node biopsy after neoadjuvant chemotherapy in biopsy-proven node-positive breast cancer: the SN FNAC study. *J Clin Oncol*. 2015;33(3):258-64. DOI:10.1200/jco.2014.55.7827.
36. Caudle AS, Yang WT, Krishnamurthy S, et al. Improved Axillary Evaluation Following Neoadjuvant Therapy for Patients With Node-Positive Breast Cancer Using Selective Evaluation of Clipped Nodes: Implementation of Targeted Axillary Dissection. *J Clin Oncol*. 2016;34(10):1072-8. DOI:10.1200/jco.2015.64.0094.

Authors:

Jun Liao, MM ^{1#}, Mengting Xu, MM ^{2#}, Ruimin Li, MD ^{3#}, Shunian Li, MM ¹, Ziwei Cao, MM ¹, Danyang Wang, MM ¹, Yongli Li, MD ¹, Meiyun Wang, MD ¹, Hongna Tan, MD ^{1*}

Institutions:

1. Department of Radiology, People's Hospital of Zhengzhou University & Henan

485 Provincial People's Hospital, No. 7 Wei Wu Road, Jinshui District, Zhengzhou
486 450003, China;

487 2. Department of Radiology, People's Hospital of Henan University & Henan
488 Provincial People's Hospital, Zhengzhou 450003, China;

489 3. Department of Radiology, Fudan University Shanghai Cancer Center, Shanghai
490 200032, China;

491

492 **Equal contribution:** Jun Liao, Mengting Xu and Ruimin Li contributed equally to this
493 work.

494

495 ***Corresponding author:**

496 Hongna Tan, No. 7 Wei Wu Road, Jinshui District, Zhengzhou 450003, China;

497 E-mail: natan2000@126.com; Tel: +86 13526766042

498

499 **Author contributions**

500 Conception and design: Jun Liao and Hongna Tan.

501 Administrative support: Yongli Li.

502 Provision of study materials or patients: Yongli Li and Meiyun Wang.

503 Collection and assembly of data: Mengting Xu, Shunian Li, Ziwei Cao, and Danyang
504 Wang.

505 Data analysis and interpretation: Ruimin Li, Mengting Xu, Shunian Li and Ziwei Cao.

506 Manuscript writing: Jun Liao and Hongna Tan.

Manuscript revision: All authors.

Final approval of manuscript: All authors.

Accountable for all aspects of the work: All authors

Funding

This study was supported by Medical Science and Technological Project of Henan Province (No. LHGJ20220055).

Ethics approval and consent to participate

The study was approved by the Ethics Committee of Henan Provincial People's Hospital (Reference Number: 2022-124). This study is a retrospective analysis and has received an exemption from the informed consent requirement. All methods were carried out in accordance with relevant guidelines and regulations or the Declaration of Helsinki.

Declaration of competing interests

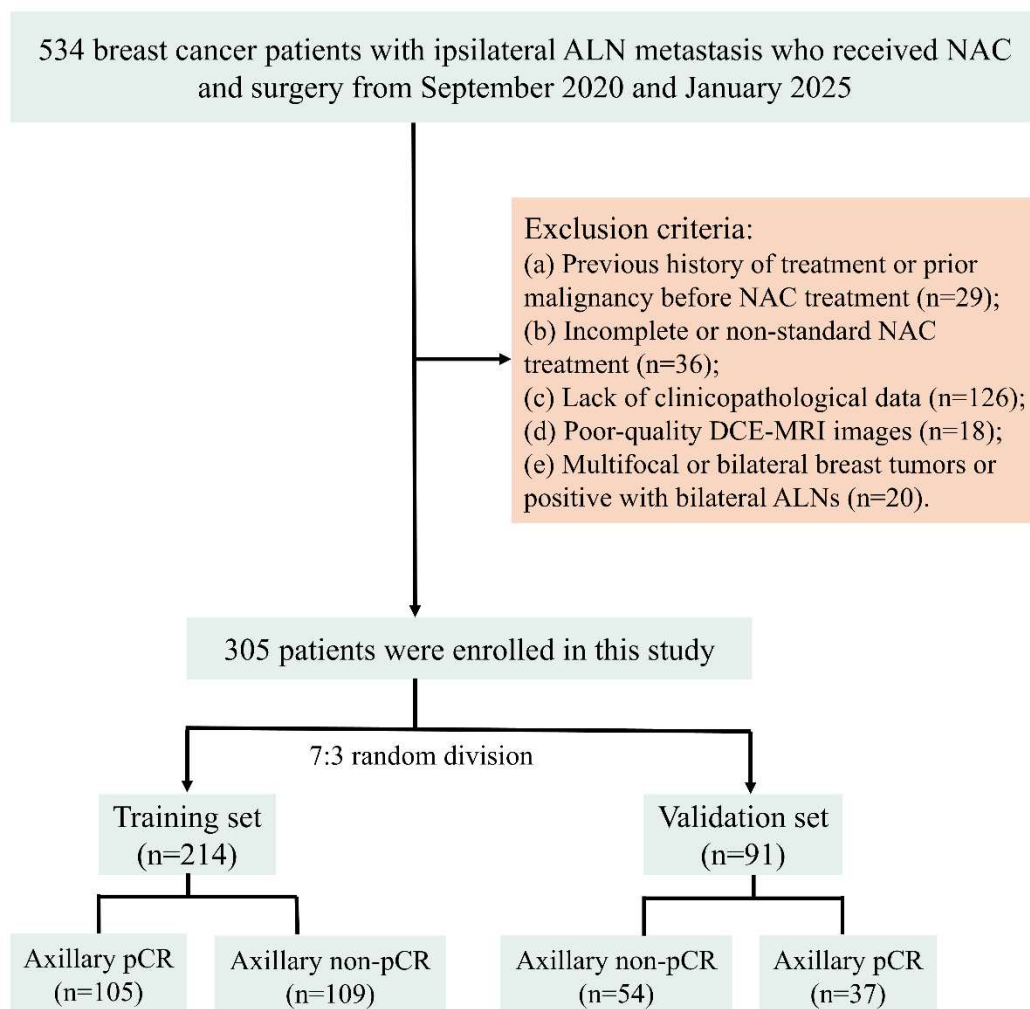
All authors declare there is no conflict of interest.

Acknowledgements

We thank the OnekeyAI platform and its developers, as well as all of the individuals who participated in this study and the technical staff for their support.

529

530 **Fig and fig's legend:**



531

532 **Figure 1:** Flowchart of patients enrolment. ALN, axillary lymph node; NAC, neoadjuvant

533 chemotherapy; DCE-MRI, dynamic contrast-enhanced MRI.

534

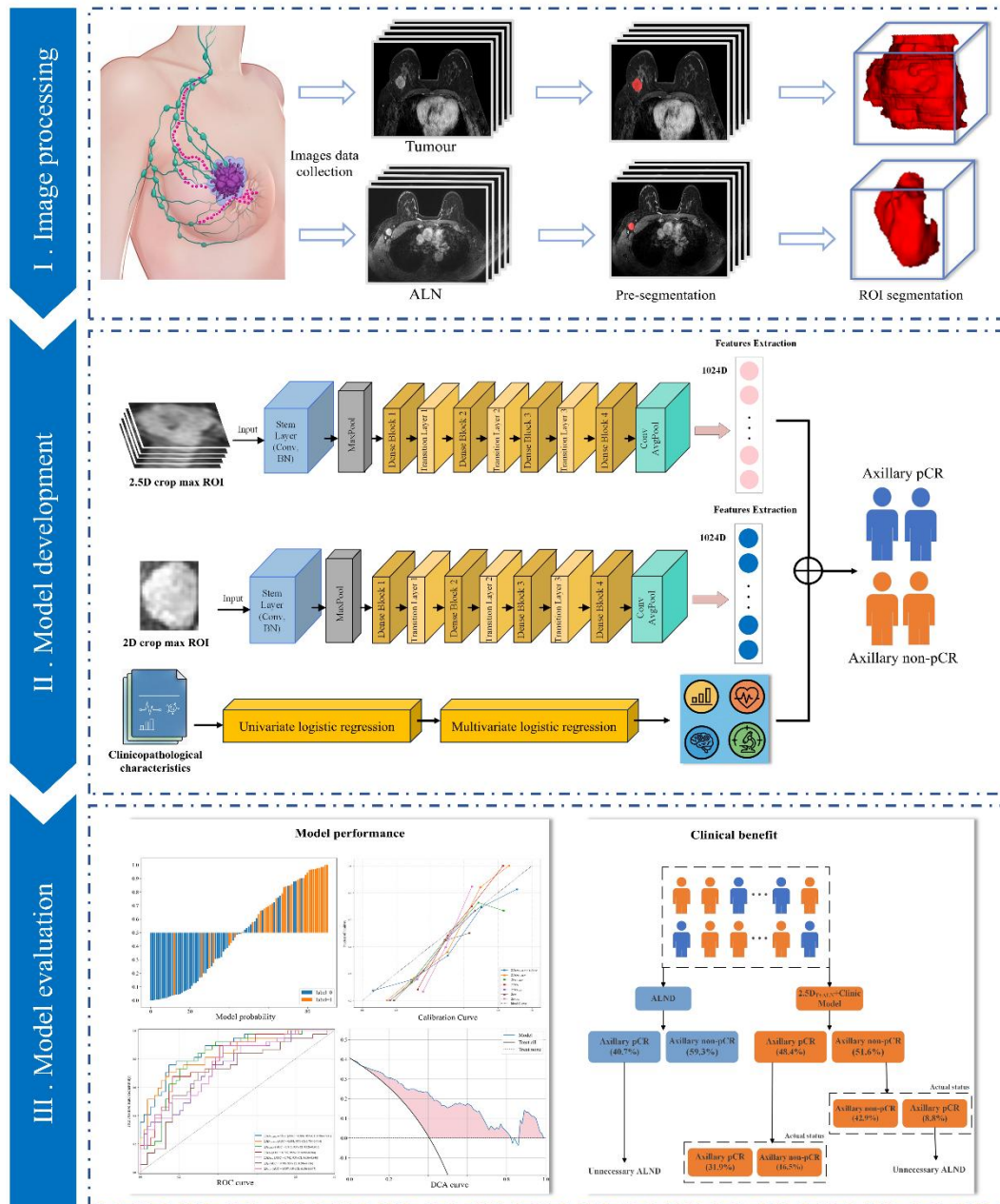


Figure 2: The overall flowchart of this study. ROC, receiver-operator characteristic; DCA, decision curve analysis; LASSO, Least Absolute Shrinkage and Selection Operator.

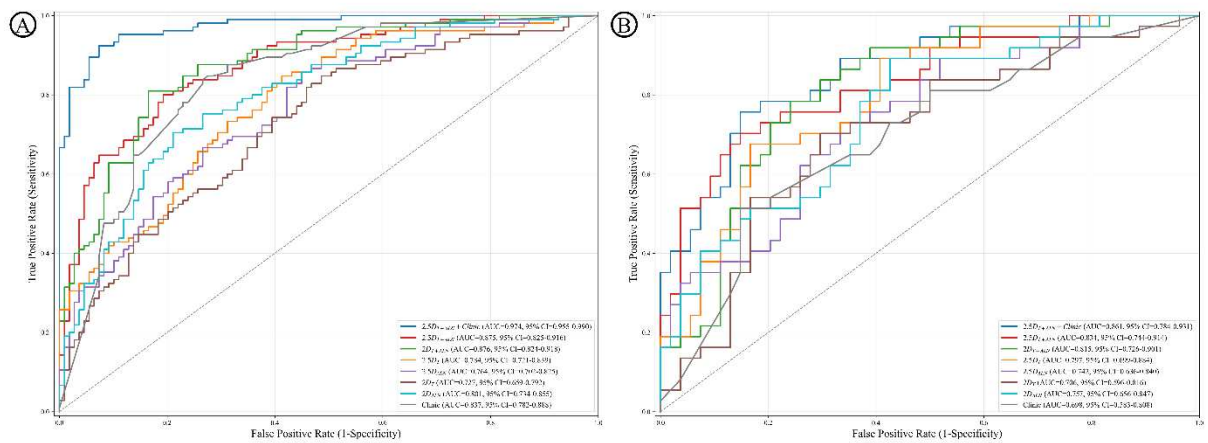


Figure 3: ROCs of models for predicting axillary pCR in breast cancer. (A) training set (n = 214) and (B) validation set (n = 91). AUC, area under the receiver operating characteristic curves

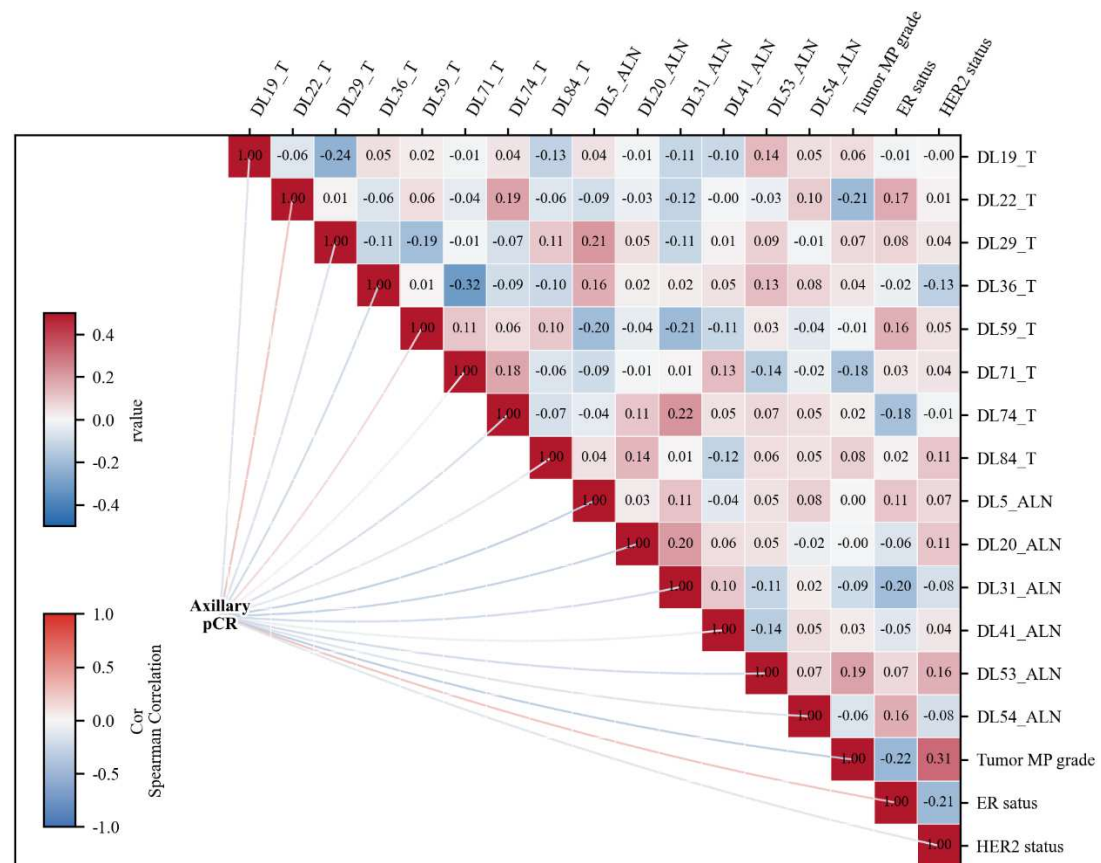


Figure 4: Spearman correlation matrices show associations among 2.5D deep learning features of the tumor and ALN, clinicopathologic characteristics, and axillary pCR.

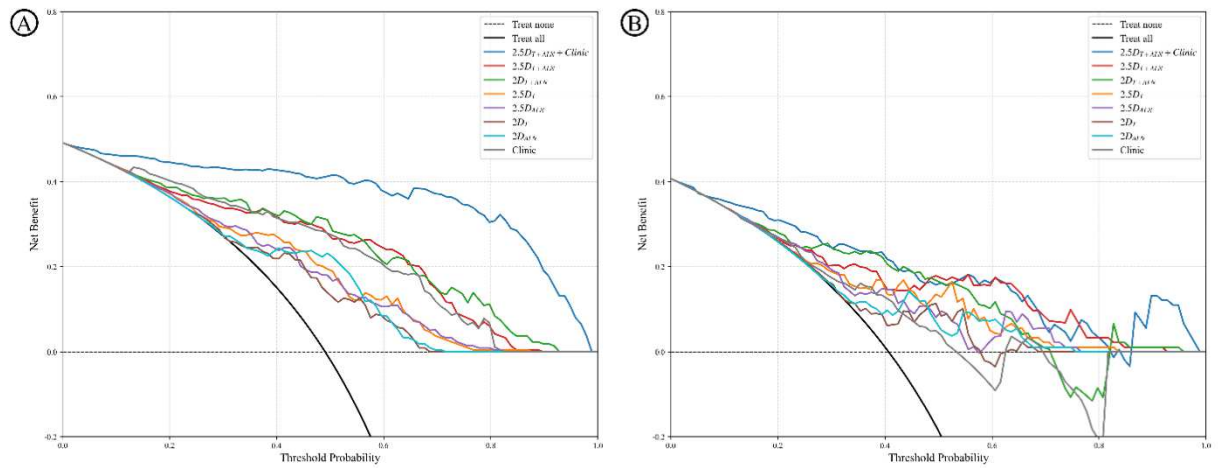


Figure 5: Decision curves of models in training and validation sets. Decision curves of the proposed deep learning models for predicting axillary pCR in the (A) training set and (B) validation set.

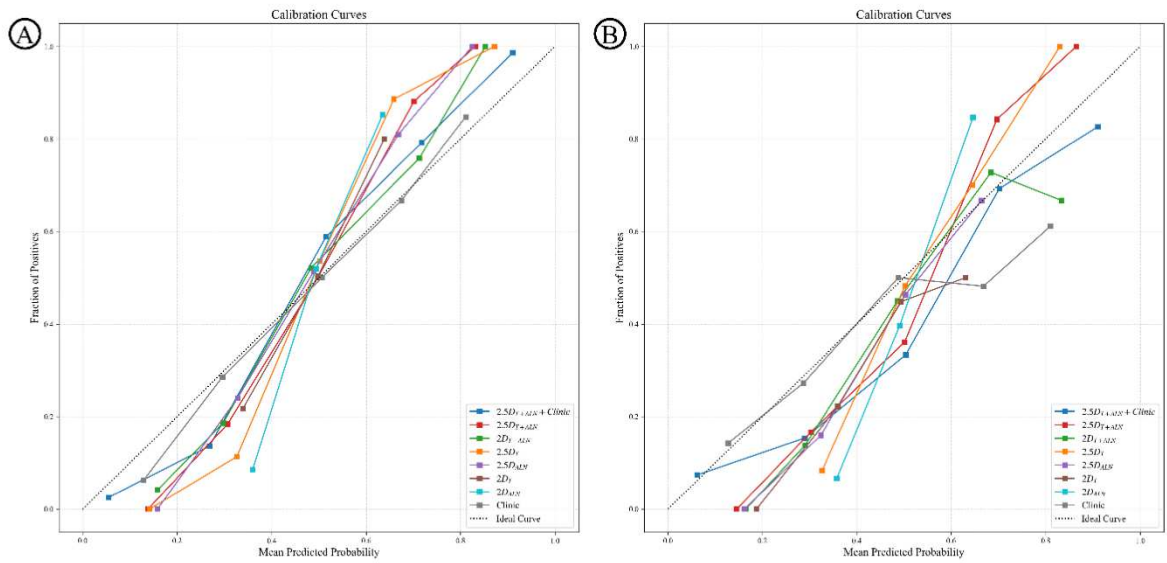


Figure 6: Calibration curves of models in training and validation sets. Calibration curves of the proposed deep learning models for predicting axillary pCR in the (A) training set and (B) validation set.

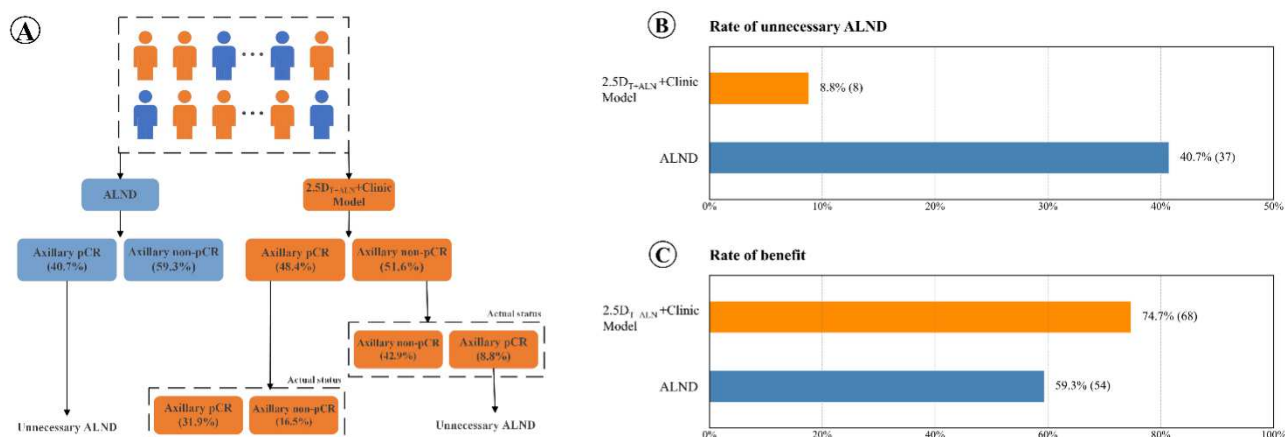


Figure 7: Clinical benefit assessment of 2.5DT+ALN+Clinic model. (A) Recommendation for ALND based on 2.5DT+ALN+Clinic model in patients with breast cancer and ALNs metastasis. (B)Rate of unnecessary ALND under two strategies: routine ALND for all patients vs ALND guided by the 2.5DT+ALN+Clinic model; (C) Clinical benefit rate under two strategies: routine ALND for all patients vs ALND guided by the 2.5DT+ALN+Clinic model. ALND, axillary lymph node dissection;

Table 1 The MRI acquisition parameters.

Scanner	TR/TE (ms)	FOV (mm)	Matrix	Slice Thickness (mm)	Slice Gap (mm)	Slices	B value (s/mm2)
GE Discovery MR750 (3.0T)	5.4/1.7	360×360	320×320	2	0	80	0/800
Siemens Skyra(3.0T)	4.5/1.6	340×340	384×384	1.5	0	80	0/1000
GE HDxt (1.5T)	3.4/1.4	240×240	256×160	2.2	0	120	0/1000

Note.—FOV = field of view, TR = repetition time, TE = echo time, ms = millisecond, mm = millimeter.

575 Table 2. Clinicopathologic characteristics of Patients in the Training and Validation sets.

Characteristics	Training set (n=214)			Validation set (n=91)		
	ALN	ALN pCR	P	ALN	ALN pCR	P
	Non-pCR(n=109)	(n=105)		Non-pCR(n=54)	(n=37)	
Age, years	47.40 ± 9.61	47.37 ± 9.68	0.916	47.87 ± 8.45	48.08 ± 10.59	0.916
ALN short diameter, mm	10.05 ± 5.93	11.65 ± 6.04	0.019	10.62 ± 6.16	12.71 ± 6.37	0.084
Tumor long diameter, mm	30.48 ± 14.84	33.58 ± 14.68	0.077	33.02 ± 14.82	30.24 ± 13.09	0.462
Lesion site (%)			0.588			0.443
Left	58 (53.2)	51 (48.6)		29 (53.7)	16 (43.2)	
Right	51 (46.8)	54 (51.4)		25 (46.3)	21 (56.8)	
Tumor MP grade (%)			<0.001			0.004
Grade 5	44 (40.4)	83 (79.1)		30 (55.6)	32 (86.5)	
Grade 1-4	65 (59.6)	22 (20.9)		24 (44.4)	5 (13.5)	
Menstrual status (%)			0.913			0.347
Pre/perimenopausal	61 (56.0)	57 (54.3)		30 (55.6)	16 (43.2)	
Postmenopausal	48 (44.0)	48 (45.7)		24 (44.4)	21 (56.8)	
Breast surgery (%)			0.885			1
BCS	23 (21.1)	24 (22.9)		10 (18.5)	7 (18.9)	
Mastectomy	86 (78.9)	81 (77.1)		44 (81.5)	30 (81.1)	
Pathological grade (%)			0.399			0.986
Grade I	2 (1.8)	2 (1.9)		2 (3.7)	1 (2.7)	
Grade II	84 (77.1)	73 (69.5)		34 (63.0)	24 (64.9)	
Grade III	23 (21.1)	30 (28.6)		18 (33.3)	12 (32.4)	
ER status (%)			<0.001			0.038
Negative	59 (54.1)	19 (18.1)		26 (48.2)	9 (24.3)	
Positive	50 (45.9)	86 (81.9)		28 (51.8)	28 (75.7)	
PR status (%)			<0.001			0.017
Negative	73 (67.0)	39 (37.1)		34 (63.0)	13 (35.1)	
Positive	36 (33.0)	66 (62.9)		20 (37.0)	24 (64.9)	
HER2 status (%)			<0.001			0.309
Negative	33 (30.3)	85 (80.9)		28 (51.8)	24 (64.9)	
Positive	76 (69.7)	20 (19.1)		26 (48.2)	13 (35.1)	
Ki-67 index (%)			0.009			0.577
Low	12 (11.0)	27 (25.7)		8 (14.8)	8 (21.6)	
High	97 (89.0)	78 (74.3)		46 (85.2)	29 (78.4)	
Molecular subtypes (%)			<0.001			0.137
Luminal A	6 (5.5)	29 (27.6)		6 (11.1)	8 (21.6)	
Luminal B	16 (14.7)	47 (44.8)		10 (18.5)	12 (32.4)	
HER2 overexpression	71 (65.1)	17 (16.2)		26 (48.2)	12 (32.4)	
Triple-negative	16 (14.7)	12 (11.4)		12 (22.2)	5 (13.6)	

576 Note.—P < 0.05 indicates that the variable distribution varied significantly between patients with axillary
577 pCR and axillary non-pCR. pCR, pathological complete response; ER, oestrogen receptor; PR,
578 progesterone receptor; HER2, human epidermal growth factor receptor; BCS, breast-conserving surgery.

Table 3. Predictors of axillary lymph node pCR to NAC using univariate and multivariable logistic regression analysis.

Variables	Univariate	logistic	P-value	Multivariate	logistic	P-value
	regression			regression		
	Odds Ratio(95%CI)			Odds Ratio(95%CI)		
Age	1.001(0.996-1.005)		0.785			
ALN short diameter	0.992(0.974-1.010)		0.482			
Tumor long diameter	0.999(0.992-1.005)		0.694			
ER status	0.581(0.434-0.779)		0.002	0.296(0.148-0.593)		0.004
PR status	0.545(0.388-0.767)		0.003	0.821(0.395-1.707)		0.657
Ki-67 index	1.244(0.969-1.597)		0.152			
Lesion site	1.004(0.870-1.158)		0.965			
Tumor MP grade	2.954(1.970-4.433)		<0.001	2.687(1.486-4.860)		0.006
Menstrual status	1.000(0.715-1.399)		1			
Breast surgery	1.062(0.823-1.370)		0.699			
Pathological grade	1.000(0.906-1.104)		1			
HER2 status	3.800(2.514-5.743)		<0.001	5.525(3.037-10.044)		<0.001
Molecular subtypes	1.131(1.039-1.232)		0.018	0.957(0.808-1.134)		0.672

Note:ER,oestrogen receptor; PR,progesterone receptor; HER2,human epidermal growth factor receptor 2; pCR, pathological complete response;NAC,neoadjuvant chemotherapy; CI, confidence interval.

Table 4. Performance Comparison of Different Models in Axillary pCR Prediction in Patients With Breast Cancer.

Set	Models	AUC	ACC(%)	SEN(%)	SPEC(%)	PPV(%)	NPV(%)	P-value
Training set	Clinic	0.838(0.783-0.892)	0.785	0.848	0.725	0.748	0.832	<0.001
	2D _T	0.727(0.659-0.794)	0.678	0.848	0.514	0.627	0.778	0.001
	2D _{ALN}	0.801(0.743-0.860)	0.748	0.705	0.789	0.763	0.735	0.001
	2.5D _T	0.784(0.724-0.844)	0.715	0.848	0.587	0.664	0.800	<0.001
	2.5D _{ALN}	0.764(0.701-0.826)	0.701	0.857	0.55	0.647	0.800	<0.001
	2D _{T+ALN}	0.876(0.830-0.922)	0.822	0.810	0.835	0.825	0.820	0.001
	2.5D _{T+ALN}	0.875(0.828-0.921)	0.804	0.800	0.807	0.800	0.807	<0.001
	2.5D _{T+ALN} +Clinic	0.974(0.956-0.991)	0.925	0.924	0.927	0.924	0.927	-
Validation set	Clinic	0.698(0.588-0.809)	0.714	0.514	0.852	0.704	0.719	<0.001
	2D _T	0.706(0.593-0.816)	0.703	0.703	0.704	0.619	0.776	0.026
	2D _{ALN}	0.757(0.658-0.856)	0.703	0.892	0.574	0.589	0.886	0.099
	2.5D _T	0.797(0.706-0.888)	0.769	0.676	0.833	0.735	0.789	0.253
	2.5D _{ALN}	0.742(0.639-0.844)	0.681	0.730	0.648	0.587	0.778	0.015
	2D _{T+ALN}	0.815(0.726-0.904)	0.769	0.784	0.759	0.690	0.837	0.431
	2.5D _{T+ALN}	0.834(0.749-0.919)	0.791	0.703	0.852	0.765	0.807	0.569
	2.5D _{T+ALN} +Clinic	0.861(0.786-0.937)	0.813	0.757	0.852	0.778	0.836	-

Note-ACC, accuracy; AUC, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV, positive predictive value; SENS, sensitivity; SPEC, specificity. P represents the AUC differences between each model and 2.5D_{T+ALN}+Clinic model (DeLong validation). Data in parentheses represent the 95% confidence interval.