Cancer Biol Med 2026. doi: 10.20892/j.issn.2095-3941.2025.0327

ORIGINAL ARTICLE

Temporal radiomics for non-invasive preoperative prediction of pathologic complete response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer

Sunyi Zheng[1]*, Shuo Wang[2]*, Ziwei Feng[3]*, Jing Liang[1], Jiaxin Liu[1], Xiaomeng Yang[1], Zhanshuo Zhang[2], Yuechen Cui[2], Jiping Xie[2], Shuxuan Fan[1], Jing Wang[4], Guoqing Liao[5], Haiyu Zhou[5], Zhaoxiang Ye[1], Jianyu Xiao[1], Lei Shi[6], Xiaonan Cui[1], Dongsheng Yue[2]

[1]Department of Radiology, Medical Artificial General Intelligence for Computation (MAGIC) Lab, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Tianjin Key Laboratory of Digestive Cancer, Tianjin Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China; [2]Department of Lung Cancer, Key Laboratory of Cancer Prevention and Therapy, National Clinical Research Center of Cancer, Lung Cancer Center, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China; [3]National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Key Laboratory of Molecular Cancer Epidemiology, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China; [4]School of Public Health, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China; [5]Department of Thoracic Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, Guangzhou 510060, China; [6]Department of Radiology, Zhejiang Cancer Hospital, Hangzhou 310022, China

ABSTRACT

Objective: This study aimed to develop and validate a temporal radiomics model based on pre- and post-treatment CT scans for the preoperative prediction of pathologic complete response (pCR) in patients with resectable non-small cell lung cancer (NSCLC) undergoing neoadjuvant chemoimmunotherapy (NCI).

Methods: Data from 263 patients with resectable NSCLC who underwent NCI followed by curative surgery and had both pre- and post-treatment CT scans were retrospectively collected. Patients from one hospital were randomly divided into training and internal test sets at a 7:3 ratio, while patients from two other hospitals served as the external test set. Radiomics features were extracted from the CT scans at both timepoints and delta features capturing the temporal changes were calculated. Radiomics models based on different features were developed using the least absolute shrinkage and selection operator for feature selection, followed by logistic regression. Model performance was evaluated using the area under the curve (AUC).

Results: The radiomics model based on delta features yielded AUCs of 0.85, 0.76, and 0.72 in the training, internal test, and external test sets, respectively, which were superior to the radiomics models based on pre-treatment features (0.74, 0.66, and 0.62, respectively) and post-treatment features (0.80, 0.76, and 0.65, respectively). By integrating the optimal features from all three feature sources, the combined model achieved further improvements in performance, with AUCs of 0.89, 0.85, and 0.78, respectively, across the three sets.

Conclusions: A CT-based radiomics model incorporating temporal features from pre- and post-treatment scans shows favorable performance for the non-invasive preoperative estimation of pCR to NCI in patients with NSCLC.

KEYWORDS Non-small cell lung cancer; radiomics; pathologic complete response; neoadjuvant therapy; computed tomography

Introduction

Lung cancer remains one of the leading causes of cancer-related mortality worldwide and continues to be a focus of global attention[1,2]. Non-small cell lung cancer (NSCLC) is the predominant subtype of lung cancer, accounting for approximately 85% of all lung cancer cases. Although surgical resection is the standard treatment for early-stage resectable NSCLC, the risks of recurrence and metastasis remain high in

*These authors contributed equally to this work.
Correspondence to: Dongsheng Yue, Xiaonan Cui and Lei Shi
E-mail: yuedongsheng_cg@163.com, cuixiaonan@tjmuch.com and shilei@zjcc.org.cn
ORCID ID: https://orcid.org/0000-0002-4937-7525
Received August 5, 2025; accepted October 30, 2025;
published online January 14, 2026.
Available at www.cancerbiomed.org
©2026 The Authors. Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

patients undergoing surgery alone, resulting in relatively poor 5-year survival rates[3]. Neoadjuvant chemotherapy has shown clinical benefits but a considerable proportion of patients still experience disease relapse or progression[4].

With the rapid advances in immunotherapy, neoadjuvant chemoimmunotherapy has emerged as a promising therapeutic strategy for NSCLC. By combining chemotherapy and immunotherapy prior to surgery, this approach aims to enhance anti-tumor immunity, shrink the tumor, improve the likelihood of achieving R0 resection, and eliminate micro-metastatic disease, thereby minimizing the risk of recurrence. Several studies have shown that neoadjuvant chemoimmunotherapy can improve pathologic complete response (pCR) rates, as well as progression-free and overall survival in patients with NSCLC[5-7]. Nevertheless, not all NSCLC patients benefit from this approach. Clinical trials have shown that only 17%–41% of patients achieve pCR[5-11]. Accurate preoperative prediction of pCR could help identify patients who may be candidates for a "wait-and-see" strategy, as has been explored in rectal and esophageal cancers. This approach may spare high-risk patients, such as those with poor physical status or centrally located tumors requiring pneumonectomy, from unnecessary surgical intervention. In addition, extensive lymph node dissection might negatively affect prognosis following immunotherapy. Limiting the extent of dissection in patients achieving pCR may help preserve immune function and improve subsequent treatment outcomes.

Currently, pCR assessment relies on invasive tissue sampling and postoperative pathologic examination. However, tumor heterogeneity may limit the accuracy of biopsy specimens, which may not fully capture the overall treatment response. Imaging modalities, such as CT and PET, have a vital role as non-invasive alternatives in evaluating treatment response. Unlike pathologic examinations, these imaging modalities enable dynamic monitoring of tumor characteristics and their changes before and after treatment, thereby potentially reducing reliance on tissue sampling. Despite these features of imaging modalities, discrepancies between radiologic and pathologic assessments persist. For example, 33% of patients classified with stable disease and 73% with a partial response on radiologic imaging in the NADIM trial were shown to have achieved a pCR[12]. This finding underscores the need for a reliable method to bridge the gap between radiologic and pathologic evaluations.

Radiomics has shown considerable potential in predicting treatment response in lung cancer[13-15]. For example, a study showed that radiomics features extracted from PET-CT achieved an AUC of 0.82 in predicting pCR after neoadjuvant chemoimmunotherapy, performing better than models using CT or PET alone[14]. In addition, Yang et al. reported that radiomics features extracted from pre- or post-treatment CT scans can be used to predict a pCR after neoadjuvant or conversion chemoimmunotherapy[15]. However, the value of temporal radiomics features from both pre- and post-treatment CT scans for pCR prediction has not been reported.

To this end, the current study aimed to develop and evaluate a temporal CT-based radiomics model that incorporates pre-treatment features, post-treatment features, and the feature differences for the non-invasive preoperative prediction of a pCR in NSCLC patients undergoing neoadjuvant chemoimmunotherapy.

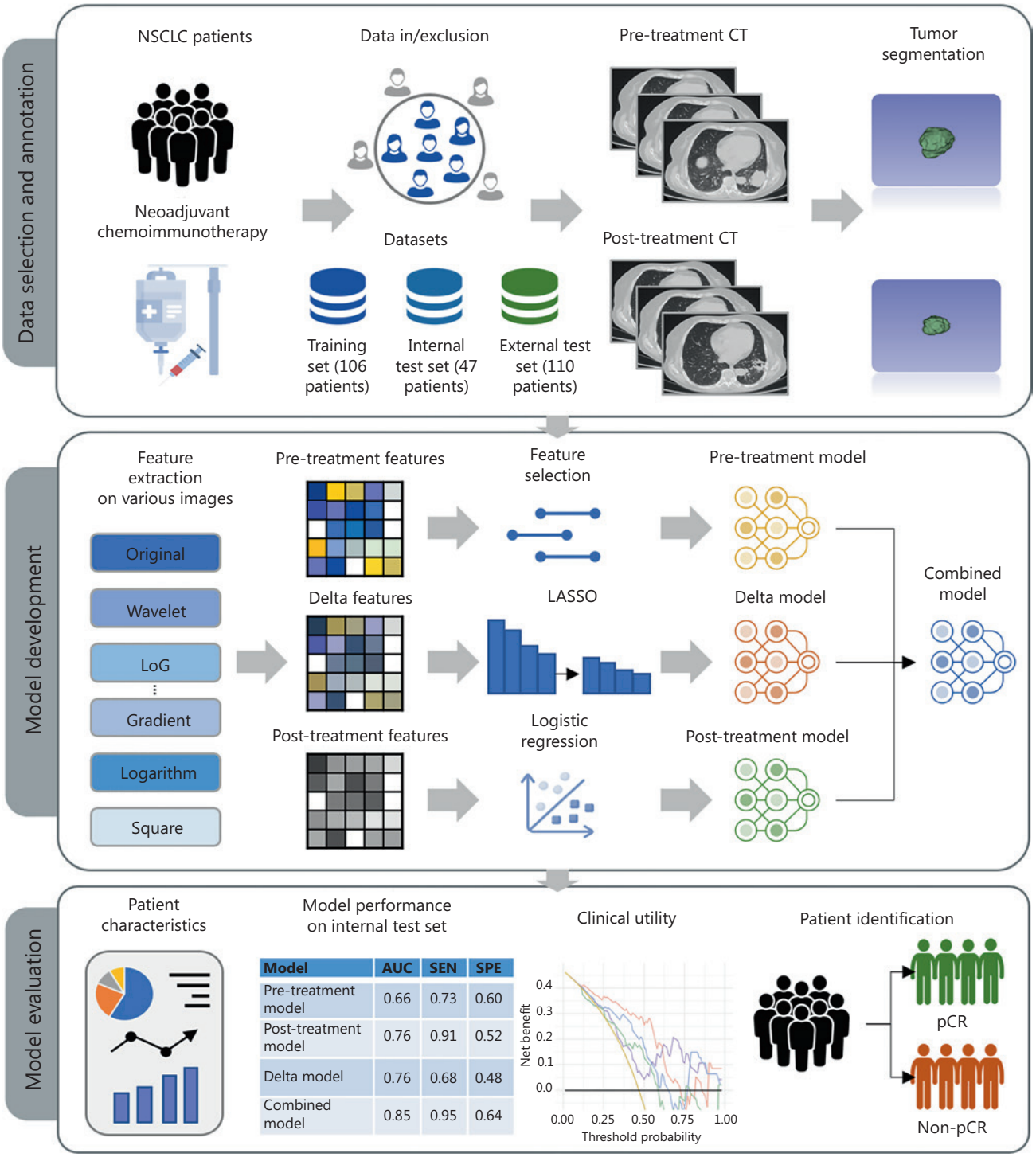
Materials and methods

Study cohorts

This retrospective study was approved by the Ethics Committees of Tianjin Medical University Cancer Institute and Hospital (Hospital 1, BC20252896), Guangdong Provincial People's Hospital (Hospital 2, KY202509302), and Zhejiang Cancer Hospital (Hospital 3, IRB2023951). Ethical approval was waived by the hospitals due to the retrospective nature of the study. Data were collected from 165 NSCLC patients treated at Hospital 1 between August 2018 and July 2021 and from 127 patients treated at Hospitals 2 and 3. All patients underwent curative surgery following neoadjuvant chemoimmunotherapy. The patient inclusion criteria were as follows: (1) a confirmed diagnosis of NSCLC by endoscopic bronchoscopy or CT-guided needle biopsy; (2) clinical stage II–III disease based on the TNM classification; (3) receipt of neoadjuvant chemoimmunotherapy prior to surgery; and (4) availability of both pre- and post-treatment non-enhanced CT scans performed before surgery. The exclusion criteria included the following: (1) poor-quality CT images; (2) absence of pre- or post-treatment CT scans; (3) tumor boundaries that were unclear or difficult to delineate; and (4) incomplete clinical or pathologic data. A detailed flowchart of the patient selection process is presented in **Figure S1**. Baseline clinical data collected prior to treatment included age, gender, smoking history, pathologic subtype, and clinical stage. The flowchart of study design is shown in **Figure 1**.

Neoadjuvant administration

All patients received a standardized neoadjuvant chemoimmunotherapy regimen consisting of platinum-based chemotherapy combined with immune checkpoint inhibitors. The chemotherapy backbone typically included cisplatin, lobaplatin, nedaplatin, or carboplatin. Immunotherapy agents primarily targeted the PD-(L)1 axis and included drugs, such as nivolumab, pembrolizumab, sintilimab, camrelizumab, tislelizumab, envafolimab, SHR-1701, durvalumab, penpulimab, adebrelimab, and toripalimab. In addition, a small subset of patients received non-PD-(L)1-based agents as part of



Model	AUC	SEN	SPE
Pre-treatment model	0.66	0.73	0.60
Post-treatment model	0.76	0.91	0.52
Delta model	0.76	0.68	0.48
Combined model	0.85	0.95	0.64

Figure 1 Flowchart of the study design. Tumors from non-small cell lung cancer (NSCLC) patients who underwent neoadjuvant chemoimmunotherapy were segmented on pre- and post-treatment CT scans. Radiomics features were extracted separately from the pre- and post-treatment images and delta-radiomics features were calculated as the relative change between them. These features were used to build three models including pre-treatment, post-treatment, and delta-radiomics models. The optimal features from each model were integrated to develop a combined model. All models were evaluated on internal and external test sets using the area under the curve and decision curve analysis to predict pathologic complete response (pCR). SEN and SPE represent sensitivity and specificity, respectively.

combination regimens, including bevacizumab, ipilimumab, and IBI110. Immunotherapy was administered every 3 weeks and surgical resection was performed within 4–6 weeks after completion of neoadjuvant therapy. The specific drug regimen, number of treatment cycles, and timing of surgery were individualized based on multidisciplinary team discussions, which accounted for the clinical status and treatment response of the patient.

Definition of pathologic complete response

Two pathologists independently evaluated treatment response on pathologic sections. The evaluation was performed in accordance with the IASLC guidelines[16], which define pCR as the absence of viable tumor cells in the primary tumor bed and regional lymph nodes. Any discrepancies between the two assessments were resolved by joint review and consensus discussion.

CT acquisition

Non-enhanced CT scans were acquired using GE Discovery CT750 HD, GE LightSpeed 16, and GE LightSpeed 8 scanners (GE Healthcare, Chicago, IL, USA); Siemens Somatom Sensation 64 (Siemens Healthineers, Erlangen, Bavaria, Germany); Philips Brilliance iCT 256 (Philips Healthcare, Best, North Brabant, the Netherlands); and United Imaging uCT 760 (United Imaging Healthcare, Shanghai, Shanghai, China). The scanning range extended from the lung apex-to-below the diaphragm. The tube voltage was set at 120 kV and the tube current was automatically modulated. Images were reconstructed with a slice thickness of 1.25 mm using the Stnd and Lung reconstruction kernels for the GE CT systems. In contrast, a 1.0- or 1.5-mm slice thickness with B70f, B30, and B31f reconstruction kernels was used for the Siemens CT scanner. The Philips CT scanner applied a slice thickness of 1 mm with B, C, YA, and YB reconstruction kernels and the United Imaging CT scanner utilized the B_SHARP_C reconstruction kernel with a slice thickness of 1 mm. The soft kernels included B30, B31f, Stnd, B, and C, whereas the hard kernels included B70f, Lung, YA, YB, and B_SHARP_C.

Tumor annotation and radiomics feature extraction

Tumors were semi-automatically delineated using 3D Slicer (version 5.6.2; Surgical Planning Laboratory, Brigham and Women's Hospital, Boston, MA, USA) by two radiologists, each with > 10 years of experience. A senior radiologist reviewed and manually refined the delineations on both pre- and post-treatment CT scans after the initial segmentation. Both radiologists were blinded to the clinical and pathologic patient information.

Radiomics features were automatically extracted using PyRadiomics (version 3.0.1; Computational Imaging and Bioinformatics Lab, Brigham and Women's Hospital, Boston, MA, USA) from the segmented tumor regions on the original and a series of filtered images. The applied filters included wavelet, Laplacian of Gaussian, gradient, logarithm, square, square root, exponential, and local binary pattern (lbp) filters in two- and three-dimensional forms. These filters are commonly used to highlight different aspects of the image. For example, wavelet transforms decompose the image into multiple frequency components, capturing both spatial and frequency information. The Laplacian of Gaussian filter enhances fine edges and subtle boundaries. The gradient filter emphasizes intensity transitions, while the logarithm, square, square root, and exponential filters modify intensity distributions to reveal additional contrast. Local binary pattern filters extract local texture patterns and are particularly useful for quantifying tissue heterogeneity. By applying these filters, various transformed versions of the original image are generated, enabling the extraction of complementary features that may not be visible on the unprocessed images.

Radiomics features were extracted and categorized into first-order statistics and shape- and texture-based features from the original and filtered images. The first-order feature class contains 19 types of features that describe the overall distribution of voxel intensities. Shape-based features quantify the size, shape, and surface characteristics of the tumor, including 16 three-dimensional and 10 two-dimensional feature types. Texture-based features reflect the spatial arrangement and relationships of voxel intensities, providing insight into tumor heterogeneity. These features were derived from multiple statistical matrices, including 24 feature types from the gray level co-occurrence matrix (GLCM), 16 from the gray level run length matrix (GLRLM), 16 from the gray level size zone matrix (GLSZM), 14 from the gray level dependence matrix (GLDM), and 5 from the neighboring gray tone difference matrix (NGTDM). Each matrix captures different aspects of texture, such as uniformity, granularity, and spatial complexity. Detailed computational formulas for the features are available at https://pyradiomics.readthedocs.io/en/latest/features.html.

All CT scans were resampled to an isotropic voxel size of $1 \times 1 \times 1$ mm^3 before feature extraction to enhance the robustness of the radiomics features. Radiomics features of 30 cases were randomly selected for consistency assessment to ensure feature reproducibility using the intraclass correlation coefficient. Intra- and inter-observer segmentations were independently performed by two radiologists on the same subset of cases to ensure feature reproducibility. Features with an intraclass correlation coefficient > 0.75 were considered stable and included in the subsequent analysis[17].

Delta-radiomics features

Radiomics features were independently extracted at each time point for patients with pre- and post-treatment CT scans. Delta-radiomics features were calculated as the relative change from pre- to post-treatment and defined as the difference between post- and pre-treatment values normalized by the pre-treatment value. Similarly, to verify the reproducibility of delta features derived from pre- and post-treatment scans, the same 30 cases were evaluated by calculating the intraclass correlation coefficient between the initial segmentation and the segmentation performed 1 month later. The delta features with a coefficient > 0.75 were considered stable and retained for further analysis.

Radiomics model development

Four radiomics models were developed based on different feature sources, including pre-treatment, post-treatment, delta, and a combined set of optimal features drawn from all three feature sources. All data were standardized using Z-score normalization before model development to minimize bias caused by differences in feature scales. Feature selection was carried out in four sequential steps. First, features with an intraclass correlation coefficient < 0.75 were excluded to ensure reliability. Second, features not significantly associated with the outcome were removed using one-way analysis of variance (ANOVA) with a threshold of 0.2. Third, Spearman correlation analysis was performed to reduce redundancy and 1 was retained when the correlation between 2 features exceeded 0.7. Finally, the least absolute shrinkage and selection operator (LASSO) method with 10-fold cross-validation was applied to identify the most predictive features. The final model for predicting pCR was built using weighted binary logistic regression with backward selection. Class weights were automatically adjusted inversely to the frequencies to address class imbalance. Accordingly, classes with larger sample sizes were assigned smaller weights, whereas classes with smaller sample sizes were assigned larger weights.

Model interpretability

The Shapley additive explanations (SHAP) method was applied to improve interpretability of the radiomics model[18]. The SHAP method was used to analyse the contribution of each radiomic feature within the model, providing a visual representation of how features influence the prediction outcome. A positive value indicates an increase in the predicted probability, while a negative value indicates a decrease. This method allows for evaluation of the overall relevance of each feature and the role in individual predictions.

Statistical analysis

Statistical analyses were performed using Python (version 3.10.11; Python Software Foundation, Wilmington, DE, USA) and R (version 4.4.2; R Foundation for Statistical Computing, Vienna, Austria). Categorical variables were compared between different patient groups using a chi-squared test. An independent t-test or Kruskal–Wallis H test was applied for continuous variables depending on data distribution[19-21]. Model performance was assessed by comparing the area under the curve (AUC), sensitivity, and specificity. The optimal classification threshold was identified by maximizing the sum of the sensitivity and specificity. Sensitivity was defined as the proportion of true positives among all actual positives at this threshold and specificity was defined as the proportion of true negatives among all actual negatives. The DeLong test was used to evaluate statistical differences in AUCs between models[22]. Boxplots of predicted probabilities for pCR across radiomics models were generated and group differences between pCR and non-pCR cases were assessed using the Mann–Whitney U test. The clinical value of the model was evaluated using decision curve analysis by assessing the net benefit across a range of threshold probabilities. A two-sided P value < 0.05 was considered statistically significant.

Results

Patient characteristics

A total of 263 patients who met the inclusion and exclusion criteria were enrolled in the study. Data from 153 patients at Hospital 1 were randomly divided into training and test sets in a 7:3 ratio, while the remaining 110 patients from Hospitals 2 and 3 served as the external test set. The clinical characteristics of patients in sets are summarized in **Table 1**. The majority of patients in the training, internal, and external test sets were male, accounting for 82.1%, 87.2%, and 89.1% of patients, respectively. A large proportion of patients in the training (80.2%) and internal test sets (89.4%) had a history of cigarette smoking compared to 62.7% in the external test set. Squamous cell carcinoma was the predominant pathologic subtype (72.6% of patients in the training set, 74.5% in the internal test set, and 66.4% in the external test set). Clinical stage III tumors were observed in 56.6% of patients in the training set, 53.2% in the internal test set, and 65.5% in the external test set. Most patients received anti-PD-1 or anti-PD-L1 therapy with rates of 92.5% in the training set, 87.2% in the internal test set, and 100% in the external test set. A pCR was achieved in 47.2% of patients in the training set, 46.8% in the internal test set, and 35.5% in the external test set.

Table 1 Patient characteristics in the training and test sets

Variables	Training set (*n* = 106)			Internal test set (*n* = 47)			External test set (*n* = 110)		
	Non-pCR (*n* = 56)	pCR (*n* = 50)	*P* value	Non-pCR (*n* = 25)	pCR (*n* = 22)	*P* value	Non-pCR (*n* = 71)	pCR (*n* = 39)	*P* value
Age (years)	58.73 ± 6.91	62.06 ± 6.10	0.010	62.20 ± 6.51	62.36 ± 7.73	0.938	62.2 ± 9.0	65.4 ± 7.2	0.092
Gender			0.320			0.252			0.053
Male	44 (78.6%)	43 (86.0%)		20 (80.0%)	21 (95.5%)		60 (84.5%)	38 (97.4%)	
Female	12 (21.4%)	7 (14.0%)		5 (20.0)	1 (4.5%)		11 (15.5%)	1 (2.6%)	
Smoking history			0.156			0.426			0.669
No	14 (25.0%)	7 (14.0%)		4 (16.0%)	1 (4.5%)		28 (39.4%)	13 (33.3%)	
Yes	42 (75.0%)	43 (86.0%)		21 (84.0%)	21 (95.5%)		43 (60.6%)	26 (66.7%)	
Pathology			< 0.001			0.278			0.794
Non-SCC	25 (44.6%)	4 (8.0%)		8 (32.0%)	4 (18.2%)		25 (35.2%)	12 (30.8%)	
SCC	31 (55.4%)	46 (92.0%)		17 (68.0%)	18 (81.8%)		46 (64.8%)	27 (69.2%)	
ICIs			1.000			1.000			1.000
Anti-PD-(L)1	52 (92.9%)	46 (92.0%)		22 (88.0%)	19 (86.4%)		71 (100.0%)	39 (100.0%)	
Non-PD-(L)1	4 (7.1%)	4 (8.0%)		3 (12.0%)	3 (13.6%)		0 (0.0%)	0 (0.0%)	
Clinical stage			0.784			0.447			0.683
II	25 (44.6%)	21 (42.0%)		13 (52.0%)	9 (40.9%)		26 (36.6%)	12 (30.8%)	
III	31 (55.4%)	29 (58.0%)		12 (48.0%)	13 (59.1%)		45 (63.4%)	27 (69.2%)	

SCC, ICIs, PD-1, and PD-L1 refer to squamous cell carcinoma, immune checkpoint inhibitors, programmed cell death protein 1, and programmed death-ligand 1, respectively.
P < 0.05 was considered statistically significant.

Pre-treatment radiomics model

Three features were selected for the model based on radiomics features extracted from pre-treatment CT scans through a sequential process involving ANOVA, Spearman correlation analysis, LASSO regression, and binary logistic regression. The selected features were original_glcm_ClusterProminence, wavelet-HHL_firstorder_Skewness, and wavelet-HHL_glcm maximal correlation coefficient (MCC). The pre-treatment radiomics model achieved an AUC of 0.74 (95% CI: 0.65–0.84) in the training set, 0.66 (95% CI: 0.50–0.83) in the internal test set, and 0.62 (95% CI: 0.51–0.73) in the external test set. Sensitivity and specificity were 0.70 and 0.75 in the training set, 0.73 and 0.60 in the internal test set, and 0.62 and 0.58 in the external test set, respectively. Detailed model performance is shown in **Figure 2**. Model performance on external test subsets from the two hospitals are shown in **Tables S1 and S2**.

Post-treatment radiomics model

Similarly, three radiomics features extracted from post-treatment CT scans were shown to be predictive of pCR. The selected features included log-sigma-5-0-mm-3D_glcm_MCC, original_shape_Flatness, and lbp-3D-m1_firstorder_Skewness. The post-treatment radiomics model yielded an AUC of 0.80 (95% CI: 0.71–0.88) in the training set, 0.76 (95% CI: 0.61–0.88) in the internal test set, and 0.65 (95% CI: 0.55–0.75) in the external test set. The model showed a sensitivity of 0.78 and a specificity of 0.71 in the training set. Sensitivity and specificity were 0.91 and 0.52 for the internal test set, respectively, while the sensitivity was 0.56 and the specificity was 0.63 in the external test set.

Delta-radiomics model

Five features were selected by ANOVA, Spearman correlation analysis, LASSO regression, and binary logistic regression when using delta-radiomics features derived from pre-treatment and post-treatment CT scans.These features included wavelet-HLL_ngtdm_Strength, log-sigma-1-0-mm-3D_glrlm_ShortRunHighGrayLevelEmphasis, lbp-3D-m2_firstorder_Median, log-sigma-5-0-mm-3D_glrlm_GrayLevelNonUniformityNormalized, and original_shape_Flatness. The delta-radiomics model achieved AUCs of 0.85

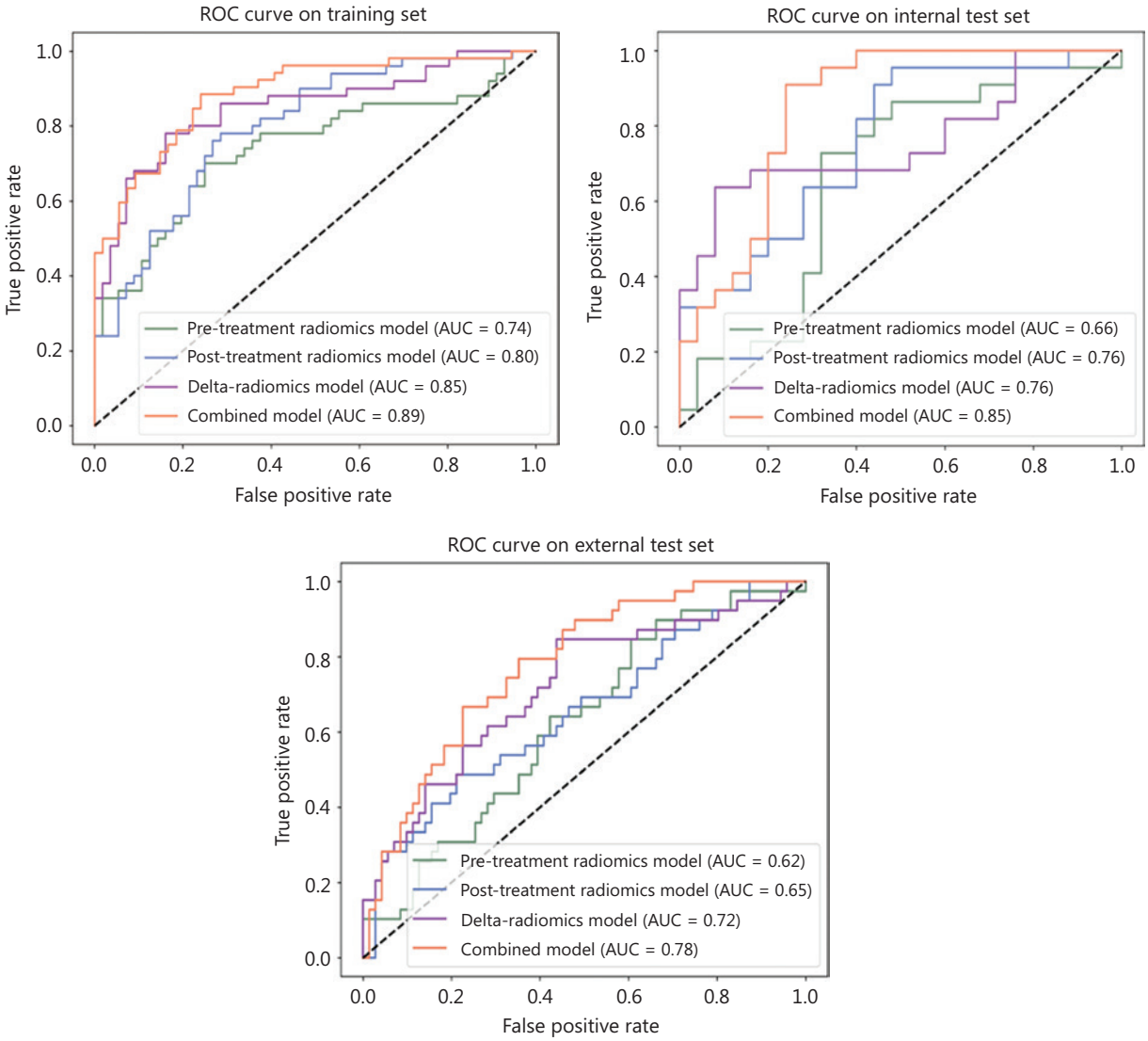


Figure 2 Continued

(95% CI: 0.77–0.91), 0.76 (95% CI: 0.60–0.90), and 0.72 (95% CI: 0.62–0.82) in the training, internal test, and external test sets, respectively. The corresponding sensitivities were 0.78, 0.68, and 0.64, respectively, and the specificities were 0.84, 0.48, and 0.66, respectively.

Combined radiomics model

A combined model incorporating the optimal features from the three aforementioned models was also developed. The selected radiomics features included log-sigma-5-0-mm-3D_glrlm_GrayLevelNonUniformityNormalized_delta, original_shape_Flatness_posttreatment, lbp-3D-m1_firstorder_Skewness_posttreatment, original_glcm_ClusterProminence_pretreatment, log-sigma-1-0-mm-3D_glrlm_ShortRunHighGrayLevelEmphasis_delta, and wavelet-HLL_ngtdm_Strength_delta. The combined model achieved an AUC of 0.89 (95% CI: 0.82–0.94) in the training set, 0.85 (95% CI: 0.74–0.95) in the internal test set, and 0.78 (95% CI: 0.69–0.86) in the external test set. The combined model showed significantly improved performance in the training, internal test, and external test sets compared to the pre-treatment model (P = 0.004, P = 0.040, and P = 0.018, respectively). In contrast, no significant differences were detected between the combined and delta-radiomics models in the internal and external test sets (P = 0.256 and 0.182, respectively). A sub-analysis stratifying the data by manufacturer, reconstruction kernel type, and slice thickness to evaluate the combined model performance is provided in **Tables S3–S5**.

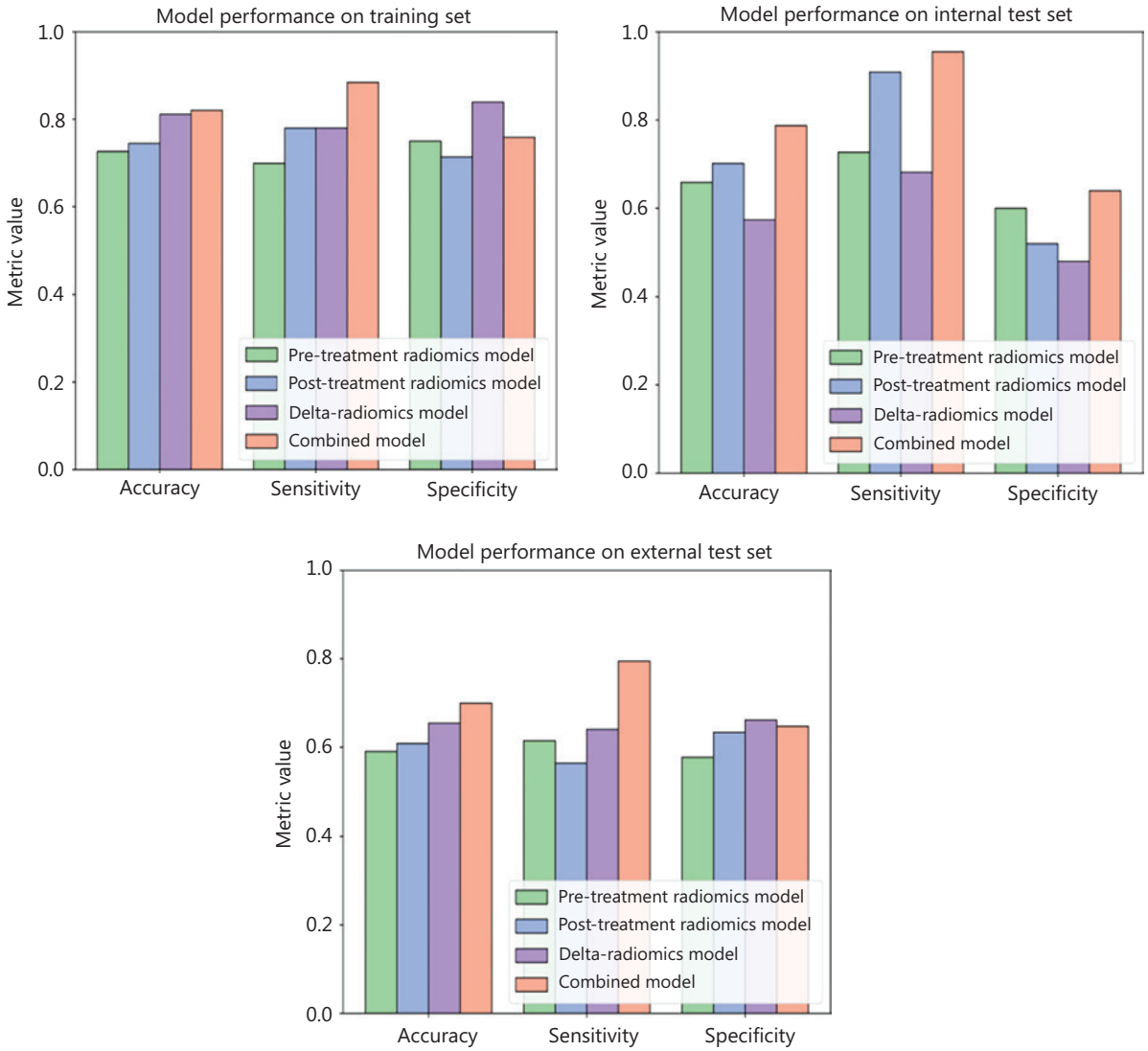


Figure 2 Comparison of model performance of four models in predicting pathologic complete response for non-small cell lung cancer patients to neoadjuvant chemoimmunotherapy. The combined model integrates the optimal radiomics features from pre-treatment, post-treatment and delta radiomics models.

Probability distributions

Boxplots of predicted probabilities for pCR across four radiomics models are displayed in **Figure 3**. The figure shows that the predicted probabilities for pCR were significantly higher in the post-treatment, delta, and combined models than for non-response in the training, internal test, and external test sets.

Decision curve analysis

The clinical usefulness of the predictive models was evaluated using decision curve analysis (**Figure 4**). The combined model showed a net clinical benefit when the threshold probability ranged between 0.12 and 0.55 in the training, internal, and external test sets compared to the treat-all and -none strategies, indicating that the model possesses a relatively broad clinical utility. The combined model provided a higher net benefit at a threshold probability of 0.25 compared to the other models in all three sets.

Model interpretability with the SHAP method

The SHAP values were computed for the selected radiomic features in the four radiomics models. The SHAP feature importance plots present the features in descending order

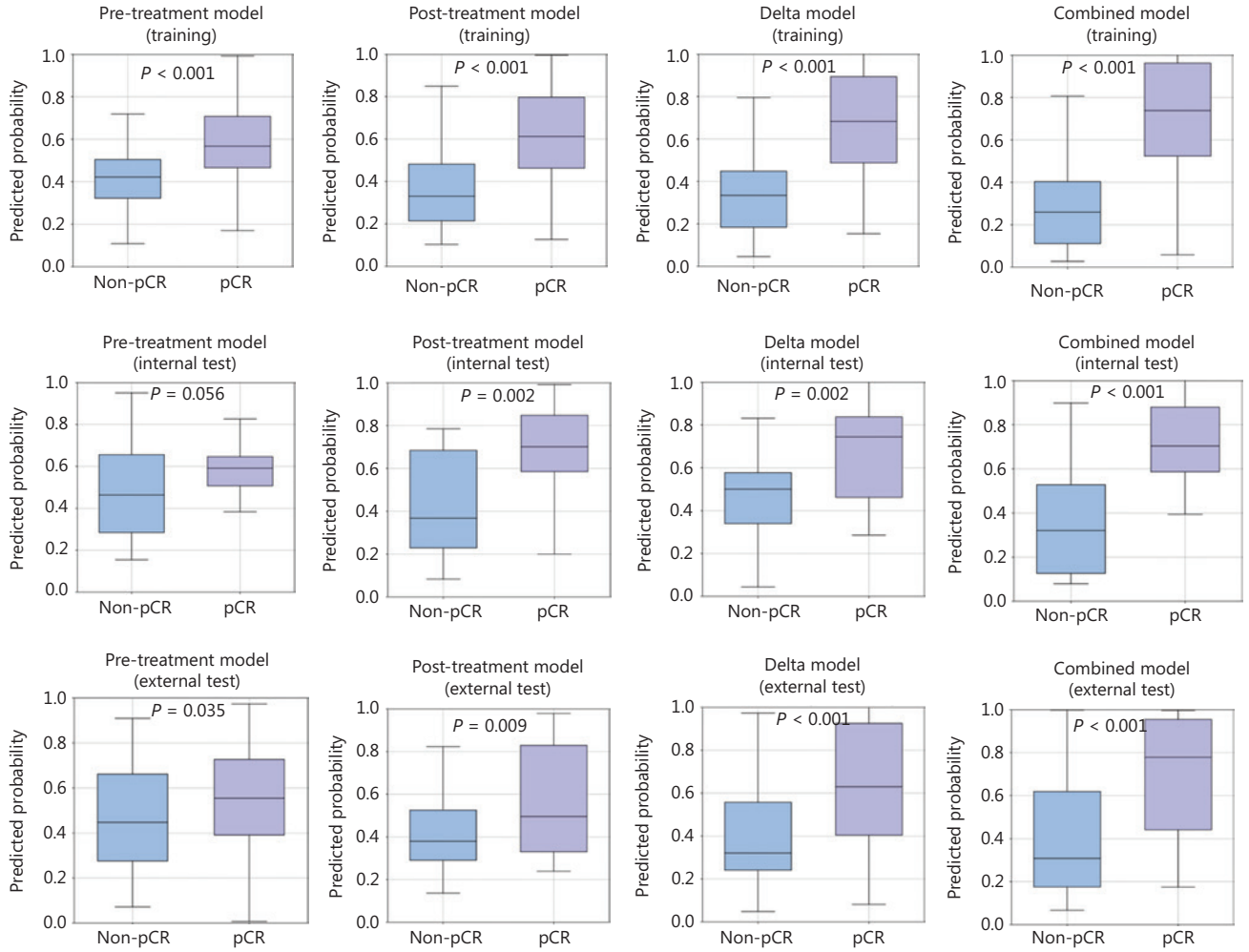


Figure 3 Boxplot comparison of predicted probabilities for pathologic complete response (pCR) across radiomics models.

based on the average contribution to the model output, as shown in **Figure 5**. Features toward the top of the plot tended to have a greater influence on the prediction. The original_glcm_ClusterProminence, lbp-3D-m1_firstorder_Skewness, wavelet-HLL_ngtdm_Strength, and original_shape_Flatness_posttreatment features were identified as having the strongest impact on the pCR prediction in the pre-treatment, post-treatment, delta, and combined models, respectively. **Figure 6** displays the SHAP summary plots, which illustrate the overall influence of individual features on the model prediction of pCR. Positive SHAP values are associated with a higher predicted probability of pCR, while negative values suggest a lower predicted probability. The magnitude of each SHAP value reflects the extent to which a feature contributed to the prediction for a given case.

SHAP waterfall plots were generated for two patients randomly selected from the combined model to illustrate individual-level interpretability, as shown in **Figure 7**. A large change in wavelet-HLL_ngtdm_Strength between pre- and post-treatment scans may reflect tumor shrinkage in the left plot, which could have contributed positively to the prediction of a pCR. In contrast, the right plot shows minimal variation in wavelet-HLL_ngtdm_Strength, suggesting limited change in tumor size or boundary characteristics following treatment. The original_shape_Flatness_posttreatment exhibited a relatively large negative influence in the right plot, indicating that the lesion with a more spherical post-treatment shape were associated with a lower probability of achieving a pCR.

Discussion

A pCR is regarded as a short-term endpoint for evaluating the efficacy of neoadjuvant chemoimmunotherapy in NSCLC. Given the variability in pCR rates among patients, accurate preoperative identification of individuals who may

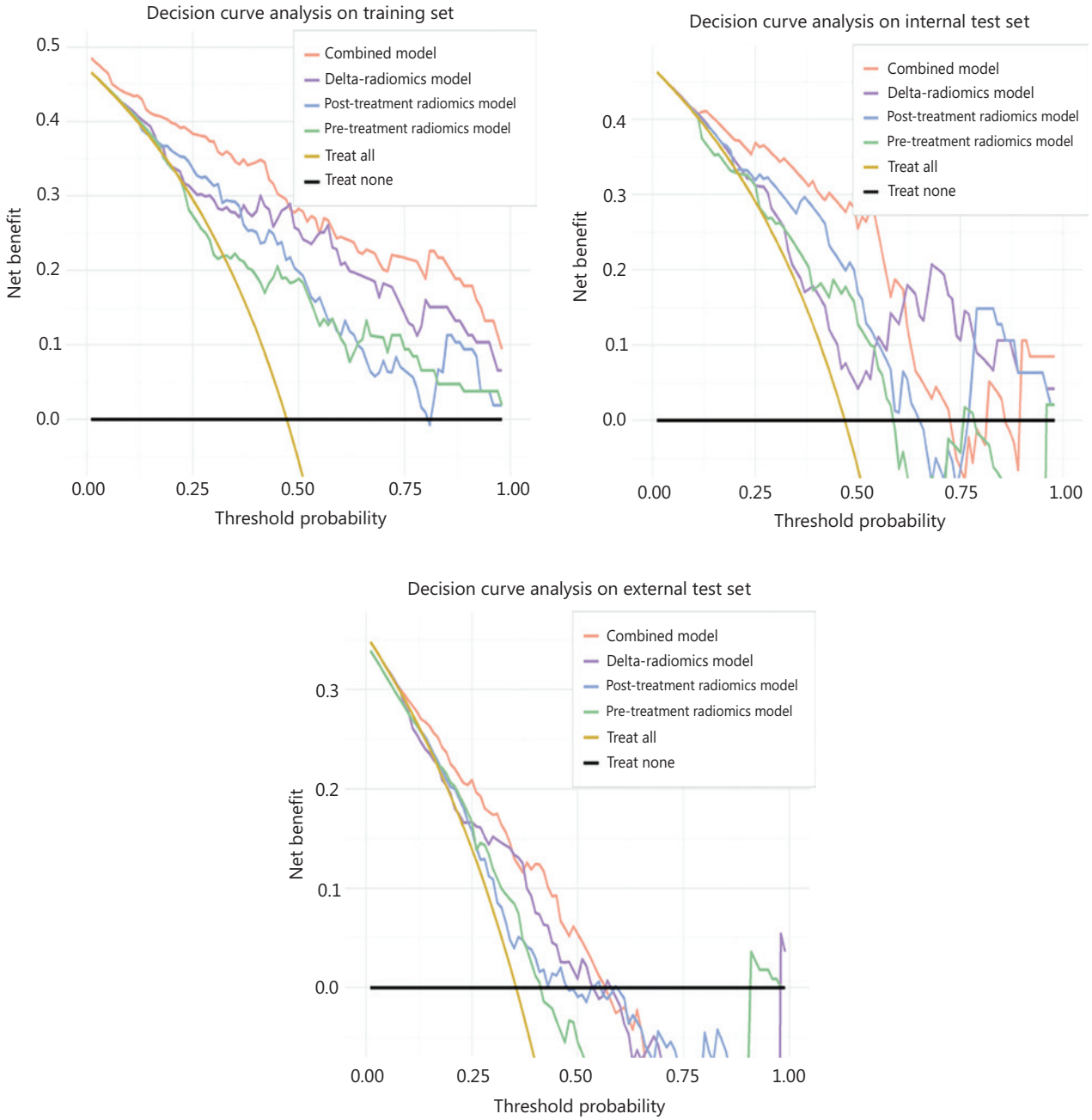


Figure 4 Decision curves of different radiomics models in prediction of pathologic complete response.

achieve a pCR could help avoid unnecessary treatment and facilitate more personalized therapeutic planning. To this end, a radiomics-based model utilizing longitudinal CT imaging was developed to predict a pCR. The experimental results showed that the combined model, which integrated temporal features extracted from pre- and post-treatment scans and the differences, achieved good AUCs of 0.89, 0.85, and 0.78 in the training, internal test, and external test sets, respectively.

Radiomics enables comprehensive and non-invasive characterization of tumor features from different CT scans during the treatment course of NSCLC. In this study radiomics features extracted from pre- and post-treatment CT scans, the differences between these features, and the combined optimal features for predicting a pCR were investigated. The post-treatment model outperformed the pre-treatment model. This finding may be attributed to an ability of the post-treatment model to capture morphologic and textural

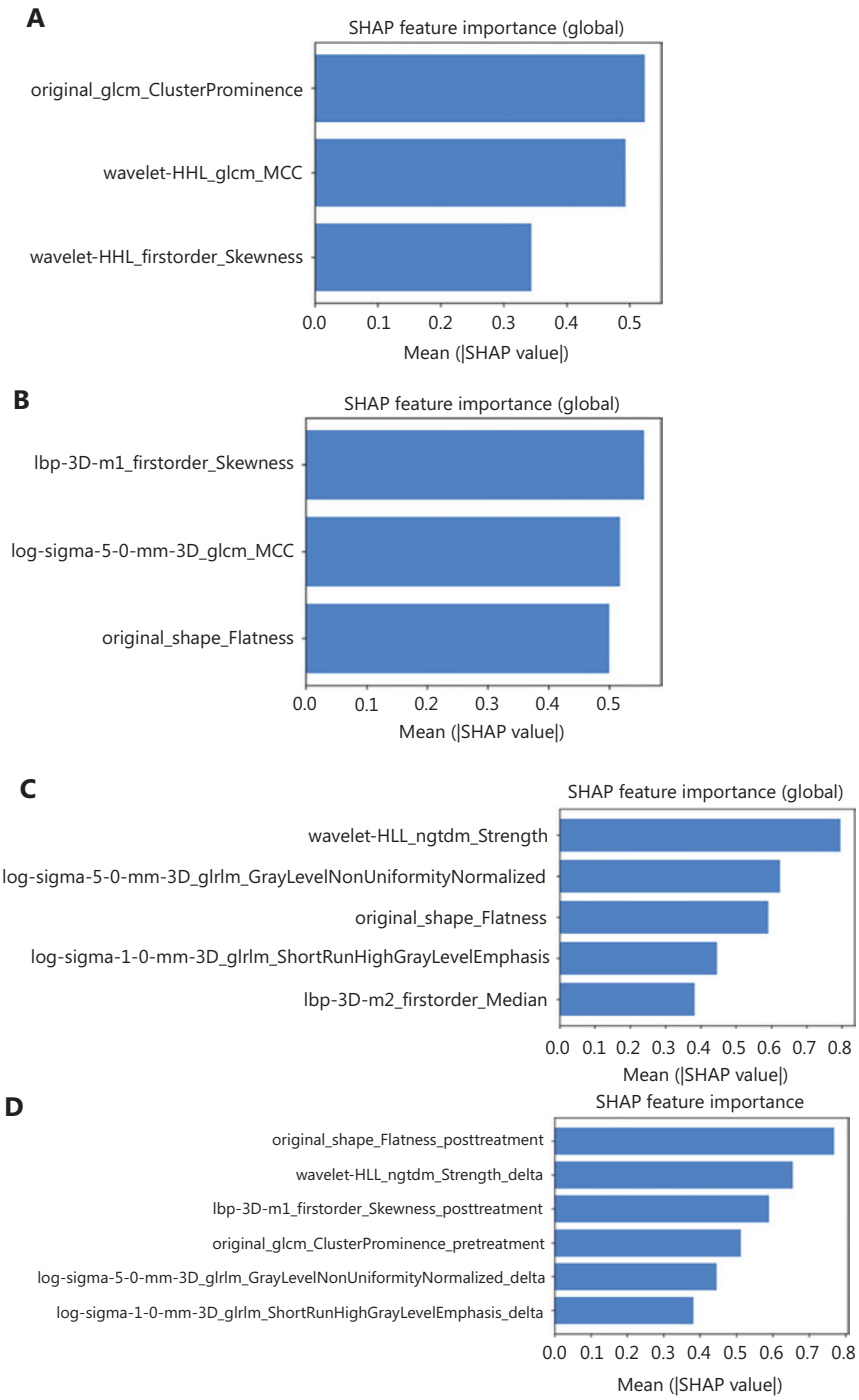


Figure 5 Feature importance plots based on Shapley Additive Explanations (SHAP) for the pre-treatment, post-treatment, delta, and combined radiomics models (A–D). Features are ranked in descending order according to the mean absolute Shapley values, reflecting the relative contribution to the overall model prediction. Features at the top have a greater influence on the model output.

alterations induced by neoadjuvant chemoimmunotherapy, such as necrosis and fibrosis. In contrast, pre-treatment features primarily reflected the intrinsic heterogeneity and structural complexity of tumors at baseline, which may be less predictive of a pCR after neoadjuvant chemoimmunotherapy compared to post-treatment characteristics. A similar trend in model performance has also been reported in previous studies utilizing both pre- and post-treatment CT

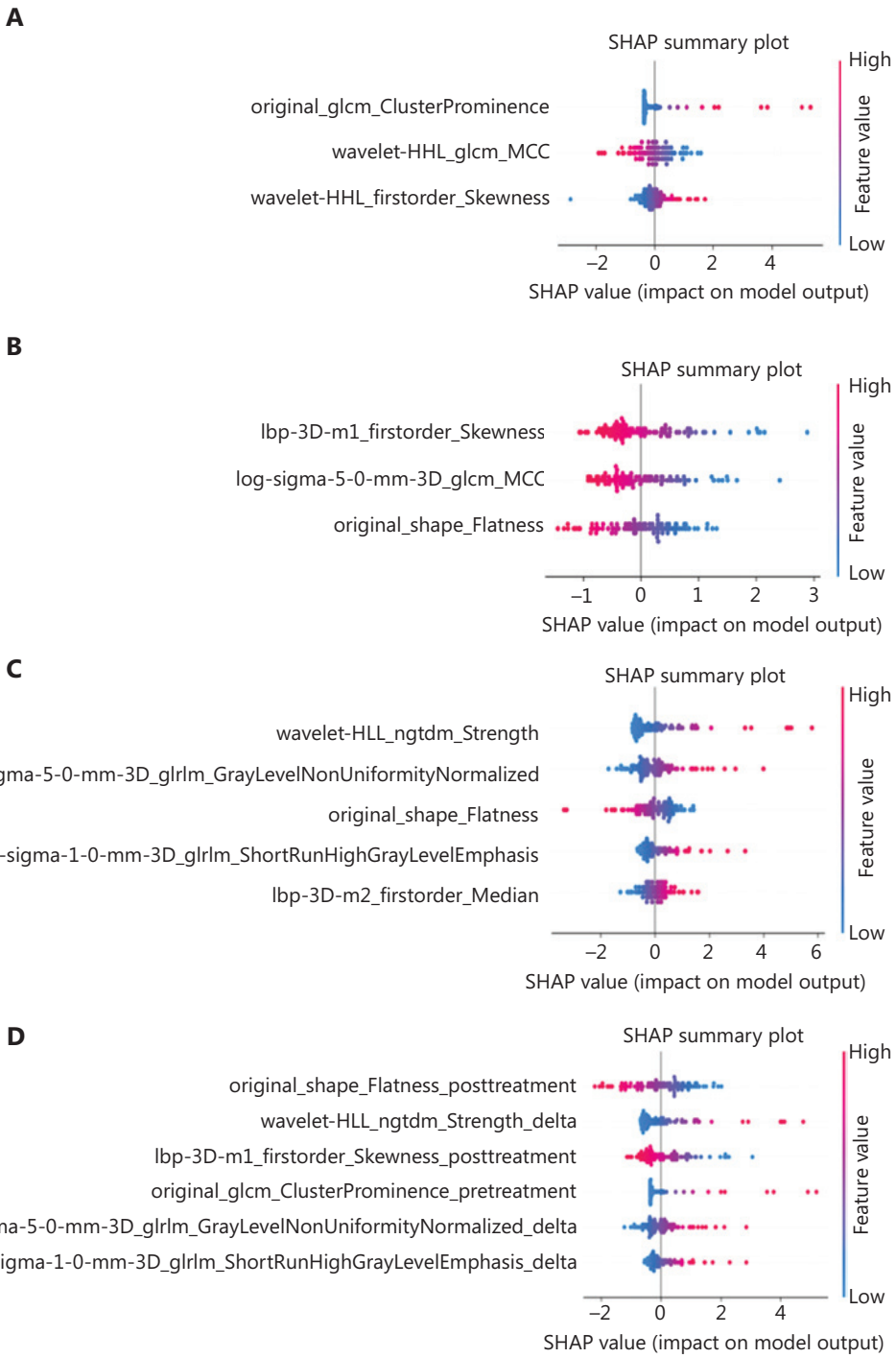


Figure 6 Summary plots based on Shapley Additive Explanations (SHAP) showing the global influence of each radiomic feature on the prediction of pathologic complete response across the pre-treatment, post-treatment, delta, and combined radiomics models (A–D). Features are ranked from top to bottom according to the overall impact. Positive SHAP values are associated with a higher predicted probability of pCR, while negative values suggest a lower predicted probability.

features for treatment response[15,23]. In addition, we calculated the feature change between pre- and post-treatment imaging to build the delta-radiomics model. The model also showed better performance than the pre-treatment model, which suggests that the dynamic transformation of tumor phenotype may serve as a powerful surrogate for treatment

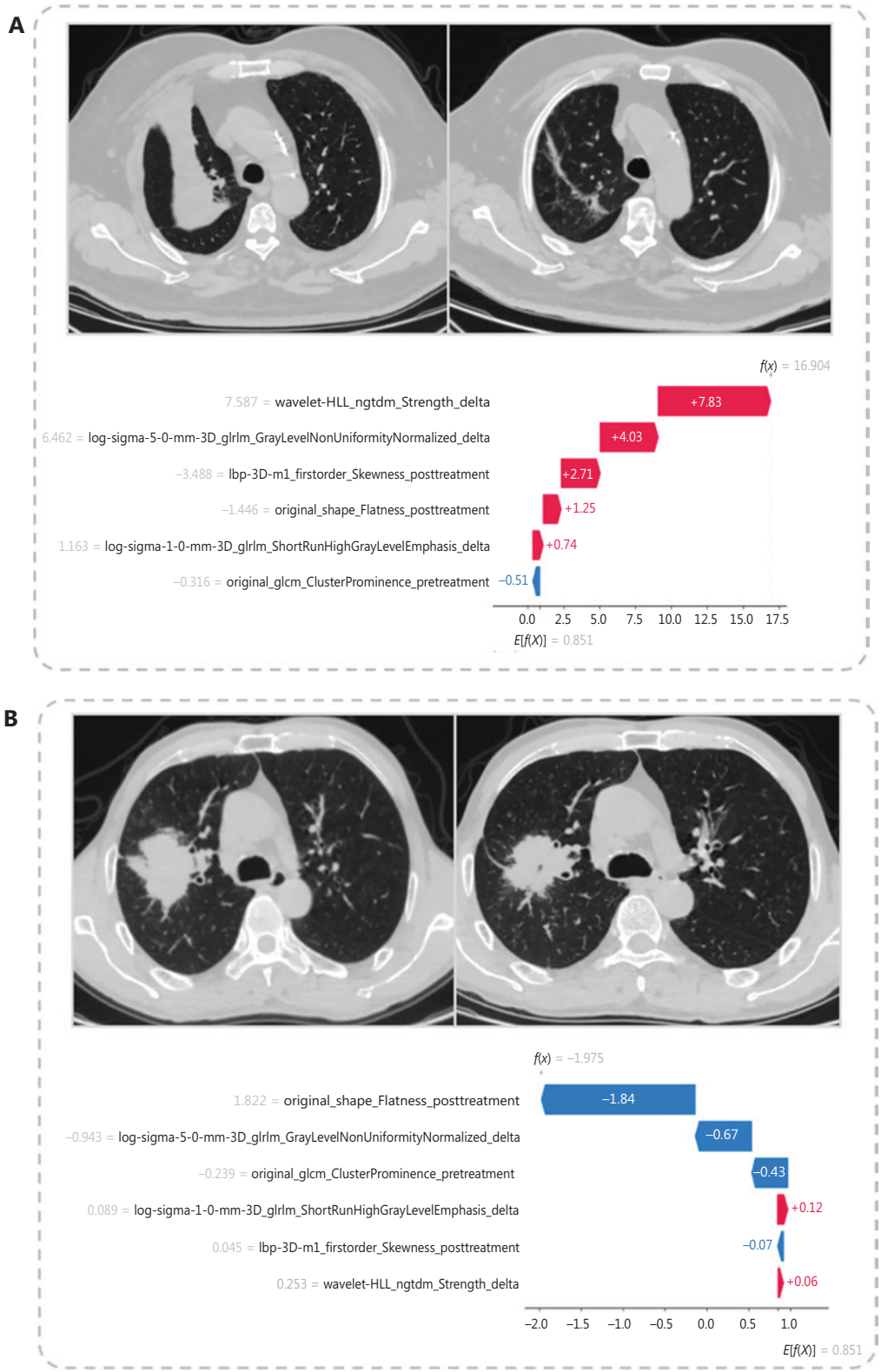


Figure 7 Waterfall plots based on Shapley Additive Explanations (SHAP) visualizing individual-level feature contributions in the combined radiomics model. Red bars indicate features that increased the predicted probability of a pCR, while blue bars indicate those that decreased

the predicted probability of a pCR. The final model output reflects the cumulative influence of all features. In the upper plot, a large change in wavelet-HLL_ngtdm_Strength may suggest tumor shrinkage, contributing to the prediction of pCR. In the lower plot, original_shape_Flatness_posttreatment exhibited a relatively large negative influence, indicating that the lesion with a more spherical post-treatment shape was associated with a lower probability of achieving a pCR.

efficacy. This finding can be the reason that the selected delta radiomic features may capture biologically relevant temporal alterations, such as heterogeneity, immune infiltration, and tumor microenvironment. Furthermore, the combined model, which integrates features from pre-treatment, post-treatment, and delta scans, achieved the best overall performance, indicating that these feature sets provide complementary information for pCR prediction.

The performance of the combined model was comparable to the performance reported in other studies focusing on the prediction of a pCR in NSCLC following neoadjuvant chemoimmunotherapy. Ye et al. developed a deep learning model using non-enhanced and contrast-enhanced pre-treatment CT scans for patients with stage IB–III NSCLC[24]. The model achieved AUCs of 0.76 and 0.80 on the test set using non-enhanced and enhanced scans, respectively. When both scan types were combined as input, the AUC improved to 0.87, highlighting the value of multimodal CT fusion. Despite the promising results, the approach was limited in patients with contraindications to contrast agents. In contrast, our combined model relied solely on non-enhanced CT images and still achieved an AUC of 0.85 for predicting a pCR. Furthermore, Yang et al. developed a radiomics model based on PET and CT imaging for stage IB–IIIB NSCLC, achieving an AUC of 0.82 in the test set[14]. The good performance might be attributed to the fact incorporation of tumor metabolic activity, which is relevant for predicting a pathologic response[25]. However, PET imaging is costly and not routinely available in all clinical settings. Notably, the CT-based method utilized in the current study is more cost-effective and feasible for widespread clinical use.

There were some limitations in this study. First, due to the retrospective nature of this study, selection bias may have occurred. Further prospective research is warranted to confirm the utility of the model in clinical practice. Second, the study primarily focused on predicting a pCR in stage II–III NSCLC patients undergoing surgical resection following neoadjuvant chemoimmunotherapy. Incorporating stage I NSCLC cases in future models may be valuable for supporting clinical decision-making and facilitating more personalized treatment planning for a broader patient population. Third, this study only considered radiomics features extracted from pre- and post-treatment CT scans. In addition to imaging data, incorporating complementary information, such as biochemical indicators, genomic profiling, and pathology-derived features, may also be valuable for predicting a pCR or other endpoints[26,27]. However, these data were not available in the present study. We plan to collect such information in a corollary study to further improve model performance and broaden the clinical applicability. Fourth, our study was limited by the small sample size for model development. Nevertheless, the model yielded an AUC of 0.85 for pCR prediction, which may offer preliminary insight into the potential predictive value of temporal radiomics features.

Conclusions

In conclusion, we developed a radiomics model that incorporated pre- and post-treatment features for the non-invasive preoperative prediction of pCR in NSCLC patients undergoing neoadjuvant chemoimmunotherapy. The model shows potential for identifying patients who are likely to achieve pCR and supporting personalized treatment planning.

Acknowledgements

We thank the high performance computing platform of Tianjin Medical University for support.

Grant support

This study was funded by grants from the National Natural Science Foundation of China (Grant No. 82573102, 82173038, 82273428, 82302180 and 82171932), Noncommunicable Chronic Diseases-National Science and Technology Major Project (Grant No. 2023ZD0501700, 2023ZD0501703, 2024ZD0520000 and 2024ZD0520002), Tianjin Key Medical Discipline (Specialty) Construction Project (Grant No. TJYXZDXK-010A), Tianjin Key Medical Discipline Construction Project (Grant No.TJYXZDXK-3-004B), Scientific Developing Foundation of Tianjin Education Commission (Grant No. 2024KJ182), Tianjin Health Research Project (Grant No. TJWJ2024QN012).

Conflict of interest statement

No potential conflicts of interest are disclosed.

Authors contributions

Conceived and designed the analysis: Sunyi Zheng, Shuo Wang, Ziwei Feng, Lei Shi, Xiaonan Cui, Dongsheng Yue.
Collected the data: Jing Liang, Jiaxin Liu, Xiaomeng Yang, Zhanshuo Zhang, Yuechen Cui, Jiping Xie, Shuxuan Fan, Guoqing Liao.
Contributed data or analysis tools: Haiyu Zhou, Zhaoxiang Ye, Jianyu Xiao, Lei Shi, Xiaonan Cui, Dongsheng Yue.
Performed the analysis: Sunyi Zheng, Shuo Wang, Ziwei Feng, Jing Wang.
Wrote the paper: Sunyi Zheng, Shuo Wang, Ziwei Feng, Jing Wang.

Data availability statement

The data generated in this study are available upon request from the corresponding author.

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin. 2025; 75: 10-45.
2. Zheng S, Liu J, Xie J, Zhang W, Bian K, Liang J, et al. Differentiating high-grade patterns and predominant subtypes for IASLC grading in invasive pulmonary adenocarcinoma using radiomics and clinical-semantic features. Cancer Imaging. 2025; 25: 42.
3. Wozniak AJ, Gadgeel SM. Adjuvant therapy for resected non-small cell lung cancer. Ther Adv Med Oncol. 2009; 1: 109-18.
4. Felip E, Altorki N, Zhou C, Vallières E, Martínez-Martí A, Rittmeyer A, et al. Overall survival with adjuvant atezolizumab after chemotherapy in resected stage II–IIIA non-small-cell lung cancer (IMpower010): a randomised, multicentre, open-label, phase III trial. Ann Oncol. 2023; 34: 907-19.
5. Yue D, Wang W, Liu H, Chen Q, Chen C, Zhang J, et al. LBA58 pathological response to neoadjuvant tislelizumab (TIS) plus platinum-doublet (PtDB) chemotherapy (CT) in resectable stage II–IIIA NSCLC patients (pts) in the phase III (Ph3) RATIONALE-315 trial. Ann Oncol. 2023; 34: S1299.
6. Forde PM, Spicer J, Lu S, Provencio M, Mitsudomi T, Awad MM, et al. Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer. N Engl J Med. 2022; 386: 1973-85.
7. Heymach JV, Harpole D, Mitsudomi T, Taube JM, Galffy G, Hochmair M, et al. Perioperative durvalumab for resectable non-small-cell lung cancer. N Engl J Med. 2023; 389: 1672-84.
8. Wakelee H, Liberman M, Kato T, Tsuboi M, Lee S-H, Gao S, et al. Perioperative pembrolizumab for early-stage non-small-cell lung cancer. N Engl J Med. 2023; 389: 491-503.
9. Lu S, Zhang W, Wu L, Wang W, Zhang P, Fang W, et al. Perioperative toripalimab plus chemotherapy for patients with resectable non-small cell lung cancer: the neotorch randomized clinical trial. JAMA. 2024; 331: 201-11.
10. Cascone T, Awad M, Spicer J, He J, Lu S, Sepesi B, et al. LBA1 CheckMate 77T: phase III study comparing neoadjuvant nivolumab (NIVO) plus chemotherapy (chemo) vs neoadjuvant placebo plus chemo followed by surgery and adjuvant NIVO or placebo for previously untreated, resectable stage II–IIIb NSCLC. Ann Oncol. 2023; 34: S1295.
11. Forde P, Spicer J, Girard N, Provencio M, Lu S, Wang C, et al. 84O Neoadjuvant nivolumab (N) + platinum-doublet chemotherapy (C) for resectable NSCLC: 3-y update from CheckMate 816. J Thorac Oncol. 2023; 18: S89-90.
12. Provencio M, Nadal E, Insa A, García-Campelo MR, Casal-Rubio J, Dómine M, et al. Neoadjuvant chemotherapy and nivolumab in resectable non-small-cell lung cancer (NADIM): an open-label, multicentre, single-arm, phase 2 trial. Lancet Oncol. 2020; 21: 1413-22.
13. Zheng S, Cui X, Ye Z. Integrating artificial intelligence into radiological cancer imaging: from diagnosis and treatment response to prognosis. Cancer Biol Med. 2025; 22: 6-13.
14. Yang M, Li X, Cai C, Liu C, Ma M, Qu W, et al. [^{18}F]FDG PET-CT radiomics signature to predict pathological complete response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer: a multicenter study. Eur Radiol. 2024; 34: 4352-63.
15. Yang N, Yue HL, Zhang BH, Chen J, Chu Q, Wang JX, et al. Predicting pathological response to neoadjuvant or conversion chemoimmunotherapy in stage IB-III non-small cell lung cancer patients using radiomic features. Thorac Cancer. 2023; 14: 2869-76.
16. Travis WD, Dacic S, Wistuba I, Sholl L, Adusumilli P, Bubendorf L, et al. IASLC multidisciplinary recommendations for pathologic assessment of lung cancer resection specimens after neoadjuvant therapy. J Thorac Oncol. 2020; 15: 709-40.
17. Yang C, Huang M, Li S, Chen J, Yang Y, Qin N, et al. Radiomics model of magnetic resonance imaging for predicting pathological grading and lymph node metastases of extrahepatic cholangiocarcinoma. Cancer Lett. 2020; 470: 1-7.
18. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Red Hook, NY, USA: Curran Associates Inc.; 2017: 4768-77.
19. Plackett RL. Karl pearson and the chi-squared test. Int Stat Rev. 1983; 51: 59-72.
20. Kim TK. T test as a parametric statistic. Korean J Anesthesiol. 2015; 68: 540-6.
21. Ostertagova E, Ostertag O, Kováč J. Methodology and application of the Kruskal–Wallis test. Appl Mech Mater. 2014; 611: 115-20.
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 837-45.
23. Bao X, Peng Q, Bian D, Ni J, Zhou S, Zhang P, et al. Short-term intra-and peri-tumoral spatiotemporal CT radiomics for predicting major pathological response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer. Eur Radiol. 2025; 35: 6052-64.

24. Ye G, Wu G, Qi Y, Li K, Wang M, Zhang C, et al. Non-invasive multimodal CT deep learning biomarker to predict pathological complete response of non-small cell lung cancer following neoadjuvant immunochemotherapy: a multicenter study. J Immunother Cancer. 2024; 12: e009348.
25. Eze C, Schmidt-Hegemann N-S, Sawicki LM, Kirchner J, Roengvoraphoj O, Käsmann L, et al. PET/CT imaging for evaluation of multimodal treatment efficacy and toxicity in advanced NSCLC—current state and future directions. Eur J Nucl Med Mol Imaging. 2021; 48: 3975-89.
26. Gao P, Xiao Q, Tan H, Song J, Fu Y, Xu J, et al. Interpretable multi-modal artificial intelligence model for predicting gastric cancer response to neoadjuvant chemotherapy. Cell Rep Med. 2024; 5: 101848.
27. Jiang C, Zhang X, Qu T, Yang X, Xiu Y, Yu X, et al. The prediction of pCR and chemosensitivity for breast cancer patients using DLG3, RADL and Pathomics signatures based on machine learning and deep learning. Transl Oncol. 2024; 46: 101985.

Cite this article as: Zheng S, Wang S, Feng Z, Liang J, Liu J, Yang X, et al. Temporal radiomics for non-invasive preoperative prediction of pathologic complete response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer. Cancer Biol Med. 2026; 23: 294-309. doi: 10.20892/j.issn.2095-3941.2025.0327