Cancer Immunology, Immunotherapy (2026) 75:42
https://doi.org/10.1007/s00262-025-04291-x

RESEARCH

# Prediction of tumor-infiltrating lymphocytes through habitat radiomics and exploration of response mechanisms in neoadjuvant immunochemotherapy-treated lung cancer

**Zuhan Geng[1] · Yihong Hu[2,3] · Shunchen Zhou[4] · Guangyao Wu[5] · Zhenyu Gong[6] · Ruiping Feng[7] · Zhangfeng Huang[1] · Peiyuan Mei[1] · Kuo Li[1] · Guanchao Ye[8] · Yongde Liao[1]**

Received: 15 September 2025 / Accepted: 28 December 2025
© The Author(s) 2026

## Abstract

**Background** Neoadjuvant immunochemotherapy (NAIC) induces tumor microenvironment remodeling in non-small cell lung cancer (NSCLC), presenting challenges for treatment response assessment. This study developed and validated a habitat radiomics approach for non-invasive prediction of tumor-infiltrating lymphocyte (TIL) status to evaluate NAIC response in NSCLC.

**Methods** This retrospective study enrolled 238 NSCLC patients following NAIC for clinical analysis, of which 201 patients met criteria for radiomics analysis. Patients were classified into TIL-positive and TIL-negative groups based on pathological assessment. Post-treatment computed tomography (CT) images were analyzed using K-means clustering to identify tumor habitat sub-regions for radiomic feature extraction. Seven machine learning algorithms were evaluated for TIL status prediction. Model interpretability was assessed through SHapley Additive exPlanations (SHAP) analysis. Single-cell RNA sequencing (scRNA-seq) data were analyzed to compare major pathological response (MPR) and non-MPR tumor microenvironments through cell type annotation, differentiation trajectory analysis, and intercellular communication network analysis.

**Results** Pre-treatment neutrophil-to-lymphocyte ratio (NLR) showed association with pathological response in multivariable analysis. The radiomics cohort was randomly divided 7:3 into training ($n = 140$) and test ($n = 61$) sets. The Random Forest model achieved an area under the receiver operating characteristic curve (AUC) of 0.823 (95% CI: 0.694–0.932) in the test set, and the habitat radiomics model stratified patients into high and low recurrence risk groups. Single-cell analysis identified immunosuppressive features in non-responding tumors, characterized by expansion of SERPINB9 + regulatory T cells (Tregs) that regulated suppressive intercellular communication networks.

**Conclusions** This study establishes a habitat radiomics model for non-invasive assessment of TIL status following neoadjuvant immunochemotherapy in NSCLC. The model shows reliable predictive performance and prognostic stratification capability, offering potential clinical utility for treatment response evaluation and patient selection.

**Keywords** Non-small cell lung cancer · Habitat radiomics · Tumor-infiltrating lymphocytes · Neoadjuvant immunochemotherapy · Tumor microenvironment

## Abbreviations

| | |
|---|---|
| ACC | Accuracy |
| AI | Artificial intelligence |
| AJCC | American joint committee on cancer |
| AUC | Area under the curve |
| BMI | Body mass index |
| BSA | Bovine serum albumin |
| CD8Tex | Exhausted CD8 + T cells |
| CEA | Carcinoembryonic antigen |
| CI | Confidence interval |
| DAPI | 4′, 6-Diamidino-2-phenylindole |
| DGE | Differential gene expression |
| GEO | Gene Expression Omnibus |
| GLCM | Gray level co-occurrence matrix |
| GLDM | Gray level dependence matrix |
| GLRLM | Gray level run length matrix |
| GLSZM | Gray level size zone matrix |
| GO | Gene ontology |
| H&E | Hematoxylin and eosin |

Zuhan Geng, Yihong Hu, and Shunchen Zhou have equally contributed to this work.

Extended author information available on the last page of the article

Published online: 27 January 2026

| | |
|---|---|
| IASLC | International association for the study of lung cancer |
| ICIs | Immune checkpoint inhibitors |
| IFN | Interferon |
| ISGs | Interferon-stimulated genes |
| KDE | Kernel density estimation |
| KNN | K-Nearest Neighbors |
| LDH | Lactate dehydrogenase |
| LT | Lymphotoxin |
| LUAD | Lung adenocarcinoma |
| LUSC | Lung squamous cell carcinoma |
| MHC | Major histocompatibility complex |
| MLP | Multi-layer perceptron |
| MPR | Major pathological response |
| mRMR | Minimum redundancy maximum relevance |
| NAIC | Neoadjuvant immunochemotherapy |
| NGTDM | Neighboring gray tone difference matrix |
| NLR | Neutrophil-to-lymphocyte ratio |
| non-MPR | Non-major pathological response |
| NPV | Negative predictive value |
| NSE | Neuron-specific enolase |
| NSCLC | Non-small cell lung cancer |
| OR | Odds ratio |
| PCA | Principal component analysis |
| pCR | Pathological complete response |
| PD-1 | Programmed death 1 |
| PD-L1 | Programmed death ligand 1 |
| PET/CT | Positron emission tomography/Computed tomography |
| PFS | Progression-free survival |
| PGE2 | Prostaglandin E2 |
| PLR | Platelet-to-lymphocyte ratio |
| PPV | Positive predictive value |
| RECIST | Response evaluation criteria in solid tumors |
| RFS | Recurrence-free survival |
| ROI | Region of interest |
| SCC | Squamous cell carcinoma antigen |
| scRNA-seq | Single-cell RNA sequencing |
| SD | Standard deviation |
| SHAP | SHapley additive explanations |
| SNN | Shared nearest neighbor |
| SVM | Support vector machine |
| TCGA | The cancer genome atlas |
| TILs | Tumor-infiltrating lymphocytes |
| TIM-3 | T-cell immunoglobulin and mucin-domain containing-3 |
| TLS | Tertiary lymphoid structure |
| TME | Tumor microenvironment |
| TNF | Tumor necrosis factor |
| TNM | Tumor, Node, Metastasis |
| Tregs | Regulatory T cells |
| TRIPOD | Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis |
| UMAP | Uniform manifold approximation and projection |
| VCAM1 | Vascular cell adhesion molecule 1 |

## Introduction

The integration of immune checkpoint inhibitors (ICIs) into neoadjuvant treatment protocols has transformed the therapeutic approach to locally advanced non-small cell lung cancer (NSCLC). Clinical trials have reported pathological response rates with combined immunochemotherapy that exceed historical outcomes with chemotherapy alone [1–3]. However, considerable heterogeneity exists in patient responses, and reliable predictive biomarkers remain limited.

The evaluation of treatment response following neoadjuvant immunochemotherapy (NAIC) presents unique complexities distinct from conventional chemotherapy assessment. Standard imaging criteria, including Response Evaluation Criteria in Solid Tumors (RECIST), demonstrate limited accuracy in predicting pathological response due to inflammatory changes, pseudoprogression phenomena, and complex tissue remodeling patterns unique to immunochemotherapy [4]. Post-treatment tumors frequently exhibit heterogeneous compositions of viable tumor cells, necrosis, fibrosis, and inflammatory infiltrates, creating substantial challenges for radiological interpretation [5]. Studies have identified distinct microenvironmental regions within tumors, including active tumor areas, necrotic zones, and fibrotic regions, each displaying unique imaging characteristics [6].

Tumor-infiltrating lymphocytes (TILs) represent key effector cells in anti-tumor immunity and have emerged as important biomarkers in immunochemotherapy. Multiple studies have established correlations between TIL density and clinical outcomes across various malignancies [7, 8]. In lung cancer specifically, patients with high TIL infiltration demonstrate superior pathological response rates and extended survival following ICI treatment [9]. However, current TIL assessment requires invasive tissue sampling, limiting its application in treatment monitoring and adaptive therapy strategies. Pre-surgical, non-invasive assessment of TIL status could identify patients who may benefit from treatment intensification or de-escalation, facilitate early evaluation of immunotherapy efficacy, and guide personalized therapeutic strategies.

Radiomics has offered novel approaches for extracting quantitative features from medical images that capture tissue properties. Radiomics extracts quantitative features from medical imaging based on the principle that images

represent data rather than mere pictures [10]. These imaging features, which include texture patterns, shape characteristics, and intensity profiles, may reveal biological processes such as cellular organization, vascular patterns, and metabolic states [11]. Previous investigations have established associations between radiomic signatures and molecular markers, immune infiltration patterns, and clinical outcomes across various malignancies [12–15]. Through unsupervised clustering of imaging features, habitat analysis can identify and quantify these sub-regions, potentially revealing characteristics invisible to conventional imaging assessment. Previous studies in NSCLC have shown that habitat-based radiomics can improve prediction accuracy for treatment outcomes [16]. Unlike conventional whole-tumor radiomics that treats the tumor as a homogeneous entity, habitat analysis can capture spatial variations in cellular composition, vascular patterns, and metabolic states within the tumor microenvironment. Research has demonstrated that PET-/CT-based habitat radiomics effectively predicted NSCLC recurrence risk [17]. Habitat-based radiomics addresses this heterogeneity through unsupervised clustering to identify tumor sub-regions with distinct imaging phenotypes, potentially reflecting underlying biological diversity.

Despite these advances, no studies have specifically investigated habitat radiomics for predicting TIL status in the post-neoadjuvant immunochemotherapy setting in lung cancer. Given the prognostic importance of TILs and the clinical need for non-invasive monitoring tools, this study aimed to develop and validate a habitat radiomics-based machine learning model for predicting TIL status in NSCLC patients following neoadjuvant immunochemotherapy and to elucidate the immunosuppressive mechanisms underlying treatment resistance through integrated single-cell transcriptomics.

## Methods

### Study design and patient population

This retrospective cohort study was conducted at Union Hospital, Tongji Medical College, Huazhong University of Science and Technology. The study protocol received approval from the Institutional Review Board (approval number UCT240116). Given the retrospective nature of the analysis, the requirement for informed consent was waived.

Patients with histologically confirmed NSCLC who received neoadjuvant immunochemotherapy combined with chemotherapy from August 2019 to February 2024 were evaluated for inclusion. Inclusion Criteria: (i) histologically confirmed stage IB-IIIB NSCLC according to the 8th edition of the American Joint Committee on Cancer (AJCC) staging system; (ii) patients who typically received at least ≥ 2 cycles of neoadjuvant immunochemotherapy prior to surgery; (iii) post-operative pathological evaluation performed on the treatment response of the tumor primary lesion and lymph nodes according to the International Association for the Study of Lung Cancer (IASLC) recommendations [18]. Exclusion criteria: (i) incomplete medical records or missing clinical data essential for analysis; (ii) inadequate CT image quality precluding radiomics analysis; (iii) interval between CT acquisition and surgery exceeding 30 days.

The single-cell RNA sequencing (scRNA-seq) data analyzed in this study were obtained from the Gene Expression Omnibus (GEO) database (accession number: GSE207422). This dataset consists of tumor samples from NSCLC patients who underwent neoadjuvant immunochemotherapy. Patients were clinically categorized based on pathological response into MPR ($n = 3$) and non-MPR ($n = 10$).

### Clinical and laboratory data collection

Clinical variables were systematically extracted from electronic medical records using a standardized data collection form. Baseline demographics included age at diagnosis, gender, body mass index (BMI), and smoking history. Tumor characteristics comprised histological subtype, primary tumor location, and clinical TNM staging.

Treatment-related variables included specific immunochemotherapy agent, including pembrolizumab, nivolumab, durvalumab, tislelizumab, camrelizumab, platinum-based doublet chemotherapy regimen, and number of completed cycles. Laboratory parameters were collected at baseline within 7 days before treatment initiation and post-treatment within 7 days before surgery. These included complete blood count with differential, comprehensive metabolic panel including liver and kidney function tests, lactate dehydrogenase (LDH), and traditional tumor markers such as carcinoembryonic antigen (CEA), squamous cell carcinoma antigen (SCC), and neuron-specific enolase (NSE). Inflammatory biomarkers were calculated as follows: neutrophil-to-lymphocyte ratio (NLR) = absolute neutrophil count/absolute lymphocyte count; PLR = platelet count/absolute lymphocyte count. Univariate and multivariate logistic regression analyses identified clinical predictors of pathological response. Recurrence-free survival (RFS) was defined as the time from the date of surgery to the date of radiologically or pathologically confirmed recurrence, death from any cause, or last follow-up. Survival analysis employed Kaplan–Meier method with log-rank test.

### Pathological assessment and TIL evaluation

Surgical specimens underwent standardized processing according to IASLC multidisciplinary recommendations for pathologic assessment following neoadjuvant therapy

[18]. Sections of 4 μm thickness were stained with hematoxylin and eosin (H&E). Pathological response assessment was performed by two experienced thoracic pathologists. The percentage of viable tumor cells was quantified across all tumor sections, with careful distinction from treatment-related changes including necrosis, fibrosis, and foam cell aggregates. MPR was defined as residual viable tumor cells ≤ 10%, while pCR represented absence of any viable tumor cells in both the primary tumor and sampled lymph nodes, consistent with IASLC criteria [18, 19].

Representative slide images were digitally scanned at 40 × or 100 × magnification and uploaded to the hospital pathology information system. TIL assessment was performed by pathologists blinded to clinical outcomes using semi-quantitative evaluation methods from pictures in the hospital information system [20, 21]. To ensure reliability, pathological reports noting extensive lymphocyte infiltration or prominent lymphoid aggregates were incorporated to minimize sampling bias. Patients were stratified into TIL-positive group with lymphocyte infiltration ≥ 10% and TIL-negative group with < 10% [20].

## CT image preprocessing and feature extraction

Post-treatment CT imaging was performed using contrast-enhanced images, acquired 60–70 s after intravenous administration of iodinated contrast medium. Post-treatment CT imaging was performed before surgery in all patients regardless of apparent imaging response. Even in patients achieving pathological complete response, post-treatment CT typically demonstrates residual imaging abnormalities including fibrosis and inflammatory changes that can be analyzed using radiomics. No patients were excluded based on the degree of imaging response. Detailed scanning parameters are provided in prior published work [22]. Regions of interests (ROIs) encompassing the tumor volume were manually outlined by a thoracic surgeon and radiologist with 10 years of imaging experience using ITK-SNAP software (version 3.8.0). Images underwent preprocessing to ensure reproducibility and minimize technical variability [22]. Images underwent isotropic resampling using cubic spline interpolation to standardize spatial resolution across scanners. Most tumors were resampled to $1 \times 1 \times 1$ mm$^3$ voxel size; larger tumors used $3 \times 3 \times 3$ mm$^3$ voxels. Intensity normalization was applied using z-score standardization within the lung window to minimize scanner-related variations.

A comprehensive set of radiomics features was extracted from each voxel within the tumor ROI using PyRadiomics with a local neighborhood approach. Tumor sub-regions were identified using K-means clustering algorithm from scikit-learn. The optimal number of clusters was determined using the Calinski–Harabasz index across a range of 3–8 clusters. Seven categories of features were calculated: shape-based features, first-order statistics, gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), gray level dependence matrix (GLDM), and neighboring gray tone difference matrix (NGTDM).

Radiomic features were extracted from each habitat sub-region. For samples with missing habitat regions, a k-nearest neighbors (KNN)-based interpolation method was used to fill missing values. Features from the habitat sub-regions were then combined for analysis. Feature selection was performed using the Mann–Whitney U test ($p < 0.05$) and Pearson correlation analysis ($|r| > 0.8$) with identical thresholds for both models, followed by LASSO regression for the habitat model (24 features) and mRMR for the whole-tumor model (all ROI, 25 features).

## Radiomic model development and interpretability

Seven machine learning algorithms were implemented: Logistic Regression, Random Forest, XGBoost, LightGBM, ExtraTrees, Multi-layer Perceptron (MLP), and GradientBoost. Hyperparameter optimization was performed through grid search on the training set. The search space for each algorithm is detailed in Supplementary Note 1. Label smoothing was applied to reduce overfitting. Fivefold cross-validation was conducted on the entire cohort for robustness assessment.

Model performance metrics included AUC, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1 score with 95% confidence intervals (CIs). Model performance metrics were calculated as follows: accuracy (ACC) = (true positives + true negatives) / total samples; positive predictive value (PPV) = true positives / (true positives + false positives); negative predictive value (NPV) = true negatives / (true negatives + false negatives). Model selection was based on AUC, with consideration for balanced sensitivity and specificity. DeLong test was used to compare AUC differences between models.

Subgroup analyses were conducted to evaluate model performance across patient subsets, including: (1) age ≤ 65 versus > 65; (2) clinical Stage II versus Stage III; and (3) smoking status. Model interpretability was assessed using SHAP analysis, which quantified each feature's impact on model predictions. Multiplex immunofluorescence staining for CD3, CD20, and FOXP3 was performed on representative samples in the test set to validate immune infiltration patterns.

## Immunofluorescence staining

Tissue specimens were fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned at 4 μm thickness. Two consecutive sections were selected for dual

immunofluorescence staining with different marker combinations. Following deparaffinization and rehydration, antigen retrieval was performed using citrate buffer (pH 6.0) at high temperature and pressure for 15 min. Sections were blocked with 3% BSA for 1 h at room temperature. Primary antibodies were applied overnight at 4 °C: CD3G antibody (1:100 dilution) and CD20 antibody (1:400 dilution). For Sect. 1, CD4 antibody (1:400, Cat No. 67786–1-Ig, Proteintech) and FOXP3 antibody (1:400, Cat No. 22228–1-AP, Proteintech) were co-incubated overnight at 4 °C, followed by secondary antibody incubation for 1 h at room temperature. For Sect. 2, CD25 antibody (1:400, Abcam, AB231441) and CD8 antibody (1:400, Abcam, ab316355) were incubated similarly, followed by the same secondary antibody combination. Nuclei were counterstained with 4′, 6-diamidino-2-phenylindole (DAPI) at 1:1000 dilution for 5–10 min and mounted with anti-fade mounting medium.

Images were acquired using a digital slide scanner with 20× or 40× objectives, with all fluorescence procedures performed in the dark to prevent photobleaching. Prominent tissue structural features served as anatomical landmarks to ensure spatial correspondence between consecutive sections. To assess the spatial distribution relationships between different immune cell subsets, images from consecutive sections were overlaid using Adobe Photoshop CS6 software. Specifically, the tissue architecture revealed by the DAPI channel was used as reference for precise alignment through translation, rotation, and minor scaling adjustments.

## scRNA-seq data processing and cell type annotation

The raw gene expression matrices were processed using the Seurat R package (v4.3). A standard quality control pipeline was applied to filter out low-quality cells. Cells were retained for downstream analysis if they expressed between 200 and 5,000 genes (nFeature_RNA) and had a mitochondrial gene content of less than 15%. Following filtration, gene expression counts for each cell were normalized using the "LogNormalize" method with a scale factor of 10,000.

The normalized data from all samples were merged, and the top 2000 highly variable genes were identified using the FindVariableFeatures function. The data were scaled, and Principal Component Analysis (PCA) was performed for linear dimensionality reduction. To correct for inter-sample batch effects, the Harmony algorithm was applied to the PCA embeddings using the RunHarmony function. The first 30 Harmony-corrected dimensions were used for all subsequent analyses. A shared nearest neighbor (SNN) graph was constructed on the Harmony-corrected embeddings using the FindNeighbors function. Cell clustering was performed using the Louvain algorithm at a resolution of 0.8 via the FindClusters function. For visualization, Uniform Manifold Approximation and Projection (UMAP) was run on the Harmony embeddings.

Cell clusters were annotated by evaluating the expression of canonical marker genes. Cluster-specific markers were identified by performing differential expression analysis using the FindAllMarkers function with Wilcoxon Rank Sum test, min.pct = 0.25, and logfc.threshold = 0.25. The identity of each cluster was assigned by cross-referencing these markers with established literature. Key markers used for annotation included: epithelial cells (EPCAM, KRT8, KRT18, KRT19); fibroblasts (COL1A1); endothelial cells (VWF); B cells (MS4A1, CD79A); plasma cells (IGHG1); and myeloid cells (LYZ, CSF3R).

## Differential gene expression and pseudotime trajectory analysis

Differential gene expression (DGE) analysis between cell populations of interest was performed using the FindMarkers function in Seurat. Genes with an adjusted $p$-value < 0.05 and an absolute log2 fold change > 0.25 were considered significant. To investigate the biological functions of these differentially expressed genes, Gene Ontology (GO) enrichment analysis was conducted using the clusterProfiler R package.

The developmental trajectory of CD4 + T cells was inferred using the Monocle 2 R package. The CD4 + T-cell subset was extracted from the Seurat object to create a CellDataSet object. Genes that were differentially expressed between the CD4 + T-cell clusters were used for ordering the cells in pseudotime. The DDRTree algorithm was used for dimensionality reduction, and the trajectory was constructed using the orderCells function.

## Cell–cell communication and survival analysis

Intercellular communication networks were inferred and analyzed using the CellChat R package. The MPR and non-MPR datasets were analyzed separately to identify condition-specific signaling patterns. CellChat's curated ligand–receptor interaction database was used to calculate the communication probabilities between T-cell subsets. Differential signaling pathways and interactions between the two conditions were subsequently identified and visualized.

To evaluate the clinical relevance of SERPINB9$^{+}$ regulatory T cells (Tregs), we performed differential gene expression analysis between SERPINB9$^{+}$ and SERPINB9$^{-}$ Treg subsets using the MAST algorithm in single-cell RNA-seq data, adjusting for donor/batch effects (immune_id), sequencing depth/detection rate (nCount_RNA, nFeature_RNA), and mitochondrial fraction (percent.mt). Differentially expressed genes were identified using stringent criteria: log2 fold change > 1, difference in detection rate

(Δpct) > 0.2, and false discovery rate (FDR) < 0.01. To preserve Treg identity in the gene signature, CD4 and FOXP3 were additionally included regardless of differential expression thresholds. The resulting gene signature score was computed in the TCGA-NSCLC cohort (LUAD and LUSC), and patients were dichotomized at the median value. Survival differences were assessed using Kaplan–Meier curves with log-rank tests and validated by multivariable Cox proportional hazards models. The prognostic performance was further validated in an external neoadjuvant immunotherapy cohort (GSE135222).

### Quality standards and statistical analysis

This study adhered to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines for prediction model development in Supplementary Table S1 [23].

Statistical analyses were performed using R (version 4.2.1) and Python (version 3.7). Continuous variables were expressed as mean ± standard deviation (SD). Categorical variables were presented as frequencies and percentages. Between-group comparisons used Student's t-test or Mann–Whitney U test for continuous variables and Chi-square or Fisher's exact test for categorical variables, with significance set at $\alpha = 0.05$. Machine learning model training was performed on a computer with NVIDIA RTX 4060Ti graphics card. Single-cell analysis and visualization were completed on local personal servers and the HPC Platform of Huazhong University of Science and Technology.

## Results

### Patient characteristics and clinical cohort analysis

From August 2019 to February 2024, a total of 243 patients with NSCLC received neoadjuvant immunochemotherapy and underwent surgical resection at our institution. After applying inclusion and exclusion criteria, 238 patients comprised the clinical analysis cohort, with 201 patients meeting all requirements for radiomics analysis (Fig. 1).

The clinical cohort demonstrated a male predominance of 89.4%, with a median age of 62 years ranging from 37 to 76. Squamous cell carcinoma represented the most prevalent histological subtype at 66.8%, followed by adenocarcinoma at 25.2%. The majority of patients presented with stage IIIA disease at 44.1%, followed by stage IIIB at 33.1% and stage IIB at 10.9%. Regarding immunochemotherapy agents, tislelizumab was most frequently administered at 42.0%, followed by pembrolizumab at 20.5%, and camrelizumab at 12.1%. MPR was achieved in 129 patients representing 54.2%, while pCR was observed in 79 patients representing 33.1% (Fig. 2A). With a maximum follow-up of 48.5 months, 65 recurrence events were observed representing 29.5%.

Univariate and multivariate analyses of clinical and laboratory parameters revealed factors associated with treatment response (Supplementary Table S2). Multivariate logistic regression identified three independent factors associated with MPR achievement. Post-treatment PLR demonstrated an adjusted odds ratio (OR) of 1.01 (95% CI: 1.01–1.01, $p = 0.008$), pre-treatment NLR showed an adjusted OR of 1.17 (95% CI: 1.01–1.37, $p = 0.047$), and baseline creatinine exhibited an adjusted OR of 1.02 (95% CI: 1.01–1.04, $p = 0.020$) (Fig. 2B).

### TIL status and pathological response correlation

TIL assessment identified 157 patients (66.0%) as TIL-positive and 81 patients (34.0%) as TIL-negative. The relationship between TIL status and pathological response demonstrated strong associations. Among TIL-positive patients, 110 achieved MPR (70.1%) compared to only 47 of TIL-negative patients (23.5%). Similarly, pCR rates were significantly higher in TIL-positive patients at 45.9 versus 8.6% in TIL-negative patients (Fig. 2C).

TIL status showed correlations with both pathological response and RFS. The Sankey diagram tracked patient distribution from TIL status through pathological response to clinical outcomes (Fig. 2D). Among TIL-positive patients, a greater proportion achieved MPR compared to TIL-negative patients. The flow widths represent patient numbers following each trajectory, showing that TIL-positive patients had a higher tendency to achieve MPR, and those achieving MPR demonstrated improved event-free survival rates.

Pathological response markers demonstrated prognostic associations in the clinical cohort (Fig. 2E). Patients achieving pCR showed improved RFS compared to non-pCR patients ($p = 0.010$), with more pronounced separation for MPR ($p < 0.001$). TIL status also provided stratification ($p = 0.001$). In the non-pCR subgroup, TIL-positive patients showed improved RFS compared to TIL-negative patients ($p = 0.0409$), demonstrating the prognostic value of TIL status. In contrast, traditional baseline clinical staging variables did not achieve prognostic stratification in this immunochemotherapy-treated population (Supplementary Figure S1). Neither clinical stage ($p = 0.4877$), pre-treatment T stage ($p = 0.3993$), pre-treatment N stage ($p = 0.1718$), lesion location ($p = 0.2050$), nor smoking status ($p = 0.5926$) demonstrated significant associations with RFS. Most stage III patients who achieved downstaging subsequently underwent surgery, potentially leading to better outcomes.

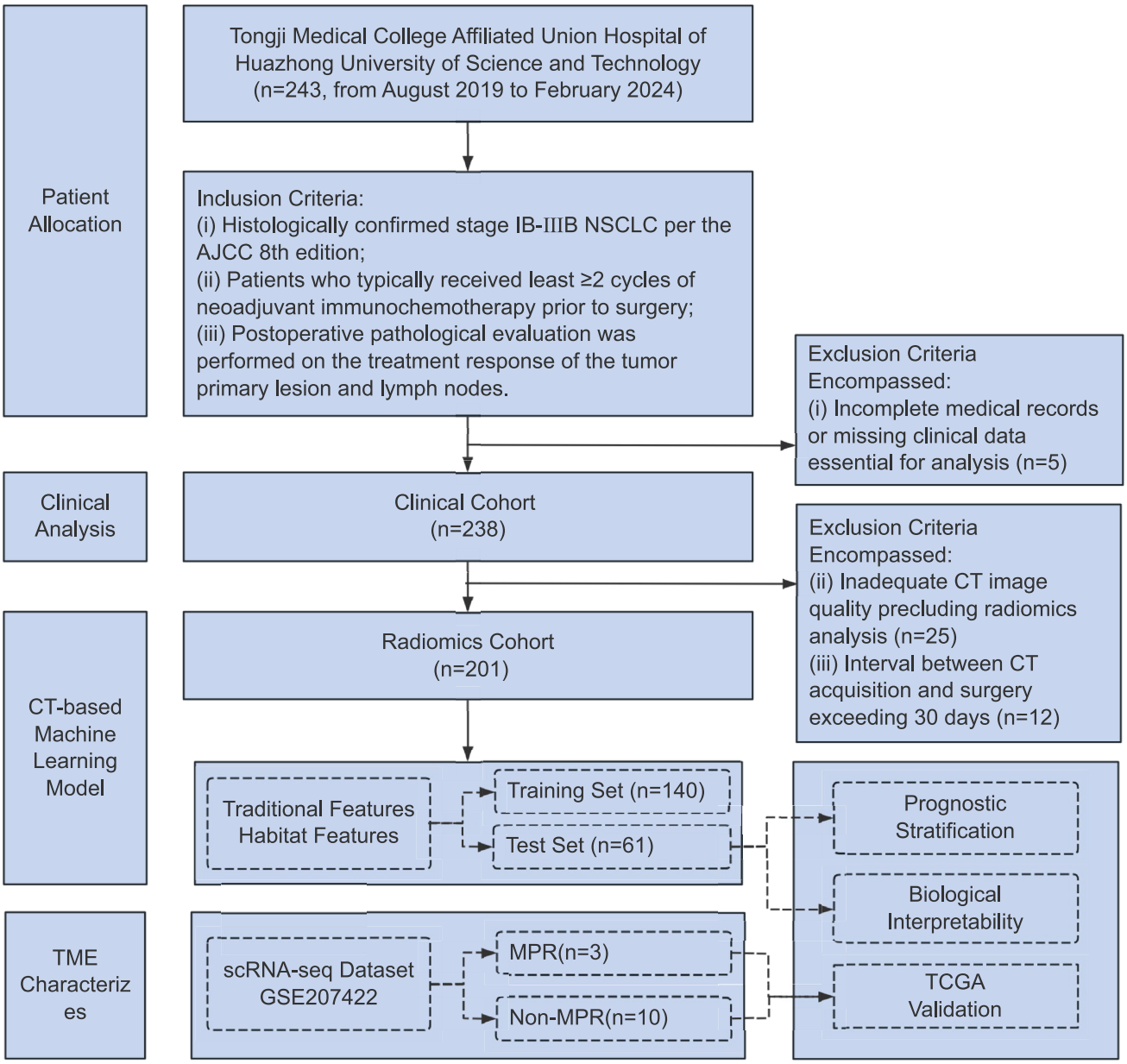


**Fig. 1** Study Enrollment and Patient Selection Flowchart. This flowchart illustrates the patient selection process and cohort allocation for the study

## Habitat identification and radiomic analysis

Figure 3A illustrates selected 6 cases showing the CT imaging and pathological response. While all patients showed varying degrees of tumor size changes on CT imaging, conventional imaging alone could not reliably distinguish between those with favorable versus poor treatment response. In the TIL-positive group, H&E staining reveals extensive lymphocyte infiltration replacing tumor tissue, minimal residual viable tumor cells, and prominent stromal fibrosis. In contrast, the TIL-negative group pathological examination revealed sheets of viable tumor cells with sparse lymphocyte infiltration and minimal treatment-related changes.

Following lesion segmentation, unsupervised clustering analysis identified three separate habitat sub-regions within post-treatment tumors (Fig. 3B). The Calinski–Harabasz index revealed maximum value at $k=3$ clusters with an index value of 393,162, with substantial decline at $k=4$ with an index value of 363,465 (shown in Fig. 3C). The resulting heatmaps displayed heterogeneous feature distributions across the lesion, where warmer colors such as red and

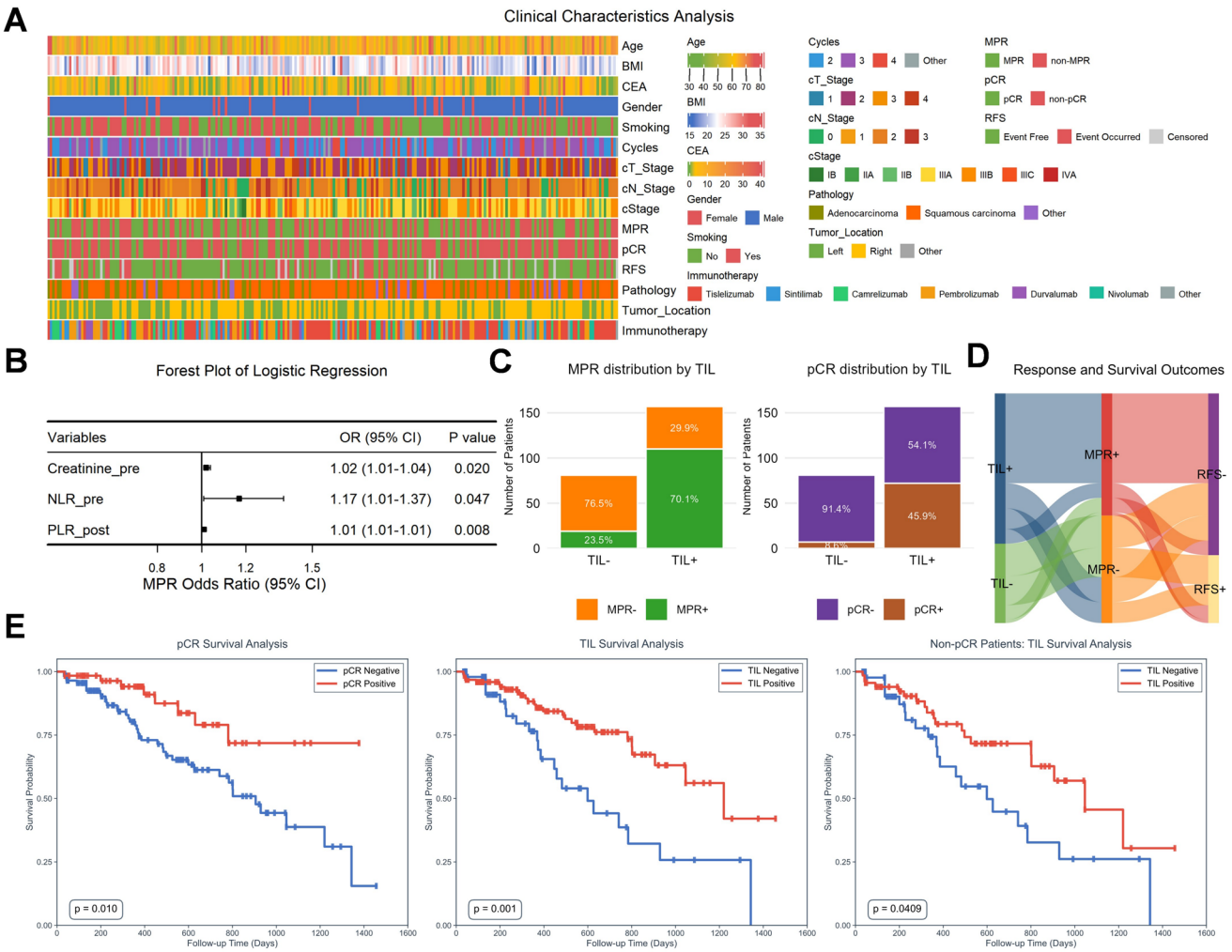


**Fig. 2** Clinical Characteristics and Pathological Response Correlations. **A** Distribution of baseline characteristics and pathological response outcomes in the clinical cohort. **B** Forest plot of multivariate logistic regression analysis. **C** Comparison of pathological response rates between TIL-positive and TIL-negative groups. **D** Sankey diagram tracks the distribution of patients across TIL, MPR, and RFS outcomes. **E** Kaplan–Meier survival curves comparing RFS across different prognostic markers. MPR, major pathological response; non-MPR, non-major pathological response; OR, odds ratio; pCR, pathological complete response; PLR, platelet-to-lymphocyte ratio; RFS, recurrence-free survival; TIL, tumor-infiltrating lymphocyte

yellow represented higher radiomics feature values, while cooler colors like blue indicated lower values following normalization (Fig. 3D).

## Machine learning model performance

The radiomics cohort was randomly divided 7:3 into training set ($n = 140$) and independent test set ($n = 61$). The distribution of clinical characteristics was balanced between training and test sets, with no significant differences in key variables (Table 1). The initial radiomic feature extraction yielded 5,505 features from the three habitat sub-regions. Following feature selection, 984 features were retained after Mann–Whitney U test, 137 features after Pearson correlation analysis, and 24 optimal features were selected by LASSO regression for final model development (Supplementary Note 2 and Figure S2A, B).

Among the 7 machine learning algorithms evaluated, Random Forest (RF) demonstrated the most balanced performance with an AUC of 0.823 (95% CI: 0.694–0.932) in the test set (Fig. 4A). The fivefold cross-validation on the entire cohort confirmed model robustness (Supplementary Figure S2C). Comprehensive evaluation of seven machine learning algorithms showed performance for TIL prediction in the test set (Table 2). The radar chart analysis revealed varied performance characteristics across algorithms, with the RF model showing balanced performance across multiple metrics in the test set, with sensitivity of 0.786, specificity of 0.737, and F1 score of 0.825 (Supplementary Figure S2D). Although XGBoost achieved a slightly

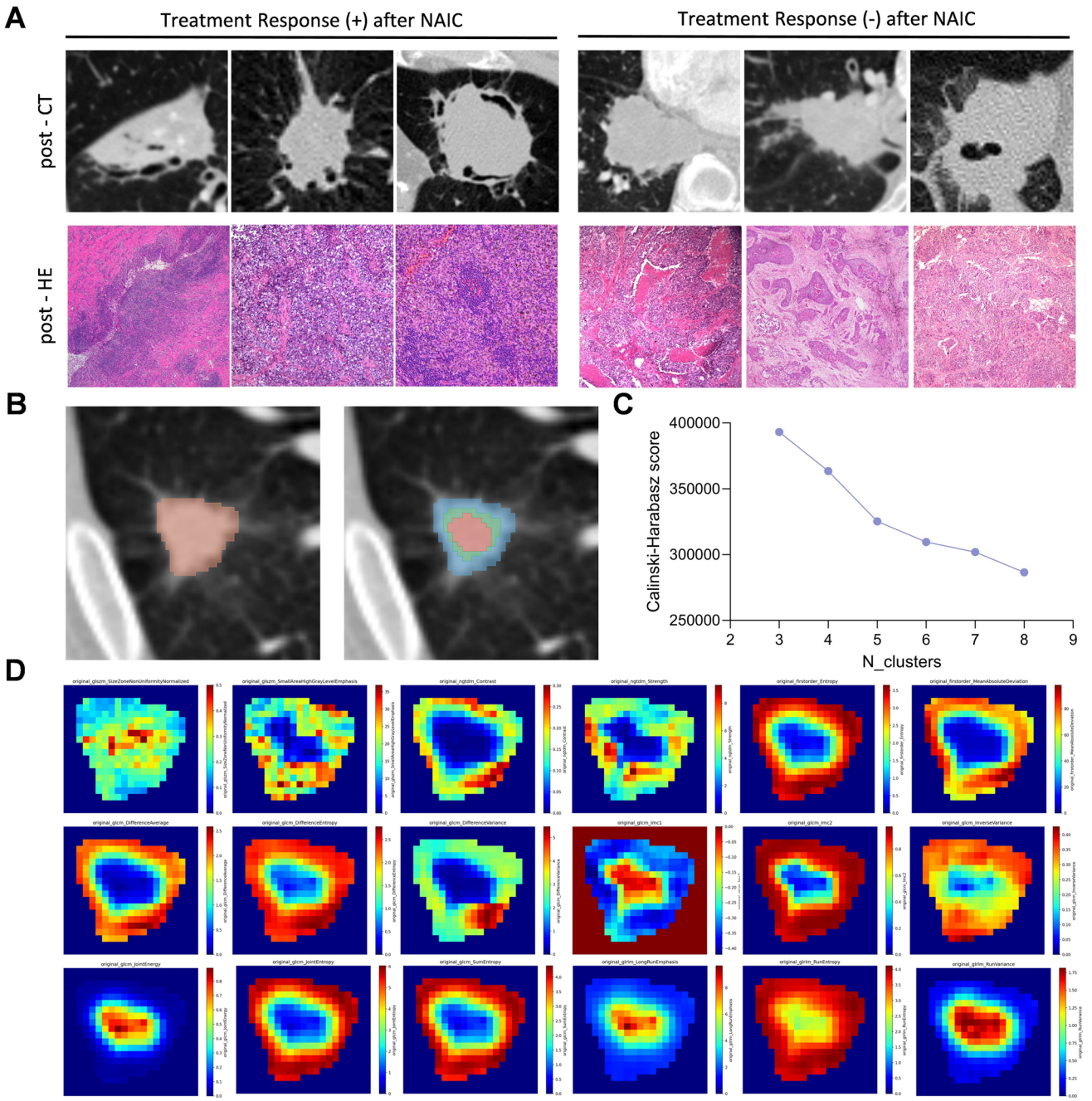


**Fig. 3** Habitat Radiomics Analysis of Post-Treatment Tumors. **A** Representative cases demonstrating computed tomography imaging characteristics and corresponding pathological findings. **B** Visualization of tumor habitat sub-regions identified through unsupervised K-means clustering. Different colors represent distinct habitat sub-regions within the tumor. **C** Calinski–Harabasz index plot for determining optimal cluster number. **D** The heatmap displays spatial heterogeneity in radiomics features, where warmer colors (red and yellow) represent higher feature values and cooler colors (blue) indicate lower values. Abbreviations: CT, computed tomography; H&E, hematoxylin and eosin; MPR, major pathological response; non-MPR, non-major pathologic response; TIL, tumor-infiltrating lymphocyte

higher AUC of 0.845 (95% CI: 0.734–0.929), its sensitivity of 0.762 and specificity of 0.632 were notably lower than RF. Therefore, RF was selected as the optimal model for its superior balance between sensitivity and specificity. Confusion matrices demonstrated the classification accuracy of each model, while probability distribution histograms illustrated the separation between predicted classes (Fig. 4B, C). DeLong test results comparing model performance are

**Table 1** Baseline characteristics of training and test sets

| Variables | Total ($n=201$) | Training set (n = 140) | Test set ($n=61$) | $P$ |
|---|---|---|---|---|
| Age, mean ± SD | 61.02 ± 7.15 | 60.84 ± 7.47 | 61.43 ± 6.40 | 0.596 |
| BMI, mean ± SD | 22.74 ± 2.95 | 22.78 ± 3.03 | 22.64 ± 2.78 | 0.753 |
| CEA, mean ± SD | 5.20 ± 7.97 | 5.61 ± 8.98 | 4.27 ± 4.86 | 0.273 |
| Gender, n (%) | | | | 0.584 |
| Male | 181 (90.05) | 125 (89.29) | 56(91.80) | |
| Female | 20 (9.95) | 15 (10.71) | 5 (8.20) | |
| Smoking, n (%) | | | | 0.871 |
| Yes | 107 (53.23) | 74 (52.86) | 33 (54.10) | |
| No | 94 (46.77) | 66 (47.14) | 28 (45.90) | |
| Pathology, n (%) | | | | 0.325 |
| Squamous carcinoma | 137 (68.16) | 94 (67.14) | 43 (70.49) | |
| Adenocarcinoma | 49 (24.38) | 33 (23.57) | 16 (26.23) | |
| Other | 15 (7.46) | 13 (9.29) | 2 (3.28) | |
| Location of lesion, n (%) | | | | 0.573 |
| Left | 83 (41.29) | 56 (40.00) | 27 (44.26) | |
| Right | 118 (58.71) | 84 (60.00) | 34 (55.74) | |
| Cycles of ICIs, n (%) | | | | 0.007* |
| 2 | 53 (26.37) | 44 (31.43) | 9 (14.75) | |
| 3 | 111 (55.22) | 75 (53.57) | 36 (59.02) | |
| 4 | 25 (12.44) | 17 (12.14) | 8 (13.11) | |
| Other | 12 (5.97) | 4 (2.86) | 8 (13.11) | |
| Immunotherapeutic agents, n (%) | | | | 0.220 |
| Tislelizumab | 85 (42.29) | 62 (44.29) | 23 (37.70) | |
| Pembrolizumab | 39 (19.40) | 31 (22.14) | 8 (13.11) | |
| Sintilimab | 26 (12.94) | 14 (10.00) | 12 (19.67) | |
| Camrelizumab | 23 (11.44) | 17 (12.14) | 6 (9.84) | |
| Nivolumab | 8 (3.98) | 5 (3.57) | 3 (4.92) | |
| Durvalumab | 7 (3.48) | 4 (2.86) | 3 (4.92) | |
| Other | 13 (6.47) | 7 (5.00) | 6 (9.84) | |
| cT Stage, n (%) | | | | 0.452 |
| 1 | 19 (9.45) | 16 (11.43) | 3 (4.92) | |
| 2 | 75 (37.31) | 49 (35.00) | 26 (42.62) | |
| 3 | 62 (30.85) | 44 (31.43) | 18 (29.51) | |
| 4 | 45 (22.39) | 31 (22.14) | 14 (22.95) | |
| cN Stage, n (%) | | | | 0.029* |
| 0 | 32 (15.92) | 20 (14.29) | 12 (19.67) | |
| 1 | 28 (13.93) | 24 (17.14) | 4 (6.56) | |
| 2 | 118 (58.71) | 76 (54.29) | 42 (68.85) | |
| 3 | 23 (11.44) | 20 (14.29) | 3 (4.92) | |
| cStage, n (%) | | | | 0.691 |
| IIA | 11 (5.47) | 7 (5.00) | 4 (6.56) | |
| IIB | 20 (9.95) | 15 (10.71) | 5 (8.20) | |
| IIIA | 89 (44.28) | 60 (42.86) | 29 (47.54) | |
| IIIB | 72 (35.82) | 50 (35.71) | 22 (36.07) | |
| IIIC | 9 (4.48) | 8 (5.71)) | 1 (1.64) | |
| RFS, n (%) | | | | 0.476 |
| Event-free | 135 (67.16) | 92 (65.71) | 43 (70.49) | |
| Event occurred | 50 (24.88) | 38 (27.14) | 12 (19.67) | |
| Censored | 16 (7.96) | 10 (7.14) | 6 (9.84) | |
| TIL, n (%) | | | | 0.969 |
| Negative | 63 (31.34) | 44 (31.43) | 19 (31.15) | |
| Positive | 138 (68.66) | 96 (68.57) | 68.85) | |

*$P<0.05$. SD—standard deviation, BMI—body mass index, CEA—carcinoembryonic antigen, ICIs—immune checkpoint inhibitors, cT—clinical tumor, cN—clinical node, cStage: clinical stage, RFS—recurrence-free survival, TIL—tumor-infiltrating lymphocytes

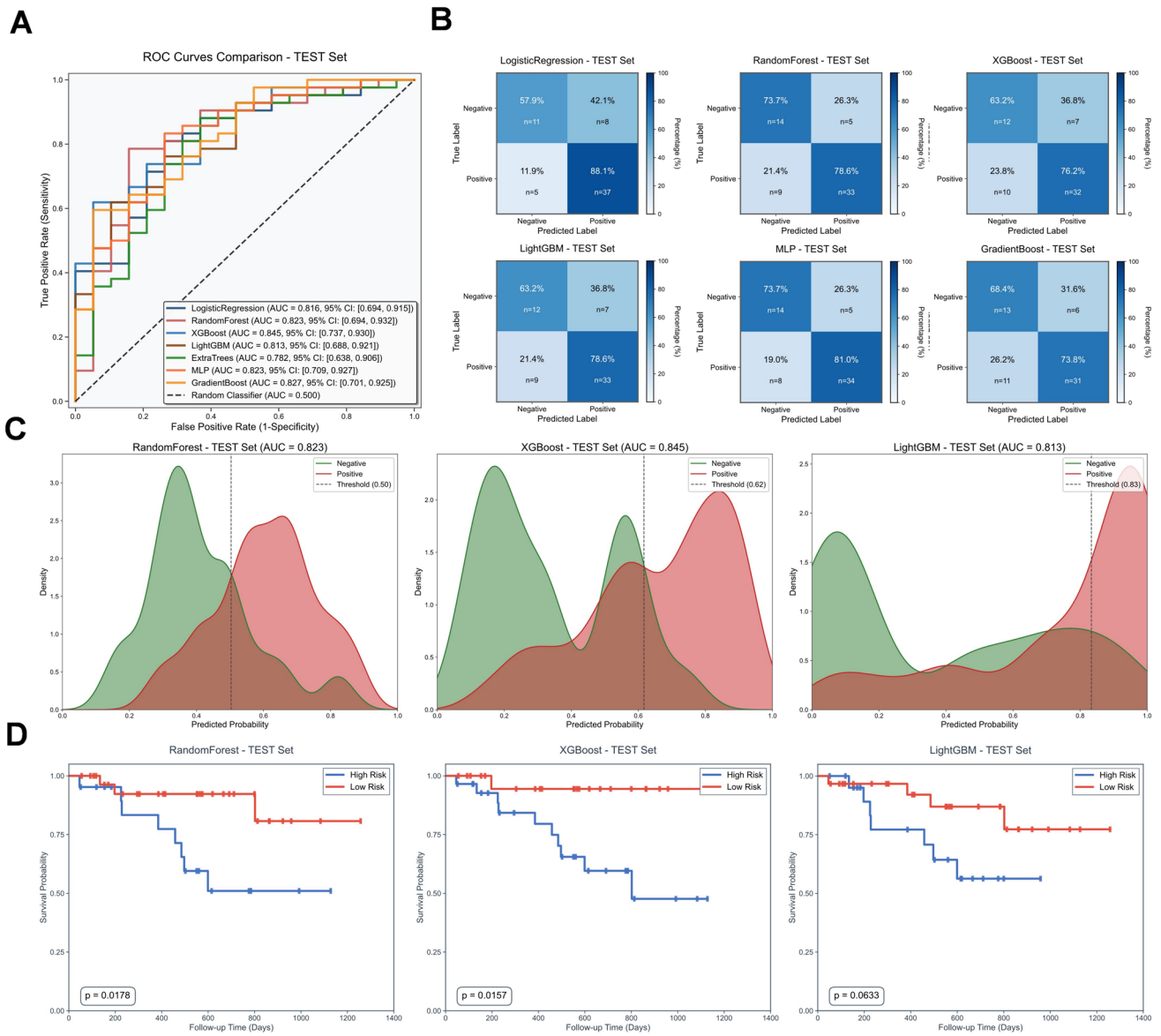


**Fig. 4** Machine Learning Model Performance and Prognostic Stratification. Receiver operating characteristic curves comparing seven machine learning algorithms for tumor-infiltrating lymphocyte status prediction in the test set. **B** Confusion matrices for each machine learning algorithm. Numbers in each cell indicate patient counts, with color intensity representing prediction accuracy. **C** Probability distribution for predicted tumor-infiltrating lymphocyte status of habitat radiomics models. Clear separation between distributions indicates good model discriminative ability. **D** Kaplan–Meier curves demonstrating prognostic stratification capability of habitat radiomics models. Abbreviations: AUC, area under the curve; CI, confidence interval; RFS, recurrence-free survival; TIL, tumor-infiltrating lymphocyte

**Table 2** Performance metrics of machine learning algorithms for TIL prediction in the test set

| Model | Accuracy | AUC | 95% CI | Sensitivity | Specificity | F1 |
|---|---|---|---|---|---|---|
| LogisticRegression | 0.787 | 0.816 | [0.695, 0.914] | 0.881 | 0.579 | 0.851 |
| RandomForest | 0.770 | 0.823 | [0.700, 0.926] | 0.786 | 0.737 | 0.825 |
| XGBoost | 0.721 | 0.845 | [0.734, 0.929] | 0.762 | 0.632 | 0.790 |
| LightGBM | 0.738 | 0.813 | [0.696, 0.917] | 0.786 | 0.632 | 0.805 |
| ExtraTrees | 0.738 | 0.782 | [0.639, 0.905] | 0.762 | 0.684 | 0.800 |
| MLP | 0.787 | 0.823 | [0.697, 0.925] | 0.810 | 0.737 | 0.840 |
| GradientBoost | 0.721 | 0.827 | [0.703, 0.926] | 0.738 | 0.684 | 0.785 |

AUC—Area under the curve, CI—confidence interval, F1—F1 score, MLP—multi-layer perceptron

shown in Supplementary Figure S2E. For comparison, the whole-tumor radiomics model achieved an AUC of 0.679 (95% CI: 0.526–0.814) (Supplementary Figure S2F and Supplementary Table S3).

Subgroup analyses revealed that the habitat radiomics model maintained performance across patient subsets (Supplementary Table S4), indicating predictive capability regardless of age subtype or disease stage. The radiomics model demonstrated reliable prognostic stratification capability in the test set (Fig. 4C). Using the optimal threshold determined by Youden index, patients were stratified into high-risk and low-risk groups based on predicted TIL status. The Random Forest model successfully stratified patients by recurrence risk, with low-risk patients showing significantly improved RFS compared to high-risk patients ($p = 0.0178$). Similar prognostic stratification was achieved with XGBoost($p = 0.0157$) and.LightGBM ($p = 0.0633$) SHAP value analysis provided insights into RF, XGBoost, and LightGBM model decision-making processes (Fig. 5A–C).

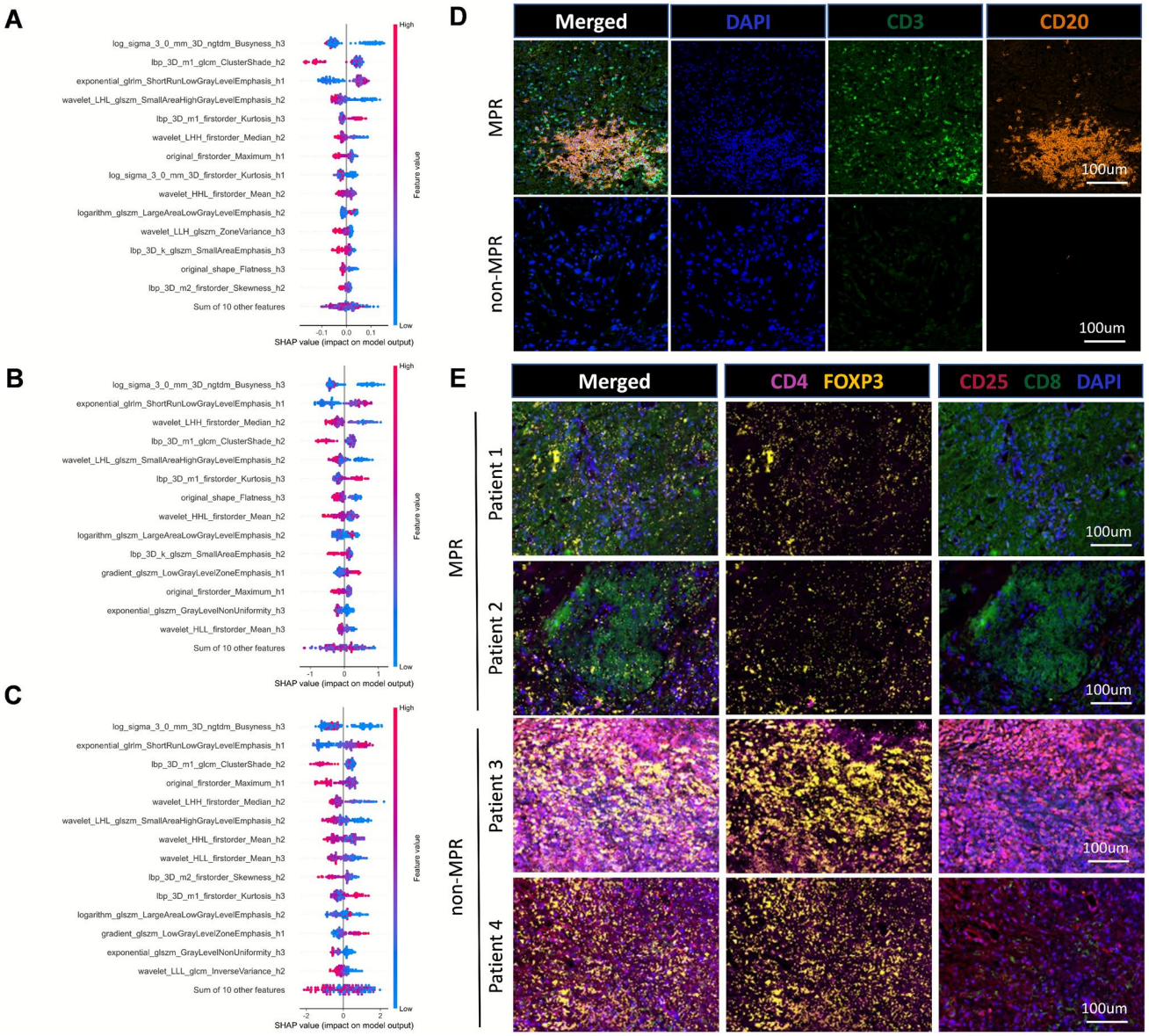


**Fig. 5** Model Interpretability and Biological Validation. (A-C) SHapley Additive exPlanations value analysis for Random Forest, XGBoost, and LightGBM models. Features are ranked by importance from top to bottom. (D-E) Multiplex immunofluorescence staining. CD, cluster of differentiation; DAPI, 4′,6-diamidino-2-phenylindole; FOXP3, forkhead box P3; MPR, major pathological response; non-MPR, non-major pathological response; SHAP, SHapley Additive exPlanations; TIL, tumor-infiltrating lymphocyte

### Biological validation

To validate the biological basis of our imaging-based TIL predictions, we performed multiplex immunofluorescence staining on representative tissue sections in the test set. Immune cell aggregates in tumor tissues from an MPR patient with high TIL infiltration (Fig. 5D). TIL-negative/non-MPR tumors exhibited elevated FOXP3 + Treg infiltration compared to TIL-positive/MPR tumors (Fig. 5E). This finding provides links between the imaging habitat phenotype and immunosuppressive cellular state, strengthening the biological interpretation that habitat radiomics captures tumor microenvironment (TME) differences. CD3 + T-cell infiltration was higher in MPR tumors compared to non-MPR, indicating more active anti-tumor immune state. In contrast, FOXP3 + Tregs were more abundant in non-MPR cases, suggesting immunosuppressive microenvironment associated with poor pathological response.

## An immunosuppressive TME driven by treg differentiation characterizes therapy resistance

To further dissect the heterogeneity underlying these immune patterns, we analyzed a public single-cell RNA sequencing dataset from lung cancer patients. Analysis of a public scRNA-seq dataset from chemoimmunochemotherapy-treated NSCLC tumors (GEO: GSE207422) revealed significant remodeling of the TME between responders and non-responders. A total of 17,533 cells from this dataset were clustered, identifying nine distinct cell populations, including major lymphoid (T cells, B cells), myeloid, and epithelial compartments (Fig. 6A, Supplementary Figure S3). The non-MPR TME showed a lower T-cell proportion with a relative increase in the myeloid compartment, an architecture consistent with an immune-excluded phenotype (Fig. 6A). To characterize the T-cell compartment in detail, we performed sub-clustering on all 6,637 identified T cells. Qualitative analysis of the T-cell compartment demonstrated a compositional shift within the non-MPR TME, characterized by expansion of both Tregs and exhausted CD8 + T cells (CD8Tex) (Fig. 6B, C). This dual enrichment indicates a state of immune suppression.

To explore the origin of this expanded Treg population, pseudotime trajectory analysis of the CD4 + T-cell lineage reconstructed a continuous differentiation pathway from a naive T-cell root toward a Treg fate (Fig. 6D). This suggests that accumulation of Tregs in the non-MPR TME arises from differentiation of naive T cells. We observed clear transcriptional reprogramming along the trajectory. As cells progressed toward the Treg fate, a naive T-cell signature (TCF7, CCR7) was replaced by genes essential for Treg identity and function, including the transcription factor FOXP3 and co-inhibitory receptors. The serine proteinase inhibitor SERPINB9 was identified as a key component of this terminal suppressive program (Supplementary Figure S4).

### *SERPINB9 + tregs drive therapeutic resistance by remodeling the communication network*

We isolated Treg cells and grouped them by SERPINB9 expression levels, categorizing them as either SERPINB9 + Treg or SERPINB9- Treg. Notably, SERPINB9 + Treg cells were significantly more abundant in non-MPR patients (Fig. 6E).

The differential expression analysis identified 44 candidate genes distinguishing SERPINB9$^{+}$ from SERPINB9$^{-}$ Tregs (full gene list and statistics in Supplementary Table S5). In the external neoadjuvant immunotherapy cohort GSE135222, SERPINB9 expression alone failed to stratify patient outcomes (log-rank $p > 0.05$; Supplementary Figure S5A-B). In contrast, the multi-gene SERPINB9$^{+}$ Treg signature achieved significant prognostic separation (log-rank $p < 0.05$; Supplementary Figure S5C). Consistent prognostic associations were observed for both progression-free survival and overall survival in the TCGA cohort (Fig. 6F; Supplementary Figure S5D). Collectively, these findings demonstrate that the SERPINB9$^{+}$ Treg-associated gene signature captures clinically meaningful immunosuppressive biology and may serve as a prognostic biomarker in NSCLC patients receiving immunotherapy.

Transcriptional profiling revealed that SERPINB9 + Tregs express a gene signature consistent with inhibition of effector T cells. Their transcriptional program includes high expression of FOXP3, alongside TNF receptor superfamily members TNFRSF9 (4-1BB) and TNFRSF4 (OX40) (Fig. 7A). High co-expression of these receptors characterizes a stable, proliferative, and suppressive tumor-infiltrating Treg phenotype. The signature also included interferon-stimulated genes (ISGs) (Fig. 7A), suggesting these Tregs actively respond to chronic Type I IFN signaling. Gene Ontology (GO) analysis confirmed their genetic program was dominated by pathways related to "regulation of T-cell activation" and "negative regulation of immune effector process." SERPINB9 + Tregs exhibit a distinct survival mechanism (Fig. 7B). High expression of SERPINB9, a stoichiometric inhibitor of Granzyme B, provides resistance to cytotoxic T-cell-mediated lysis, enabling their persistence within the inflamed tumor core.

Analysis of the intercellular communication network revealed two divergent signaling architectures defining therapeutic outcomes. The communication network in responders was dominated by pathways associated with productive adaptive immune response, specifically MHC-I and MHC-II antigen presentation pathways, along with TNF and LT (lymphotoxin) pathways. In contrast, the non-MPR network showed rewiring favoring an

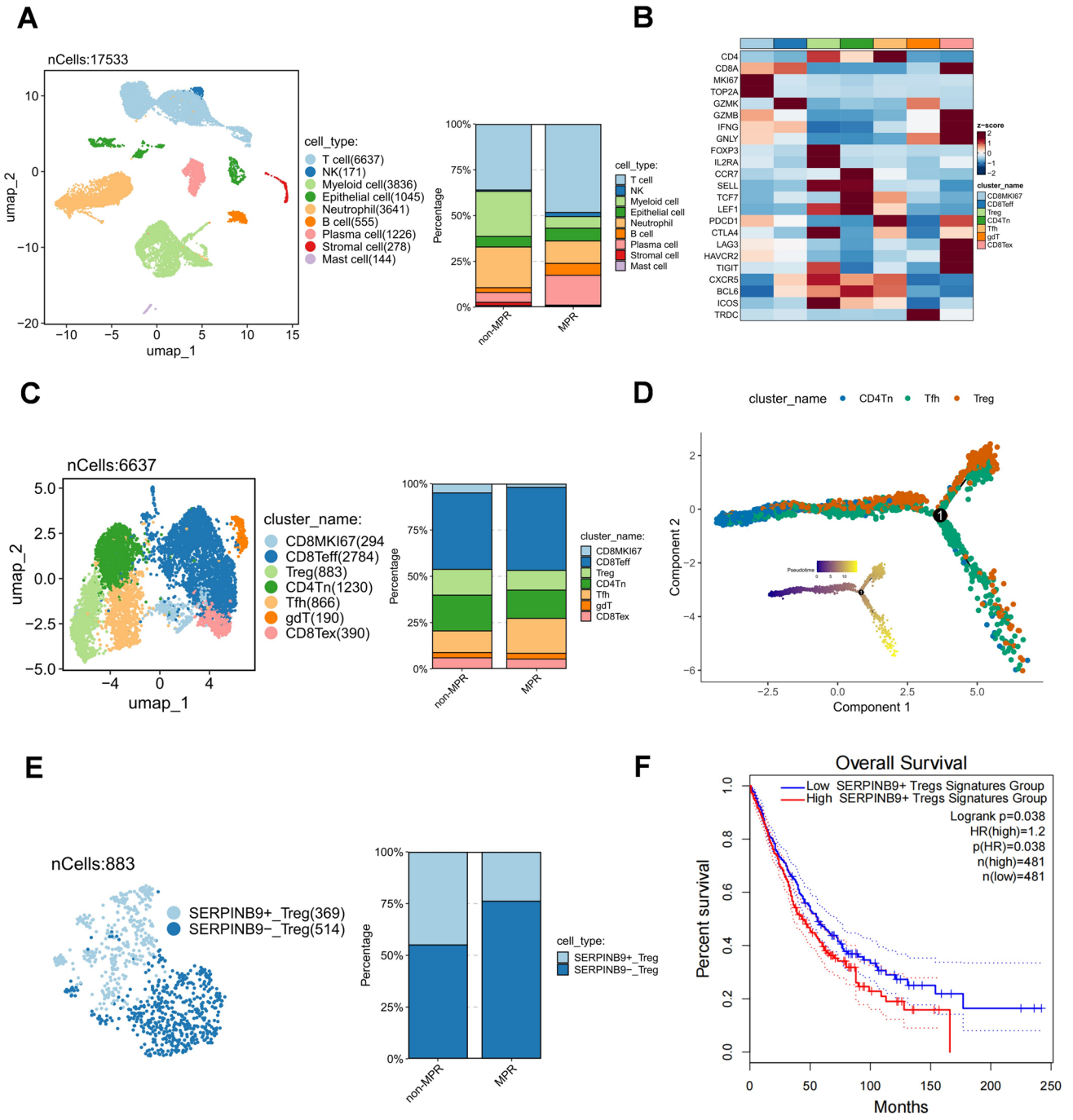


**Fig. 6** Single-Cell Transcriptomic Analysis Reveals Immunosuppressive Tumor Microenvironment. (A-C) Cells were clustered into nine distinct cell populations including T cells, B cells, plasma cells, myeloid cells, fibroblasts, endothelial cells, and epithelial cells. T cells were further classified into distinct subsets. Shown relative abundance of each T-cell subset stratified by pathological response status. **D** Pseudotime trajectory analysis of CD4-positive T-cell differentiation. **E** Comparison of SERPINB9-positive regulatory T-cell proportions between MPR and non-MPR. **F** Kaplan–Meier survival curves for SERPINB9-positive regulatory T-cell signature in TCGA. Abbreviations: CD, cluster of differentiation; CD8Tex, exhausted CD8-positive T cells; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MPR, major pathological response; non-MPR, non-major pathological response; OS, overall survival; PFS, progression-free survival; scRNA-seq, single-cell RNA sequencing; TCGA, The Cancer Genome Atlas; Treg, regulatory T cell; UMAP, Uniform Manifold Approximation and Projection

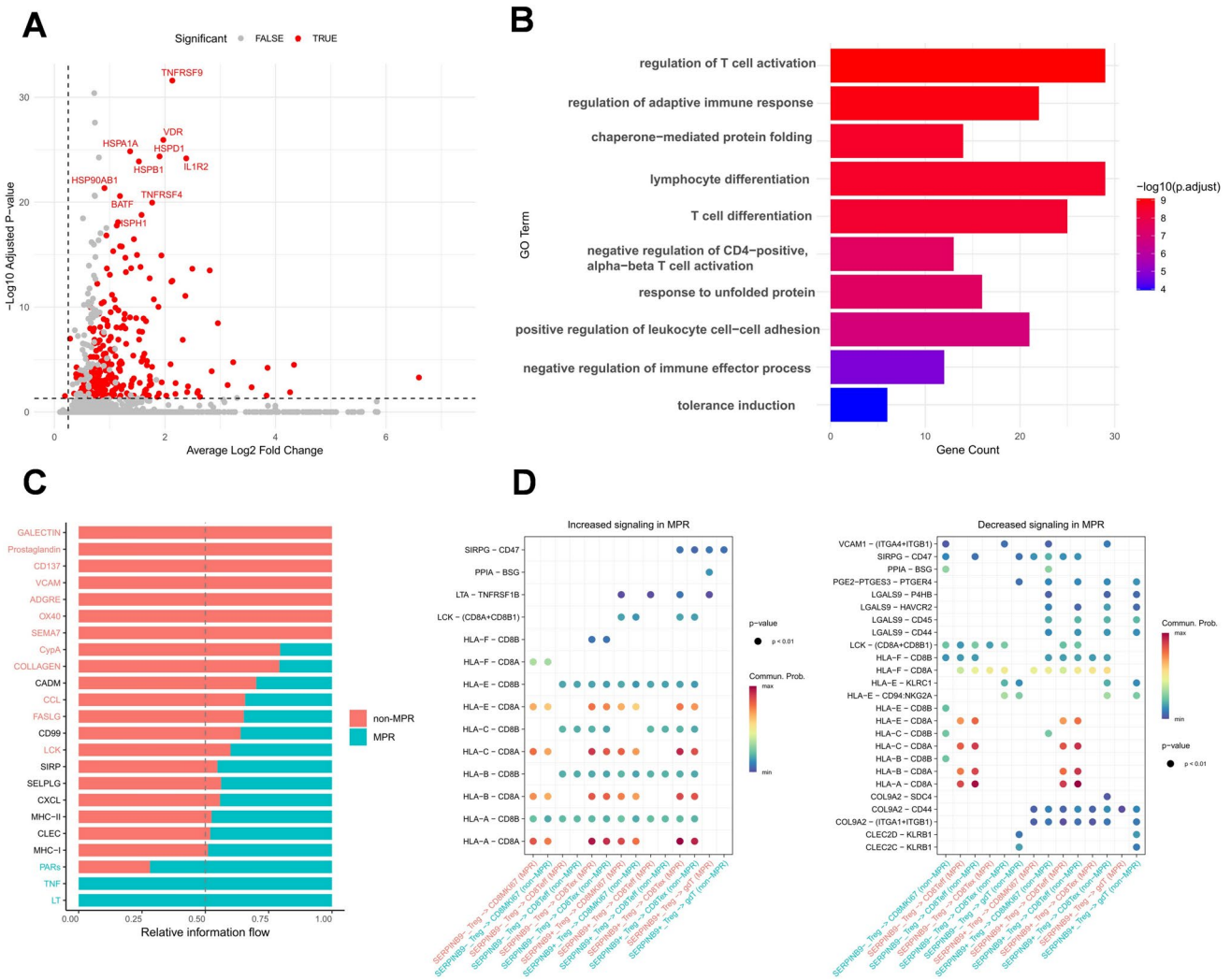


**Fig. 7** SERPINB9-Positive Regulatory T Cells Drive Immunosuppressive Communication Networks. **A** Heatmap of differentially expressed genes between SERPINB9-positive and SERPINB9-negative regulatory T cells. **B** Gene Ontology enrichment analysis of SERPINB9-positive regulatory T-cell signature. **C** Comparison of intercellular communication networks between MPR and non-MPR. **D** The heatmap compares signaling pathway activity between MPR and non-MPR across different T-cell subsets. Abbreviations: CD, cluster of differentiation; GO, Gene Ontology; HAVCR2, hepatitis A virus cellular receptor 2; LGALS9, galectin-9; MHC, major histocompatibility complex; MPR, major pathological response; non-MPR, non-major pathological response; OX40, tumor necrosis factor receptor superfamily member 4; PGE2, prostaglandin E2; TIM-3, T-cell immunoglobulin and mucin-domain containing protein 3; TNF, tumor necrosis factor; Treg, regulatory T cell

immunosuppressive state (Fig. 7C). Upregulation of T-cell inhibitory pathways, including the LGALS9-HAVCR2 (Galectin-9-TIM-3) checkpoint axis and the PGE2 signaling pathway. Remodeling of the tumor stroma toward a fibrotic state, evidenced by increased signaling from VCAM1 and COLLAGEN pathways (Fig. 7D). Dominance of CD137 (4-1BB) and OX40 pathways utilized to promote proliferation and functional stability of the suppressive SERPINB9 + Treg population rather than activating effector T cells.

## Discussion

This study developed a post-treatment CT-based habitat radiomics model for non-invasive prediction of TIL status following neoadjuvant immunochemotherapy in NSCLC patients. The Random Forest model achieved an AUC of 0.823 in the test set, demonstrating predictive performance. Through recurrence risk stratification, this model offered significant clinical translation value for precision evaluation of immunochemotherapy response.

The findings provide evidence that imaging-predicted TIL-negative tumors harbor elevated Treg infiltration, identified a SERPINB9 + Treg subset that accumulates in non-responding NSCLC tumors through differentiation from naive T cells. By bridging these scales, this study provides a non-invasive biomarker and illuminates biological mechanisms underlying treatment non-response.

Intratumoral heterogeneity critically influences treatment response. Previous studies have demonstrated that tumors can be classified into immune hot, altered, and cold phenotypes, with the immune-excluded architecture associated with poor therapeutic response [24]. TIL represents a component of anti-tumor immunity, with infiltration levels correlating with immunochemotherapy efficacy [25]. The findings demonstrated that NSCLC patients with high TIL infiltration exhibited treatment response and survival benefit following immunochemotherapy, consistent with previous investigations [7, 8]. Studies have demonstrated correlation between baseline NLR and immune checkpoint inhibitor treatment response [26]. Systematic reviews and meta-analyses have confirmed the prognostic value of NLR and PLR as biomarkers for advanced NSCLC immunochemotherapy [27]. Future studies combining local radiomics-predicted TIL status with systemic inflammatory ratios may improve prediction accuracy, providing valuable information for individualized treatment decisions.

Machine learning approaches for tumor microenvironment evaluation have been shown to stratify patients receiving neoadjuvant therapy [28]. Previous studies have established radiomics-based models for predicting CD8 + T-cell infiltration, proving the feasibility of non-invasive immune analysis [29]. Research has achieved results using conventional CT features to predict CD8 + T-cell infiltration in NSCLC [30]. AI models have been developed that predicted TIL density and analyzed spatial distribution patterns, finding that TIL spatial organization predicted progression-free survival in NSCLC patients receiving immunochemotherapy [8]. Artificial intelligence advances have enabled evaluation of multiple biomarkers. AI algorithms have been developed for assessment of PD-L1 expression and tumor-infiltrating lymphocytes, demonstrating automated biomarker evaluation feasibility [31]. Deep learning models validated in large cohorts have shown that AI-based scores predicted immunochemotherapy response with prognostic value [32]. Our imaging-based TIL prediction offers a non-invasive tool to assess the immune environment surrounding tumors.

Studies have developed habitat radiomics analysis for predicting progression-free survival and immune-related adverse reactions in NSCLC patients receiving immunochemotherapy, showing excellent performance in multicenter validation [33]. Through K-means clustering, we identified three distinct habitat sub-regions, enabling capture of spatial heterogeneity within tumors. Previous study validates this approach's value in immune microenvironment assessment through integration of radiomics features from habitat sub-regions [22]. Studies have shown that combining intratumoral and peritumoral PET/CT radiomics effectively predicted NSCLC immunochemotherapy response [34]. Recent studies have shown that quantifying intratumoral heterogeneity can predict prognosis, histological subtypes, and invasiveness in lung adenocarcinoma [35–37]. Integrating intratumoral heterogeneity metrics with immune microenvironment features may improve risk stratification and prognosis prediction in NSCLC patients receiving immunochemotherapy [38].

The transcriptional program of SERPINB9 + Tregs reveals therapeutic implications. The high expression of FOXP3 confirms their identity as regulatory T cells [39, 40]. This finding has implications for therapeutic strategies targeting these pathways, as work has shown that OX40/CD137 bispecific agonist therapy can reprogram Tregs [41]. Previous work has demonstrated that anti-SerpinB9 therapy can achieve tumor killing and enhance immunochemotherapy [42]. The biological functions of SerpinB9 extend beyond protease inhibition, encompassing roles in immune regulation [43]. The dominance of MHC-I and MHC-II pathways in responders, along with TNF and lymphotoxin signaling, supports the formation of lymphoid aggregates. These organized lymphoid aggregates have been shown to improve immunochemotherapy response and survival in multiple cancer types [44, 45].

The TIL assessment approach relied on semi-quantitative pathological evaluation, based on representative tissue sections rather than comprehensive whole-slide analysis, which may introduce some degree of inter-observer variation. The binary classification models for TIL status may not fully capture immune infiltration patterns. Studies have found that computerized tertiary lymphoid structures density served as a prognostic biomarker for resectable lung adenocarcinoma [46]. Prospective validation studies with standardized quantitative TIL assessment methods would strengthen the clinical applicability of these findings. Additionally, selection bias may exist as the cohort consisted of surgical candidates who generally had pathological responses and high TIL proportions. Multi-omics integration combining radiomics with circulating tumor DNA, PD-L1 expression, and tumor mutational burden could provide more comprehensive predictive models.

This study has limitations. First, the lack of an independent, multi-center external validation cohort represents a limitation. While the model demonstrated performance in internal testing and cross-validation, its generalizability to other institutions, scanners, and populations remains to be established. Data collection and collaboration with multiple centers are ongoing to address this gap. Second, the single-cell RNA-seq analysis was performed on

a publicly available dataset; the small sample size and imbalanced groups may limit the generalizability of cellular mechanism findings. Nonetheless, our orthogonal validation using FOXP3 in our own cohort corroborates the key finding of Treg enrichment in non-responders, lending credibility. Potential applications include contribution to surgical risk stratification, identification of patients who might benefit from additional therapy before surgery, and selection of candidates for adjuvant immunotherapy. While the model showed prognostic value for recurrence risk stratification, prospective studies are needed to determine whether radiomics-guided treatment decisions can improve clinical outcomes.

## Conclusions

In conclusion, this study developed a habitat radiomics model for non-invasive prediction of TIL status following neoadjuvant immunochemotherapy in NSCLC. The model demonstrated predictive and prognostic performance, with biological validation through single-cell transcriptomics revealing SERPINB9 + Treg-driven immunosuppressive mechanisms. By capturing tumor heterogeneity through habitat analysis and machine learning, this approach provides interpretable imaging biomarkers that could guide personalized immunochemotherapy strategies.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00262-025-04291-x.

**Acknowledgements** The computation is completed in the HPC Platform of Huazhong University of Science and Technology. The authors would like to express their sincere gratitude to the technical staff of the HPC Platform for their valuable assistance and support throughout this research.

**Author contributions** Zuhan Geng and Yihong Hu drafted the manuscript. Zuhan Geng, Guanchao Ye, and Yongde Liao conceived and designed the study. Yihong Hu and Peiyuan Mei performed the single-cell analysis. Shunchen Zhou conducted the pathological validation. Guangyao Wu carried out the image processing and data analysis. Zhenyu Gong and Zuhan Geng developed the computational models. Ruiping Feng collected and organized the clinical data. Zhangfeng Huang performed the histological staining experiments. Kuo Li, Guanchao Ye, and Yongde Liao provided resources and supervised the project. All authors reviewed, edited, and approved the final version of the manuscript.

**Funding** This research was supported by grants from the National Key Research and Development Program of China (No. 2023YFC2508603), the National Science Foundation for Young Scientists of China (Nos. 62406122, 82202148 and 82502479) and the China Postdoctoral Science Foundation (No. 2025M782272).

**Data availability** Data generated or analyzed during the study are available from the corresponding author (liaoyongde@hust.edu.cn) by request.

## Declarations

**Conflicts of interest** The authors declare no competing interests.

**Ethical approval** The study protocol was reviewed and approved by the Institutional Review Board of Union Hospital, Tongji Medical College, Huazhong University of Science and Technology (No. UCT240116). Given the retrospective design of this investigation, the ethics committee waived the requirement for individual participant consent.

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

## References

1. Forde PM, Spicer J, Lu S et al (2022) Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer. N Engl J Med 386(21):1973–1985
2. Wakelee H, Liberman M, Kato T et al (2023) Perioperative pembrolizumab for early-stage non-small-cell lung cancer. N Engl J Med 389(6):491–503
3. Provencio M, Serna-Blasco R, Nadal E et al (2022) Overall survival and biomarker analysis of neoadjuvant nivolumab plus chemotherapy in operable stage IIIA non-small-cell lung cancer (NADIM phase II trial). J Clin Oncol 40(25):2924–2933
4. Seymour L, Bogaerts J, Perrone A et al (2017) *i*RECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol 18(3):e143–e152
5. Chen DS, Mellman I (2017) Elements of cancer immunity and the cancer-immune set point. Nature 541(7637):321–330
6. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A (2015) Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. Clin Cancer Res 21(2):249–257
7. Gataa I, Mezquita L, Rossoni C et al (2021) Tumour-infiltrating lymphocyte density is associated with favourable outcome in patients with advanced non-small cell lung cancer treated with immunotherapy. Eur J Cancer 145:221–229
8. Park S, Ock CY, Kim H et al (2022) Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes as complementary biomarker for immune checkpoint inhibition in non-small-cell lung cancer. J Clin Oncol 40(17):1916–1928
9. Rakaee M, Adib E, Ricciuti B et al (2023) Association of machine learning-based assessment of tumor-infiltrating lymphocytes on standard histologic images with outcomes of immunotherapy in patients with NSCLC. JAMA Oncol 9(1):51–60
10. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278(2):563–577
11. Peng J, Zou D, Zhang X, Ma H, Han L, Yao B (2024) A novel sub-regional radiomics model to predict immunotherapy response in non-small cell lung carcinoma. J Transl Med 22(1):87

12. Trebeschi S, Drago SG, Birkbak NJ et al (2019) Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. Ann Oncol 30(6):998–1004
13. She Y, He B, Wang F et al (2022) Deep learning for predicting major pathological response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer: a multicentre study. EBioMedicine 86:104364
14. Gong XQ, Liu N, Tao YY et al (2023) Radiomics models based on multisequence MRI for predicting PD-1/PD-L1 expression in hepatocellular carcinoma. Sci Rep 13(1):7710
15. Wu J, Li C, Gensheimer M et al (2021) Radiological tumor classification across imaging modality and histology. Nat Mach Intell 3:787–798
16. Caii W, Wu X, Guo K, Chen Y, Shi Y, Chen J (2024) Integration of deep learning and habitat radiomics for predicting the response to immunotherapy in NSCLC patients. Cancer Immunol Immunother 73(8):153
17. Sujit SJ, Aminu M, Karpinets TV et al (2024) Enhancing NSCLC recurrence prediction with PET/CT habitat imaging, ctDNA, and integrative radiogenomics-blood insights. Nat Commun 15(1):3152
18. Travis WD, Dacic S, Wistuba I et al (2020) IASLC Multidisciplinary recommendations for pathologic assessment of lung cancer resection specimens after neoadjuvant therapy. J Thorac Oncol 15(5):709–740
19. Cottrell TR, Thompson ED, Forde PM et al (2018) Pathologic features of response to neoadjuvant anti-PD-1 in resected non-small-cell lung carcinoma: a proposal for quantitative immune-related pathologic response criteria (irPRC). Ann Oncol 29(8):1853–1860
20. Hendry S, Salgado R, Gevaert T et al (2017) Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in melanoma, gastrointestinal tract carcinomas, non-small cell lung carcinoma and mesothelioma, endometrial and ovarian carcinomas, squamous cell carcinoma of the head and neck, genitourinary carcinomas, and primary brain tumors. Adv Anat Pathol 24(6):311–335
21. Stein JE, Lipson EJ, Cottrell TR et al (2020) Pan-tumor pathologic scoring of response to PD-(L)1 blockade. Clin Cancer Res 26(3):545–551
22. Ye G, Wu G, Zhang C et al (2024) CT-based quantification of intratumoral heterogeneity for predicting pathologic complete response to neoadjuvant immunochemotherapy in non-small cell lung cancer. Front Immunol 15:1414954
23. Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 350:g7594
24. Galon J, Bruni D (2019) Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nat Rev Drug Discov 18(3):197–218
25. Tay C, Tanaka A, Sakaguchi S (2023) Tumor-infiltrating regulatory T cells as targets of cancer immunotherapy. Cancer Cell 41(3):450–465
26. Valero C, Lee M, Hoen D et al (2021) Pretreatment neutrophil-to-lymphocyte ratio and mutational burden as biomarkers of tumor response to immune checkpoint inhibitors. Nat Commun 12(1):729
27. Platini H, Ferdinand E, Kohar K et al (2022) Neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio as prognostic markers for advanced non-small-cell lung cancer treated with immunotherapy: a systematic review and meta-analysis. Medicina Kaunas 58(8):1069
28. Lin Y, Yu Y, Wang Q et al (2025) Machine learning model for predicting tertiary lymphoid structures and treatment response in triple-negative breast cancer. NPJ Precis Oncol 9(1):216
29. Sun R, Limkin EJ, Vakalopoulou M et al (2018) A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. Lancet Oncol 19(9):1180–1191
30. Chen Y, Xu T, Jiang C, You S, Cheng Z, Gong J (2023) CT-based radiomics signature to predict CD8+ tumor infiltrating lymphocytes in non-small-cell lung cancer. Acta Radiol 64(4):1390–1399
31. Molero A, Hernandez S, Alonso M et al (2025) Assessment of PD-L1 expression and tumour infiltrating lymphocytes in early-stage non-small cell lung carcinoma with artificial intelligence algorithms. J Clin Pathol 78(7):456–464
32. Rakaee M, Tafavvoghi M, Ricciuti B et al (2025) Deep learning model for predicting immunotherapy response in advanced non-small cell lung cancer. JAMA Oncol 11(2):109–118
33. Wu Y, Zhang W, Liang X et al (2025) Habitat radiomics analysis for progression free survival and immune-related adverse reaction prediction in non-small cell lung cancer treated by immunotherapy. J Transl Med 23(1):393
34. Lin X, Liu Z, Zhou K et al (2025) Intratumoral and peritumoral PET/CT-based radiomics for non-invasively and dynamically predicting immunotherapy response in NSCLC. Br J Cancer 132(6):558–568
35. Zuo Z, Zeng Y, Deng J et al (2025) Intratumoral heterogeneity score enhances invasiveness prediction in pulmonary ground-glass nodules via stacking ensemble machine learning. Insights Imaging 16(1):209
36. Zuo Z, Deng J, Ge W et al (2025) Quantifying intratumoral heterogeneity within sub-regions to predict high-grade patterns in clinical stage I solid lung adenocarcinoma. BMC Cancer 25(1):51
37. Zhang J, Sha J, Liu W, Zhou Y, Liu H, Zuo Z (2024) Quantification of intratumoral heterogeneity: distinguishing histological subtypes in clinical T1 stage lung adenocarcinoma presenting as pure ground-glass nodules on computed tomography. Acad Radiol 31(10):4244–4255
38. Zuo Z, Zhang G, Song P et al (2021) Survival Nomogram for Stage IB Non-Small-Cell Lung Cancer Patients, Based on the SEER Database and an External Validation Cohort. Ann Surg Oncol 28(7):3941–3950
39. Li C, Jiang P, Wei S, Xu X, Wang J (2020) Regulatory T cells in tumor microenvironment: new mechanisms, potential therapeutic strategies and future prospects. Mol Cancer 19(1):116
40. Plitas G, Konopacki C, Wu K et al (2016) Regulatory T cells exhibit distinct features in human breast cancer. Immunity 45(5):1122–1134
41. Imianowski CJ, Kuo P, Whiteside SK et al (2024) IFNγ production by functionally reprogrammed Tregs promotes antitumor efficacy of OX40/CD137 bispecific agonist therapy. Cancer Res Commun 4(8):2045–2057
42. Jiang L, Wang YJ, Zhao J et al (2020) Direct tumor killing and immunotherapy through anti-SerpinB9 therapy. Cell 183(5):1219-1233.e18
43. Huang H, Mu Y, Li S (2024) The biological function of Serpinb9 and Serpinb9-based therapy. Front Immunol 15:1422113
44. Cabrita R, Lauss M, Sanna A et al (2020) Tertiary lymphoid structures improve immunotherapy and survival in melanoma. Nature 577(7791):561–565
45. Helmink BA, Reddy SM, Gao J et al (2020) B cells and tertiary lymphoid structures promote immunotherapy response. Nature 577(7791):549–555
46. Wang Y, Lin H, Yao N et al (2023) Computerized tertiary lymphoid structures density on H&E-images is a prognostic biomarker in resectable lung adenocarcinoma. Science 26(9):107635

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

**Zuhan Geng[1] · Yihong Hu[2,3] · Shunchen Zhou[4] · Guangyao Wu[5] · Zhenyu Gong[6] · Ruiping Feng[7] · Zhangfeng Huang[1] · Peiyuan Mei[1] · Kuo Li[1] · Guanchao Ye[8] · Yongde Liao[1]**

✉ Kuo Li
lktj126@126.com

✉ Guanchao Ye
guanchaoye@zzu.edu.cn

✉ Yongde Liao
liaoyongde@hust.edu.cn

[1] Department of Thoracic Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

[2] College of Medicine, Chongqing University, Chongqing, China

[3] Radiation Oncology Centre, Chongqing University Cancer Hospital, Chongqing, China

[4] Department of Pathology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[5] Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[6] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China

[7] Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[8] Department of Thoracic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China