Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Interpretable Multi-modal Deep Learning Radiomics for Preoperative Mapping of Tertiary Lymphoid Structure Associated B-Cell Tumor Immune Microenvironment and Immunotherapy Response in Clear Cell Renal Cell Carcinoma

**Yingjie Xv**
First Affiliated Hospital of Chongqing Medical University
**Hongjian Liu**
First Affiliated Hospital of Chongqing Medical University
**Zongjie Wei**
Second People's Hospital of Yibin
**Bangxin Xiao**
First Affiliated Hospital of Chongqing Medical University
**Zhanpeng Yuan**
First Affiliated Hospital of Chongqing Medical University
**Qing Jiang**
Second Affiliated Hospital of Chongqing Medical University
**Xuan Zhang**
The Affiliated Yongchuan Hospital of Chongqing Medical University
**Feng Li**
Chongqing Three Gorges Central Hospital
**Yong Chen**
Fuling Center Hospital of Chongqing
**Ming Qiu**
The Affiliated Dazu's Hospital of Chongqing Medical University
**Xiang Peng**
First Affiliated Hospital of Chongqing Medical University
**Jiaxin Su**
Children's Hospital of Chongqing Medical University
**Weiyang He**
First Affiliated Hospital of Chongqing Medical University

**Mingzhao Xiao**

mingzhaoxiao@cqmu.edu.cn

First Affiliated Hospital of Chongqing Medical University

**Article**

**Keywords:** Clear cell renal cell carcinoma, Tertiary lymphoid structures, Deep learning radiomics, Biopathological interpretability

**Posted Date:** July 28th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10338322/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Additional Declarations:** No competing interests reported.

## Abstract

# Background

Tertiary lymphoid structures (TLSs) are associated with prognosis and immunotherapy response in clear cell renal cell carcinoma (ccRCC), but their assessment requires invasive tissue sampling. This study aimed to develop and validate a CT-based multi-modal artificial intelligence model for preoperative TLS prediction in ccRCC and to uncover the pathological and molecular determinants of the model's predictions.

# Methods

This multicenter, retrospective-prospective study enrolled 1,361 patients with ccRCC from eight tertiary hospitals, assigned to primary, test 1, test 2, prospective, multi-omics and immune checkpoint inhibitor (ICI) therapy sets. Arterial-phase CT images were used to train 2D, 2.5D, and 3D deep learning (DL) networks, alongside machine learning (ML) radiomics models. The best single-modality models were integrated via pre-fusion (PFM) and late-fusion (LFM) strategies. Performance was evaluated via the area under the receiver operating characteristic (ROC) curve (AUC). Prognostic value was analyzed via Kaplan-Meier curves. Interpretability was investigated using SHaply Additive Explanations (SHAP) algorithm, HoVer-Net cell segmentation, and bulk RNA sequencing analysis.

# Results

Among the 22 single-modality models, 2.5D-DenseNet161, 2D-ResNet50, Radiomics-XGBoost, and 3D-ResNet101 performed best in test set 1 (AUCs: 0.899, 0.888, 0.878, and 0.756, respectively), which were adopted for fusion models' establishment. The SVM-based LFM surpassed the PFM (AUC: 0.938 vs. 0.919) and maintained stability across the test 2, prospective, and multi-omics sets (AUC range, 0.877–0.918). LFM-predicted TLS positive significantly correlated with prolonged progression-free survival (all $P < 0.05$) and effectively discriminated clinical responders from non-responders for ICI therapy (AUC: 0.807). Model predicted TLS positive showed higher inflammatory cell densities ($P < 0.05$), upregulation of the CXCL13-CXCR5 axis and B-cell-related markers, enrichment of humoral immune response pathways and increased naive and memory B-cell infiltration ($P < 0.05$).

# Conclusions

The multi-modal LFM provides a non-invasive surrogate for TLS assessment in ccRCC and may support preoperative risk stratification and immunotherapy decision-making.

## Introduction

Renal cell carcinoma (RCC) represents a major urological malignancy, with clear cell renal cell carcinoma (ccRCC) accounting for approximately 70–80% of all cases[1]. With the widespread adoption of medical imaging screening and health check-ups, an increasing number of ccRCC cases are being detected[2]. Current localized treatments primarily involve surgical resection, while locally advanced or metastatic cases rely on immune checkpoint inhibitors (ICIs) combined with anti-angiogenic therapies[3,4]. However, clinical responses to these treatments vary significantly among patients.

Traditional prognostic indicators, such as TNM staging, WHO/ISUP nuclear grading and sarcomatous differentiation inadequately capture the complexity of the tumor microenvironment[5–7]. While emerging biomarkers including PD-L1 expression, tumor mutational burden (TMB), and immune cell infiltration rate offer deeper insights, their clinical application is constrained by high costs and absolute dependence on tissue sampling[8–10]. Consequently, there is an ongoing need for robust and non-invasive prognostic markers.

Tertiary lymphoid structures (TLS) are ectopic lymphoid aggregates that form within chronic inflammatory or neoplastic microenvironments[11,12]. Comprising distinct B-cell and T-cell zones, mature dendritic cells, and germinal centers, TLS orchestrate local anti-tumor immunity through enhanced antigen presentation and lymphocyte activation[13,14]. In ccRCC, the presence and maturation of TLS strongly correlate with a reduced risk of recurrence and improved ICI responsiveness[15,16]. Despite their prognostic value, assessing TLS currently dictates histopathological evaluation of surgical or biopsy specimens. This invasive approach is intrinsically flawed by spatial heterogeneity, procedural risks such as hemorrhage, and subjective interpretation. Therefore, establishing a preoperative, non-invasive surrogate for TLS status is of considerable clinical utility.

Contrast-enhanced computed tomography (CECT) is the standard imaging modality for the diagnosis of ccRCC. Beyond gross anatomical data, CECT captures subvisual textural features associated with tumor angiogenesis, necrosis, cellular density, and micro-environment composition. While CT-based radiomics and deep learning (DL) have demonstrated utility in predicting TLS status across multiple malignancies (including liver tumor, breast tumor, and lung adenocarcinoma), this application remains unexplored in ccRCC[17–20]. In addition, existing studies predominantly rely on single-modality models, lacking a systematic comparison and fusion across 2D, 2.5D, and 3D DL architectures or traditional machine learning (ML) radiomics. Critically, prior models frequently lack rigorous validation in prospective and immunotherapy cohorts, as well as biological interpretability at the transcriptomic or pathological levels.

This study aimed to develop a multi-modal framework for preoperative TLS assessment in ccRCC by benchmarking various deep learning (2D, 2.5D, and 3D) and radiomic models, whose prognostic and immunotherapeutic utility was validated across external and prospective cohorts. We further integrated digital pathology and transcriptomics to elucidate the biopathological underpinnings that support the model's prediction reliability. Our study may provide a reliable non-invasive tool for TLS assessment and support personalized clinical decision-making in ccRCC.

## Methods and Materials

# Study Design

This multicenter, retrospective and prospective cohort study was conducted in accordance with the Declaration of Helsinki[21]. The study protocol was approved by the Ethics Committee of the First Affiliated Hospital of Chongqing Medical University (Approval No. ZZ2026-0271-01). The requirement for informed consent was waived by our institutional ethics committee for the retrospective cohort, and all patients in the prospective cohort provided written informed consent.

The overall cohort comprised 1,361 patients from 8 tertiary hospitals in China, divided into a primary set, two external test sets, a prospective validation set, an immune checkpoint inhibitor (ICI) therapy cohort, and a multi-omics cohort. The research workflow encompassed three phases (Fig. 1): In Phase 1, deep learning (DL) and traditional radiomics models were developed utilizing 2D, 2.5D, and 3D tumor regions; the optimal models in each category were identified using test set 1. In Phase 2, early (feature-level) and late (decision-level) fusion strategies were applied to integrate data modalities and establish a unified fusion model. In Phase 3, the best-performance fusion model was evaluated across the multicenter external, prospective, ICI-treated, and multi-omics cohorts. Finally, the pathological and transcriptomic underpinnings of the model's predictions were investigated via digital pathology and bulk RNA sequencing.

## Patient cohort

Patients who underwent radical or partial nephrectomies (RN/PN) and were pathologically diagnosed with ccRCC at 8 tertiary hospitals (Center 1–8; Appendix file 1) between January 2016 and June 2025 were retrospectively included in this study and formed the TLS cohort (n = 1077).

Patients from Center 1 (n = 417), 2 (n = 165), 3 (n = 83), and 4 (n = 21) constituted the primary set (n = 686). Patients from Center 5 (n = 176) and 6 (n = 68) formed test set 1 (n = 244), while those from Center 7 (n = 79) and 8 (n = 68) formed test set 2 (n = 147). For the prospective cohort, patients treated at Center 1 (n = 59) and 8 (n = 8) between October 2025 and April 2026 were enrolled (n = 67). The ICI cohort (n = 152) comprised patients from Center 1 (n = 104), 2 (n = 37), and 8 (n = 11) treated between January 2023 and December 2025. The multi-omics cohort included 65 patients treated at Center 1 between January and September 2025. Detailed inclusion and exclusion criteria are outlined in Fig. 2.

## Clinico-pathological characteristics and clinical outcome

Clinico-pathological data including age, sex, tumor laterality, surgical procedure, tumor size, WHO/ISUP nuclear grade[22], intratumoral necrosis, and pathological TNM staging[23] were collected from electronic medical records.

TLS investigated in this study was defined as an intratumoral lymphoid aggregate exhibiting distinct follicular structures[24]. TLS status was independently evaluated by two uropathologists with over 10 years of experience. Based on established criteria, tumors containing at least one mature or immature TLS were classified as TLS-positive; all others were deemed TLS-negative. Discrepancies were resolved by a third senior pathologist (> 25 years of experience).

The primary clinical endpoint was progression-free survival (PFS), defined as the interval from surgery or initiation of systemic therapy to the first radiologically or pathologically confirmed disease progression, recurrence, death, or last follow-up[25]. In the ICI cohort, treatment response was evaluated using RECIST 1.1 criteria[26]. Patients achieving a complete or partial response (CR/PR) were classified as responders, whereas those with stable or progressive disease (SD/PD) were categorized as non-responders.

## CT acquisitions and image pre-procession

Preoperative abdominal CT scans were performed within two weeks prior to surgery or biopsy. Images were exported in DICOM format from institutional picture archiving and communication system (PACS). Scanning parameters are detailed in Table S1. Arterial phase images were selected for analysis and converted to NIfTI format. To minimize scanner-induced variations, intensity normalization was applied using a standard abdominal soft-tissue window (width: 300 HU, level: -30 HU). Images were subsequently resampled to 1×1×1 $mm^3$ isotropic voxels via cubic spline interpolation.

Tumor regions were automatically segmented utilizing the 3D nnU-Net framework. The resulting masks were manually reviewed and corrected using ITK-SNAP (Version 3.8.0) by a radiologist with 19 years of abdominal imaging experience (Reader A). Using the 3D tumor regions of interest (ROIs), 1,315 radiomic features encompassing first-order, shape, and texture features were extracted via PyRadiomics. Extracted features underwent z-score standardization followed by a stepwise dimensionality reduction process: (1) inter- and intra-observer intraclass correlation coefficients (threshold ≥ 0.80), (2) Mann-Whitney U testing, (3) minimum redundancy maximum relevance (mRMR), and (4) Elastic Net regression (Appendix file 2; Figs. S1-S3).

## Model Construction and Fusion Strategy

Three kinds of DL models (2D, 2.5D, and 3D) and traditional radiomics models were developed. The 2D models utilized axial slices containing the maximum tumor cross-section, resized to 256×256 pixels. The 2.5D models incorporated three orthogonal views (axial, sagittal, coronal), each resized to 256×256 pixels. The 3D models analyzed a cubic volume centered on the tumor mask with a 5-mm outward extension. The 2D and 2.5D DL models employed six architectures (Densnet161, Densnet201, ResNet101, ResNet50, vgg16_bn, vgg19) initialized with ImageNet weights, while the 3D DL model evaluated five architectures (ResNet101, ResNet50, ResNet34, ResNet18, ResNet10) pre-trained on Med3D.The associated codes were upload in a public-available website

(https://github.com/GaogeXv/CT_DL_ccRCC_tls2026). All DL models were trained using stochastic gradient descent for 50 epochs with an initial learning rate of 0.01. For the radiomics model, dimensionally reduced features were fed into five machine learning (ML) algorithms (XGBoost, RF, SVM, LR, and KNN). All models were trained and hyperparameter-tuned on the primary set. The model demonstrating the highest predictive efficacy on test set 1 within each category was selected for downstream fusion analysis.

Pre-fusion strategy: Features extracted from the final average pooling layers of the optimal 2D, 2.5D, and 3D DL models were concatenated with the standardized radiomic features. Following dimensionality reduction (Appendix file 3; Figs. S4–S6), these integrated features were processed through the five ML algorithms to identify the optimal pre-fusion model (PFM).

Late-fusion strategy: Risk scores generated by the optimal 2D DL, 2.5D DL, 3D DL, and radiomics models were stacked and evaluated using the same five ML algorithms to determine the best late fusion model (LFM).

# Model performances

Predictive performance was assessed using receiver operating characteristic (ROC) curves and the area under the curve (AUC). Accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive (NPV) values were calculated. Calibration curves were generated to evaluate the agreement between predicted probabilities and actual outcomes. Decision curve analysis (DCA) was applied to quantify potential clinical net benefit. To ensure interpretability, SHapley Additive exPlanations (SHAP) were used to determine the contribution and directionality of individual input variables.

## WSI collection and pathological feature extraction

Hematoxylin and eosin (H&E)-stained sections were digitized using a high-resolution slide scanner at 0.25 μm/pixel to generate whole-slide images (WSIs). Image preprocessing involved Vahadane stain normalization, background removal, blank area exclusion, and tissue region extraction via the Otsu method. WSIs were partitioned into 10,000×10,000-pixel blocks, which were further subdivided into 2,048×2,048-pixel tiles. These tiles were subsequently input as 270×270-pixel patches into a pre-trained HoVer-Net model (hovernet_original_consep_type_tf2pytorch) for nuclear segmentation and classification. For each detected nucleus, the bounding box, center of mass, contour coordinates, cell type ID, and prediction probability were recorded. The count and spatial density of each cell type were quantified using QuPath (Version 0.5.1).

## Transcriptomic data analysis

Total RNA was extracted using the TRIzol method. Following quality control, cDNA libraries were constructed using the Illumina TruSeq RNA Library Prep Kit and sequenced on an Illumina NovaSeq 6000

platform (PE150), yielding 50–100 million reads per sample. Differential expression analysis was performed using “edgeR” to compare LFM-predicted TLS-positive versus TLS-negative samples. The screening thresholds were defined as |log2FC| > 1 and a false discovery rate (FDR) < 0.05. Differentially expressed genes (DEGs) were subsequently subjected to Gene Ontology (GO) enrichment analyses. Immune cell deconvolution was conducted utilizing CIBERSORTx with the LM22 reference matrix to assess variations in immune cell infiltration between the LFM-predicted groups[27].

# Statistical analysis

Continuous variables were reported as median (interquartile range). Categorical variables were presented as frequencies and percentages. Group comparisons were performed using the Mann-Whitney U test. Survival outcomes were evaluated using the Kaplan-Meier method and Cox proportional hazards regression. All statistical analyses were conducted using R (version 4.2.0) and SPSS (version 26.0). A two-tailed $P < 0.05$ was considered statistically significant.

## Results

# Patient cohorts and clinico-pathological features

A total of 1,361 patients with ccRCC from eight centers were included in this study. Of these, 1,077 comprised the TLS cohort, 67 formed the prospective set (Table S2), 152 were allocated to the ICI efficacy prediction set (Table S3), and 65 were included in the multi-omics set (Table S4). Within the TLS cohort, 397 (36.9%) patients were pathologically confirmed as TLS-positive, the mean age was 58.00 (50.00–68.00) years, and the cohort included 682 (63.3%) males. Per the study protocol, the TLS cohort was divided into a primary set (n = 686), test set 1 (n = 244), and test set 2 (n = 147). No statistically significant differences in clinico-pathological characteristics were observed among the primary set, test set 1, and test set 2 (Table 1; all $P > 0.05$).

Table 1
Clinic-pathological characteristics of patients in the primary, test 1, and test 2 sets

| Variables | All (n = 1077) | Primary (n = 686) | Test 1 (n = 244) | Test 2 (n = 147) | *P* value |
|---|---|---|---|---|---|
| **Age, years** | 58.00 (50.00–68.00) | 59.00 (51.00–68.00) | 58.00 (49.00–67.00) | 56.00 (49.00–67.00) | 0.146 |
| **Tumor size, cm** | 4.20 (3.00–6.00) | 4.30 (3.00–6.00) | 4.00 (3.00–6.00) | 4.00 (3.00-5.65) | 0.273 |
| **Sex, n (%)** | | | | | 0.896 |
| Female | 395 (36.7) | 251 (36.6) | 92 (37.7) | 52 (35.4) | |
| Male | 682 (63.3) | 435 (63.4) | 152 (62.3) | 95 (64.6) | |
| **Tumor laterality, n (%)** | | | | | 0.927 |
| Left | 540 (50.1) | 342 (49.9) | 125 (51.2) | 73 (49.7) | |
| Right | 537 (49.9) | 344 (50.1) | 119 (48.8) | 74 (50.3) | |
| **Surgical procedure, n (%)** | | | | | 0.075 |
| PN | 574 (53.3) | 348 (50.7) | 143 (58.6) | 83 (56.5) | |
| RN | 503 (46.7) | 338 (49.3) | 101 (41.4) | 64 (43.5) | |
| **WHO/IUSP grade, n (%)** | | | | | |
| I | 286 (26.6) | 179 (26.1) | 66 (27.0) | 41 (27.9) | |
| II | 492 (45.7) | 303 (44.2) | 118 (48.4) | 71 (48.3) | |
| III | 264 (24.5) | 176 (25.7) | 55 (22.5) | 33 (22.4) | |
| IV | 35 (3.2) | 28 (4.1) | 5 (2.0) | 2 (1.4) | |
| **Presence of necrosis, n (%)** | | | | | 0.882 |
| No | 813 (75.5) | 521 (75.9) | 183 (75.0) | 109 (74.1) | |
| Yes | 264 (24.5) | 165 (24.1) | 61 (25.0) | 38 (25.9) | |
| **pT stage, n (%)** | | | | | 0.304 |
| 1 | 846 (78.8) | 527 (76.8) | 197 (80.7) | 122 (83.0) | |
| 2 | 103 (9.6) | 69 (10.1) | 20 (8.2) | 14 (9.5) | |
| 3 | 128 (11.6) | 90 (13.1) | 27 (11.1) | 11 (7.5) | |

| Variables | All (n = 1077) | Primary (n = 686) | Test 1 (n = 244) | Test 2 (n = 147) | *P* value |
|---|---|---|---|---|---|
| **pN status, n (%)** | | | | | 0.249 |
| 0 | 1049 (97.4) | 664 (96.8) | 240 (98.4) | 145 (98.6) | |
| 1 | 28 (2.6) | 22 (3.2) | 4 (1.6) | 2 (1.4) | |
| **Presence of TLS, n (%)** | | | | | 0.608 |
| No | 680 (63.1) | 436 (63.6) | 148 (60.7) | 96 (65.3) | |
| Yes | 397 (36.9) | 250 (36.4) | 96 (39.3) | 51 (34.7) | |

# Predictive performances of single-modality models

Among the 2D DL models, ResNet50 demonstrated the highest predictive capability on test set 1, yielding an AUC of 0.888 (95% confidence interval [CI]: 0.834–0.946) and an accuracy of 0.845 (Fig. 3A). For the 2.5D models, DenseNet161 achieved an AUC of 0.899 (95% CI: 0.858–0.939) and an accuracy of 0.862 (Fig. 3B). Among the 3D models, ResNet101 yielded an AUC of 0.756 (95% CI: 0.719–0.817) and an accuracy of 0.739 (Fig. 3C). For traditional radiomics, the XGBoost model achieved an AUC of 0.878 (95% CI: 0.839-0.900) and an accuracy of 0.808 (Fig. 3D). Detailed performance metrics for all models are summarized in Fig. 3E and Tables S5-S8. Model-derived risk scores were positively correlated with TLS presence and favorable clinical outcomes (Fig. 3F).

# Performance of fusion models in external and prospective validations

For early fusion, features extracted from the final average pooling layers of the optimized 2D-ResNet50, 2.5D-DenseNet161, and 3D-ResNet101 models were concatenated with radiomics features. Following dimensionality reduction, 20 feature clusters significantly associated with TLS status were retained (Figs. S4-S6). When evaluated across five ML algorithms, the RF classifier demonstrated the highest performance (AUC = 0.919; 95% CI: 0.886–0.951) on test set 1 and was established as the final PFM (Fig. 4A; Table 2). Risk scores generated by the optimal 2D, 2.5D, 3D, and radiomics models were aggregated using a stacked ensemble approach for late fusion. The SVM-based late fusion model (LFM) yielded the highest predictive value on test set 1, with an AUC of 0.938 (95% CI: 0.911–0.961) (Fig. 4B; Table 2). Calibration and decision curve analyses confirmed that the RF-based PFM and SVM-based LFM provided the highest calibration accuracy and maximum net clinical benefit among the early and late fusion strategies, respectively (Fig. 4C-F, S7-S10).

Table 2
Predictive performances of the pre-fusion and late-fusion models in the primary and test 1 sets

| | Model | Dataset | AUC | 95% CI |
|---|---|---|---|---|
| **Pre-fusion** | RF | Primary set | 0.937 | 0.899–0.947 |
| | SVM | Primary set | 0.923 | 0.895–0.946 |
| | XGBoost | Primary set | 0.914 | 0.880–0.921 |
| | LR | Primary set | 0.921 | 0.891–0.959 |
| | KNN | Primary set | 0.915 | 0.866–0.953 |
| | RF | Test set 1 | 0.919 | 0.886–0.951 |
| | SVM | Test set 1 | 0.91 | 0.883–0.942 |
| | XGBoost | Test set 1 | 0.892 | 0.874–0.929 |
| | LR | Test set 1 | 0.886 | 0.855–0.923 |
| | KNN | Test set 1 | 0.898 | 0.871–0.940 |
| **Late-fusion** | SVM | Primary set | 0.956 | 0.921–0.977 |
| | LR | Primary set | 0.941 | 0.912–0.965 |
| | XGBoost | Primary set | 0.933 | 0.900-0.954 |
| | RF | Primary set | 0.912 | 0.887–0.946 |
| | KNN | Primary set | 0.925 | 0.893–0.950 |
| | SVM | Test set 1 | 0.938 | 0.911–0.961 |
| | LR | Test set 1 | 0.927 | 0.901–0.949 |
| | XGBoost | Test set 1 | 0.899 | 0.863–0.923 |
| | RF | Test set 1 | 0.887 | 0.861–0.911 |
| | KNN | Test set 1 | 0.896 | 0.869–0.928 |

Given its superior performance, the LFM was evaluated on test set 2, the prospective set, and the multi-omics set, yielding AUCs of 0.918 (95% CI: 0.867–0.960), 0.881 (95% CI: 0.791–0.957), and 0.877 (95% CI: 0.780–0.960), respectively (Fig. 4G, H). Patients experiencing disease progression or mortality exhibited significantly lower LFM risk scores compared to those with stable disease or no progression across test set 2, the prospective set, and the multi-omics set (all $P < 0.05$; Fig. 4I-K).

# Prognostic value and ICI therapy response prediction abilities of LFM

Kaplan-Meier survival analysis indicated that patients classified as TLS-positive by the LFM experienced significantly prolonged PFS compared to LFM-predicted TLS-negative patients across the primary set, test set 1, and test set 2 (all $P < 0.05$; Fig. 5A-C). In the prospective and multi-omics sets, no statistically significant differences in PFS were observed between the LFM-predicted TLS-positive and TLS-negative groups ($P = 0.059$ and $P = 0.441$, respectively; Fig. 5D, E).

As shown in Fig. 5G-I, the LFM predicted ICI treatment response in advanced ccRCC patients with an AUC of 0.807 (95% CI: 0.729–0.876). A significant difference in PFS was observed between LFM-predicted responders and non-responders ($P < 0.05$; Fig. 5F). Heatmaps of clinico-pathological features and waterfall plots illustrating LFM-predicted ICI responses are provided in Fig. 5J and 5K.

# Pathological and biological interpretability of the LFM

The biological basis of the LFM predictions was evaluated across model-specific, pathological, and transcriptomic dimensions. SHAP analysis indicated that the 2.5D-DenseNet161 features contributed most significantly to the final prediction, followed by the Radiomics-XGBoost features. Except for the 2D-ResNet50 inputs, all component risk scores were positively correlated with an LFM-predicted TLS-positive status (Fig. S11).

At the pathological level, HoVer-Net-annotated WSIs and cellular classifications demonstrated that LFM-predicted TLS-positive patients exhibited significantly higher densities of inflammatory and necrotic cells compared to TLS-negative patients (all $P < 0.05$; Fig. 6A, B). These findings were corroborated by transcriptomic profiling: differential expression analysis identified 138 significantly DEGs. Classical TLS and B-cell-associated genes, including CXCL13, CXCR5, MS4A1, and PDCD1, were significantly upregulated in TLS-positive patients (Fig. 6C). Functional enrichment analysis revealed that these DEGs were heavily enriched in humoral immune response pathways (Fig. 6D). Furthermore, transcriptomic deconvolution demonstrated significantly elevated proportions of naive and memory B cells in LFM-predicted TLS-positive patients ($P < 0.05$; Fig. 6E and 6F). Representative cases illustrating high concordance between LFM-predicted TLS status and corresponding biopathological ground truth are presented in Fig. 6G and 6H.

# Discussion

In this study, we developed and validated a multimodal framework combining CT-based DL and traditional radiomics to non-invasively assess TLS status in ccRCC. The established LFM achieved robust predictive accuracy across multiple external datasets and a prospective cohort. Beyond TLS detection, the LFM successfully stratified patient prognosis and predicted therapeutic benefits from ICIs. Furthermore, matched pathological and transcriptomic analyses corroborated the biological underpinnings of the LFM, providing a solid mechanistic basis for this AI-driven approach.

Former studies have demonstrated that the presence of TLS is associated with prolonged PFS and overall survival in various solid tumors, including ccRCC, as well as improved responses to

immunotherapy[28–31]. However, current TLS assessment relies entirely on histopathological examination of invasive surgical or biopsy specimens. Accurately identifying TLS status prior to surgery or systemic therapy remains a clinical challenge. Characterized by their capacity to mine high-dimensional data, DL and radiomics can decode intratumoral heterogeneity that remains imperceptible to the human eye, thereby offering a promising non-invasive paradigm for assessing TLS status prior to therapeutic intervention[32,33]. Pioneering studies have reported excellent predictive accuracies of radiomics and DL models in assessing tumor TLS status. For example, Long[20] et al. established an MRI radiomics model for peritumoral TLS in hepatocellular carcinoma (AUC = 0.91), while Lin[17] et al. successfully predicted both TLS presence and NAT efficacy in triple-negative breast cancer (AUC = 0.922). Extending beyond these paradigms, comparable non-invasive models have shown consistent reliability across diverse solid tumors, including gallbladder cancer, soft tissue sarcomas, and lung adenocarcinomas, typically achieving AUCs between 0.747 and 0.891[34–37].

Consistent with these prior investigations, our results confirm that macroscopic CT images contain distinct quantitative features that reliably reflect the microscopic formation of ccRCCs' TLS, with models' prediction AUCs ranged from 0.886 to 0.938. However, current studies frequently relied on relatively small, single-center cohorts, which restricts model generalizability. Furthermore, most existing models utilized a single approach (such as traditional radiomics alone or 2D DL), without systematically integrating multi-dimensional spatial data. Crucially, they rarely provided biological interpretability at the pathological or transcriptomic levels. To address these limitations, the present study established a large-scale cohort of 1,361 patients across eight medical centers. A distinct feature of our study is the multi-modal fusion strategy, which combined 2D, 2.5D, and 3D DL architectures with conventional radiomics. The constructed LFM achieved an AUC of 0.938 on test set 1 and maintained high predictive accuracy across independent datasets, yielding AUCs of 0.918 in test set 2, 0.881 in the prospective cohort, and 0.877 in the multi-omics cohort.

To elucidate the biological basis of the LFM predictions, this study utilized digital pathology and RNA sequencing. Pathological analysis indicated that LFM-predicted TLS-positive tumors exhibited higher densities of inflammatory cells and cell death. Transcriptomic analysis revealed an upregulation of classical B-cell-associated genes, including CXCL13, CXCR5, MS4A1, and PDCD1 in the TLS-positive group. The CXCL13-CXCR5 axis is known to mediate B-cell recruitment during TLS neogenesis[38–40]. Furthermore, immune deconvolution showed higher proportions of infiltrating naive and memory B cells in TLS-positive patients. These observations suggest that the LFM effectively captures imaging phenotypes related to local immune-inflammatory activity, rather than merely measuring gross anatomical differences.

Clinically, the LFM demonstrated utility in guiding prognostic evaluation. Patients predicted as TLS-positive by the LFM experienced longer PFS in the primary set, test set 1, and test set 2. In the advanced ccRCC cohort, the LFM successfully distinguished between responders and non-responders to ICI combined with TKI therapy (AUC: 0.807). This suggests that the model could assist clinicians in

identifying patients who are likely to benefit from immunotherapy before surgical intervention. In the prospective and multi-omics cohorts, survival differences between predicted TLS-positive and negative patients did not reach statistical significance. This result is likely attributable to the small sample sizes and relatively short follow-up periods in these specific subsets.

This study has several limitations. First, the retrospective design of the primary cohorts introduces potential selection bias. Second, all participating centers are located in China, and external validation in diverse populations is required to confirm broader generalizability. Third, the pathological assessment of TLS was based on H&E staining. Although evaluated by experienced pathologists, this method is subject to observer variability. Future studies utilizing multiplex immunofluorescence may improve precise TLS subtyping. Finally, the sample size for transcriptomic analysis was relatively small. The proposed biological mechanisms are preliminary and require further validation through in vivo and in vitro functional experiments.

## Conclusions

In summary, this study constructed a multi-modal DL radiomics model based on preoperative contrast-enhanced CT to predict TLS status in ccRCC. Validated across retrospective and prospective cohorts, the model demonstrated stable predictive accuracy and showed potential value for prognostic stratification and ICI response prediction. Integrated pathological and transcriptomic analyses further support its biological relevance and suggest that this non-invasive tool may help guide clinical decision-making and personalized management in ccRCC.

## Abbreviations

| Abbreviation | Full term |
|---|---|
| 2D | Two-dimensional |
| 2.5D | Two-and-a-half-dimensional |
| 3D | Three-dimensional |
| AI | Artificial intelligence |
| AUC | Area under the receiver operating characteristic curve |
| ccRCC | Clear cell renal cell carcinoma |
| cDNA | Complementary DNA |
| CECT | Contrast-enhanced computed tomography |
| CI | Confidence interval |
| CR | Complete response |
| CT | Computed tomography |
| DCA | Decision curve analysis |
| DEGs | Differentially expressed genes |
| DL | Deep learning |
| DICOM | Digital Imaging and Communications in Medicine |
| FDR | False discovery rate |
| GO | Gene Ontology |
| H&E | Hematoxylin and eosin |
| ICI | Immune checkpoint inhibitor |
| KNN | k-nearest neighbors |
| LFM | Late-fusion model |
| LR | Logistic regression |
| ML | Machine learning |
| mRMR | Minimum redundancy maximum relevance |
| MRI | Magnetic resonance imaging |
| NAT | Neoadjuvant therapy |
| NIfTI | Neuroimaging Informatics Technology Initiative |
| NPV | Negative predictive value |

| Abbreviation | Full term |
|---|---|
| PACS | Picture archiving and communication system |
| PD | Progressive disease |
| PD-L1 | Programmed death-ligand 1 |
| PFM | Pre-fusion model |
| PFS | Progression-free survival |
| PN | Partial nephrectomy |
| PPV | Positive predictive value |
| PR | Partial response |
| RCC | Renal cell carcinoma |
| RECIST | Response Evaluation Criteria in Solid Tumors |
| RF | Random forest |
| RN | Radical nephrectomy |
| ROC | Receiver operating characteristic |
| ROI | Region of interest |
| ROIs | Regions of interest |
| SD | Stable disease |
| SHAP | SHapley Additive exPlanations |
| SPSS | Statistical Package for the Social Sciences |
| SVM | Support vector machine |
| TKI | Tyrosine kinase inhibitor |
| TLS | Tertiary lymphoid structure |
| TMB | Tumor mutational burden |
| TNM | Tumor-node-metastasis |
| WHO/ISUP | World Health Organization/International Society of Urological Pathology |
| WSI | Whole-slide image |
| XGBoost | Extreme Gradient Boosting |

# Declarations

**Authors' contributions:** YX: Conceptualization, methodology, software, validation, formal analysis, investigation, writing-original draft, writing-review and editing, visualization. HL: Methodology, software, validation, writing-review and editing. ZW: Methodology, validation, formal analysis, writing - review and editing. BX: Software, validation, writing-review and editing. ZY: Methodology, validation, writing-review and editing. QJ: Data curation, validation, writing-review and editing. XZ: Data curation, validation, writing-review and editing. FL: Data curation, writing-review and editing. YC: Data curation, writing-review and editing. MQ: Data curation, writing-review and editing. XP: Methodology, validation, writing-review and editing. JS: Methodology, software, validation, formal analysis, writing-review and editing. WH: Conceptualization, resources, writing-review and editing, supervision. MX: Conceptualization, resources, writing-review and editing, supervision.

**Acknowledgments:** The authors sincerely thank the staff of the Departments of Urology, Pathology, and Radiology at the First Affiliated Hospital of Chongqing Medical University for their support and assistance throughout this study. This study was funded by the Sichuan Science and Technology Program (Grant No. 2026NSFSC1911), and the Health Commission of Sichuan Province Medical Science and Technology Program (Grant No. 25QNMP109).

**Competing interests:** The authors have declared no conflict of interest.

**Data availability:** The data generated and analyzed during the current study are available from the corresponding author upon reasonable request, subject to institutional and ethical approval.

**Code availability:** The model architecture code has been deposited on GitHub and is publicly available at: https://github.com/GaogeXv/CT_DL_ccRCC_tls2026.

# References

1. Rose, T. L. & Kim, W. Y. Renal Cell Carcinoma: A Review. *JAMA* 332, 1001–1010, doi:10.1001/jama.2024.12848 (2024).
2. Capitanio, U. *et al.* Epidemiology of Renal Cell Carcinoma. *Eur Urol* 75, 74–84, doi:10.1016/j.eururo.2018.08.036 (2018).
3. Bex, A. *et al.* European Association of Urology Guidelines on Renal Cell Carcinoma: The 2025 Update. *Eur Urol* 87, 683–696, doi:10.1016/j.eururo.2025.02.020 (2025).
4. Bedke, J. *et al.* Updated European Association of Urology Guidelines on the Use of Adjuvant Immune Checkpoint Inhibitors and Subsequent Therapy for Renal Cell Carcinoma. *Eur Urol* 87, 491–496, doi:10.1016/j.eururo.2025.01.014 (2025).
5. Rosellini, M. *et al.* Prognostic and predictive biomarkers for immunotherapy in advanced renal cell carcinoma. *Nat Rev Urol* 20, 133–157, doi:10.1038/s41585-022-00676-0 (2022).
6. Fotia, G. *et al.* Current Status of Predictive Biomarker Development in Metastatic Renal Cell Carcinoma. *Curr Oncol Rep* 25, 671–677, doi:10.1007/s11912-023-01395-4 (2023).

7. Usher-Smith, J. A. *et al.* Risk models for recurrence and survival after kidney cancer: a systematic review. *BJU Int* 130, 562–579, doi:10.1111/bju.15673 (2022).
8. Harshman, L. C., Drake, C. G. & Choueiri, T. K. PD-1 blockade in renal cell carcinoma: to equilibrium and beyond. *Cancer Immunol Res* 2, 1132–1141, doi:10.1158/2326-6066.Cir-14-0193 (2014).
9. Attalla, K. *et al.* Prevalence and Landscape of Actionable Genomic Alterations in Renal Cell Carcinoma. *Clin Cancer Res* 27, 5595–5606, doi:10.1158/1078-0432.Ccr-20-4058 (2021).
10. Díaz-Montero, C. M., Rini, B. I. & Finke, J. H. The immunology of renal cell carcinoma. *Nat Rev Nephrol* 16, 721–735, doi:10.1038/s41581-020-0316-3 (2020).
11. Teillaud, J.-L., Houel, A., Panouillot, M., Riffard, C. & Dieu-Nosjean, M.-C. Tertiary lymphoid structures in anticancer immunity. *Nat Rev Cancer* 24, 629–646, doi:10.1038/s41568-024-00728-0 (2024).
12. Zhou, S. *et al.* Interferon Regulatory Factor 4 Recruits Immature B Cells to Signal Tertiary Lymphoid Structure Immaturity and Progression of Clear Cell Renal Cell Carcinoma. *International Journal of Biological Sciences* 21, 3827–3851, doi:10.7150/ijbs.113737 (2025).
13. Fridman, W. H. *et al.* B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. *Nat Rev Clin Oncol* 19, 441–457, doi:10.1038/s41571-022-00619-z (2022).
14. Zhang, Y. *et al.* Tertiary lymphoid structural heterogeneity determines tumour immunity and prospects for clinical application. *Mol Cancer* 23, 75, doi:10.1186/s12943-024-01980-6 (2024).
15. Wang, Y.-Q. *et al.* Integrated analysis of tertiary lymphoid structures and immune infiltration in ccRCC microenvironment revealed their clinical significances: a multicenter cohort study. *J Immunother Cancer* 12, doi:10.1136/jitc-2023-008613 (2024).
16. Xu, W. *et al.* Unveiling the impact of tertiary lymphoid structures on immunotherapeutic responses of clear cell renal cell carcinoma. *MedComm (2020)* 5, e461, doi:10.1002/mco2.461 (2024).
17. Lin, Y. *et al.* Machine learning model for predicting tertiary lymphoid structures and treatment response in triple-negative breast cancer. *NPJ Precis Oncol* 9, 216, doi:10.1038/s41698-025-01012-6 (2025).
18. Wu, W. *et al.* Artificial Intelligence and Multimodality Data Integration Decipher Tertiary Lymphoid Structure Maturity in Gastric Cancer. *Cancer Res* 85, 5128–5140, doi:10.1158/0008-5472.CAN-25-1793 (2025).
19. Zhao, X. *et al.* Preoperative assessment of tertiary lymphoid structures in stage I lung adenocarcinoma using CT radiomics: a multicenter retrospective cohort study. *Cancer Imaging* 24, 167, doi:10.1186/s40644-024-00813-5 (2024).
20. Long, S. *et al.* Spatial patterns and MRI-based radiomic prediction of high peritumoral tertiary lymphoid structure density in hepatocellular carcinoma: a multicenter study. *J Immunother Cancer* 12, doi:10.1136/jitc-2024-009879 (2024).
21. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 310, 2191–2194, doi:10.1001/jama.2013.281053 (2013).

22. Delahunt, B., Eble, J. N., Egevad, L. & Samaratunga, H. Grading of renal cell carcinoma. *Histopathology* 74, doi:10.1111/his.13735 (2019).
23. Paner, G. P. *et al.* Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers. *Eur Urol* 73, 560–569, doi:10.1016/j.eururo.2017.12.018 (2018).
24. Cui, X. *et al.* Tertiary lymphoid structures as a biomarker in immunotherapy and beyond: Advancing towards clinical application. *Cancer Lett* 613, 217491, doi:10.1016/j.canlet.2025.217491 (2025).
25. Bellera, C. A. *et al.* Protocol of the Definition for the Assessment of Time-to-event Endpoints in CANcer trials (DATECAN) project: formal consensus method for the development of guidelines for standardised time-to-event endpoints' definitions in cancer clinical trials. *Eur J Cancer* 49, 769–781, doi:10.1016/j.ejca.2012.09.035 (2013).
26. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 45, 228–247, doi:10.1016/j.ejca.2008.10.026 (2009).
27. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 12, 453–457, doi:10.1038/nmeth.3337 (2015).
28. Schumacher, T. N. & Thommen, D. S. Tertiary lymphoid structures in cancer. *Science* 375, eabf9419, doi:10.1126/science.abf9419 (2022).
29. Lin, J. *et al.* Tertiary Lymphoid Structures are Linked to Enhanced Antitumor Immunity and Better Prognosis in Muscle-Invasive Bladder Cancer. *Adv Sci (Weinh)* 12, e2410998, doi:10.1002/advs.202410998 (2024).
30. Shu, D. H. *et al.* Immunotherapy response induces divergent tertiary lymphoid structure morphologies in hepatocellular carcinoma. *Nat Immunol* 25, 2110–2123, doi:10.1038/s41590-024-01992-w (2024).
31. Hugaboom, M. B. *et al.* Presence of Tertiary Lymphoid Structures and Exhausted Tissue-Resident T Cells Determines Clinical Response to PD-1 Blockade in Renal Cell Carcinoma. *Cancer Discov* 15, 948–968, doi:10.1158/2159-8290.CD-24-0991 (2025).
32. Bera, K., Braman, N., Gupta, A., Velcheti, V. & Madabhushi, A. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol* 19, 132–146, doi:10.1038/s41571-021-00560-7 (2021).
33. Prelaj, A. *et al.* Artificial intelligence for predictive biomarker discovery in immuno-oncology: a systematic review. *Ann Oncol* 35, 29–65, doi:10.1016/j.annonc.2023.10.125 (2023).
34. Li, J. *et al.* The Absence of Intra-Tumoral Tertiary Lymphoid Structures is Associated with a Worse Prognosis and mTOR Signaling Activation in Hepatocellular Carcinoma with Liver Transplantation: A Multicenter Retrospective Study. *Adv Sci (Weinh)* 11, e2309348, doi:10.1002/advs.202309348 (2024).
35. Xu, Y. *et al.* An MRI-based model for preoperative prediction of tertiary lymphoid structures in patients with gallbladder cancer. *Insights Imaging* 16, 189, doi:10.1186/s13244-025-02007-4 (2025).
36. Wang, T. *et al.* MRI-derived radiomics model for predicting intratumoral tertiary lymphoid structures in soft tissue sarcoma. *Insights Imaging* 16, 201, doi:10.1186/s13244-025-02086-3 (2025).

37. Li, Y. *et al.* A novel hybrid model for predicting tertiary lymphoid structures and targeted immunotherapy outcomes in hepatocellular carcinoma: a multicenter retrospective study. *Eur Radiol* 35, 3206–3222, doi:10.1007/s00330-024-11255-9 (2024).
38. Wang, Y. *et al.* Single-cell and spatial transcriptomics implicate a prognostic function of tertiary lymphoid structures in gastric cancer. *Nat Commun* 16, 10435, doi:10.1038/s41467-025-65421-8 (2025).
39. Tokunaga, R. *et al.* B cell and B cell-related pathways for novel cancer treatments. *Cancer Treat Rev* 73, 10–19, doi:10.1016/j.ctrv.2018.12.001 (2018).
40. Molina-Alejandre, M. *et al.* Plasmatic CXCL13 as a biomarker of primary resistance to immunotherapy and impaired peripheral immunity in non-small cell lung cancer. *Biomark Res* 14, 49, doi:10.1186/s40364-026-00929-9 (2026).

# Figures

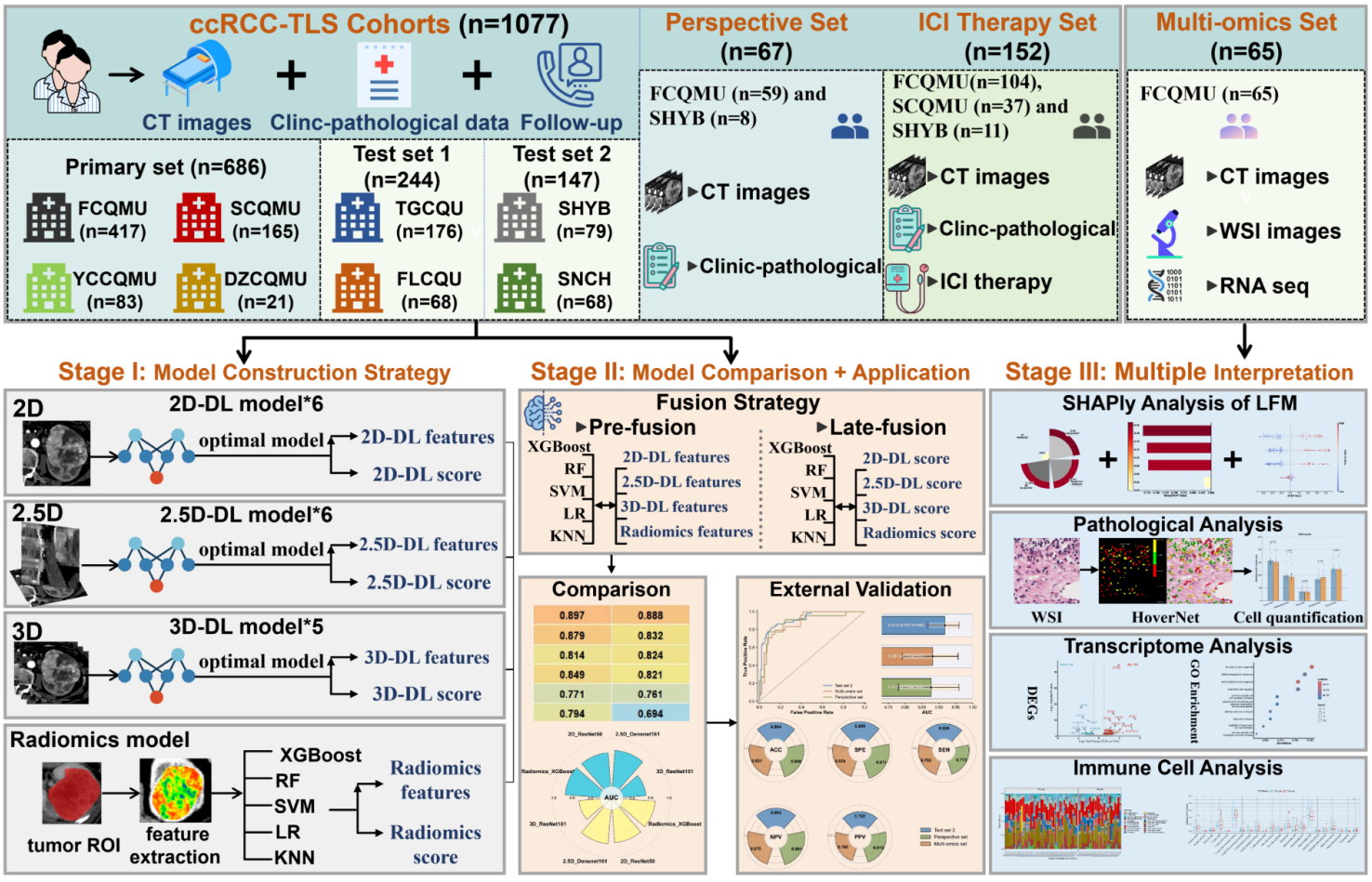



**Figure 1**

**Overall workflow of the study.** This multicenter, retrospective-prospective cohort study enrolled patients with ccRCC from eight tertiary hospitals. The cohort was partitioned into the primary set, test set 1, test set 2, prospective test set, ICI therapy set, and multi-omics test set. **Stage 1:** Preoperative arterial-phase

contrast-enhanced CT (CECT) images were utilized to develop 2D, 2.5D, and 3D deep learning (DL) models, alongside traditional radiomics models. DL features were extracted following the removal of the fully connected layers. **Stage 2:**Five machine learning algorithms were employed for early fusion (feature-level integration) of the extracted features and late fusion (decision-level integration) of the Stage 1 models. Predictive performances were systematically evaluated. **Stage 3:** SHAP analysis was applied to interpret the optimal late fusion model (LFM). Differences in cellular composition, transcriptomic profiles, and immune cell infiltration between LFM-predicted TLS-positive and negative patients were explored in the multi-omics cohort.

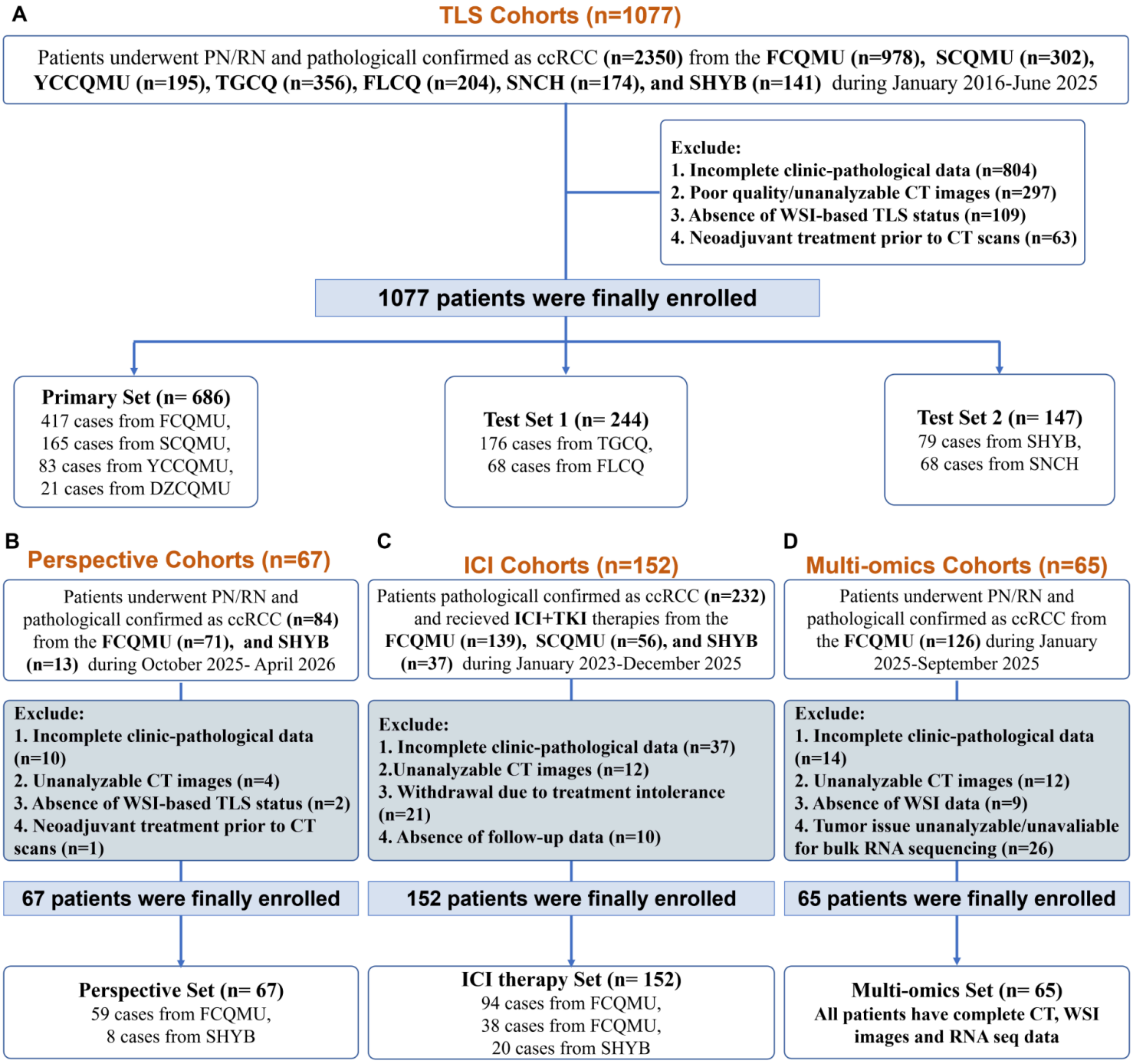
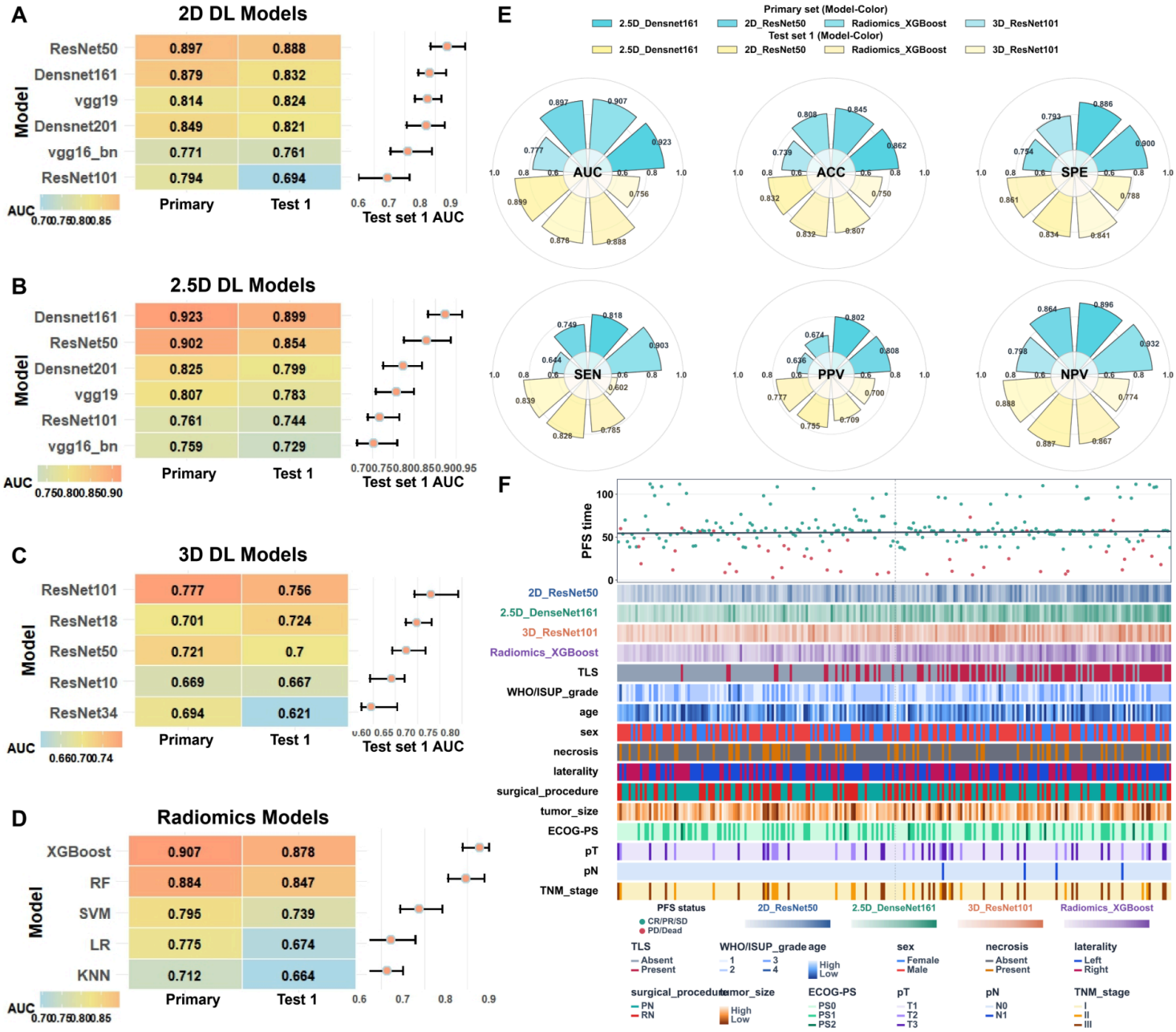



**Figure 2**

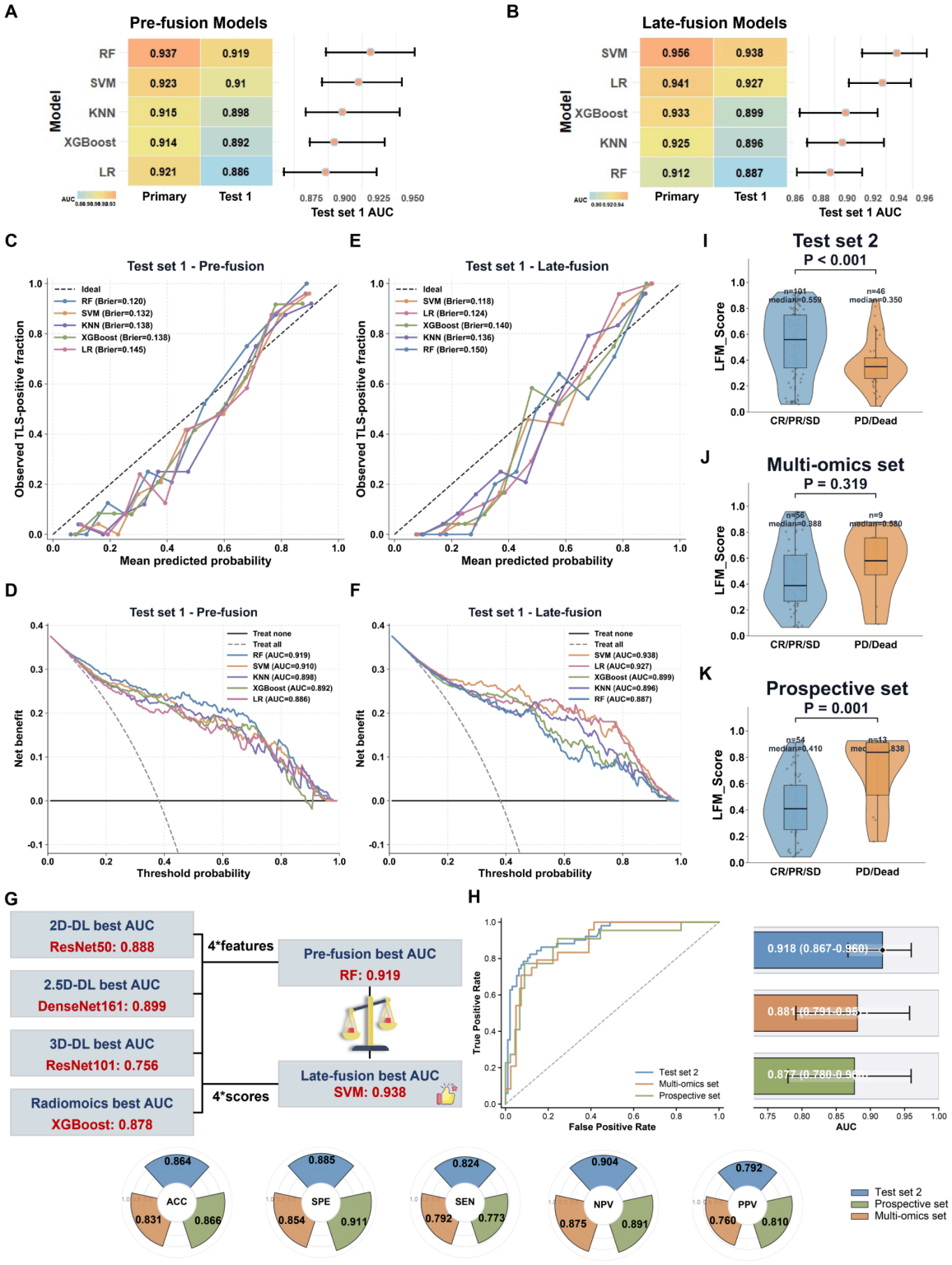
**Flowcharts of patient inclusion and exclusion**. **(A)** The primary TLS cohort. **(B)** The prospective test cohort. **(C)** The ICI efficacy prediction cohort. **(D)** The multi-omics test cohort.




## Figure 3

**Development and performance evaluation of single-modality models**. Area under the curve (AUC) values demonstrating the predictive performance for TLS status across the primary set and test set 1 utilizing **(A)** six 2D DL models, **(B)** six 2.5D DL models, **(C)** five 3D DL models, and **(D)** five traditional machine learning (ML) radiomics models. **(E)** Petal plot comparing the area under the curve (AUC), accuracy (ACC), sensitivity (SEN), specificity (SPE), negative predictive value (NPV), and positive predictive value (PPV) among the optimal 2D-DL, 2.5D-DL, 3D-DL, and ML-radiomics models. **(F)**

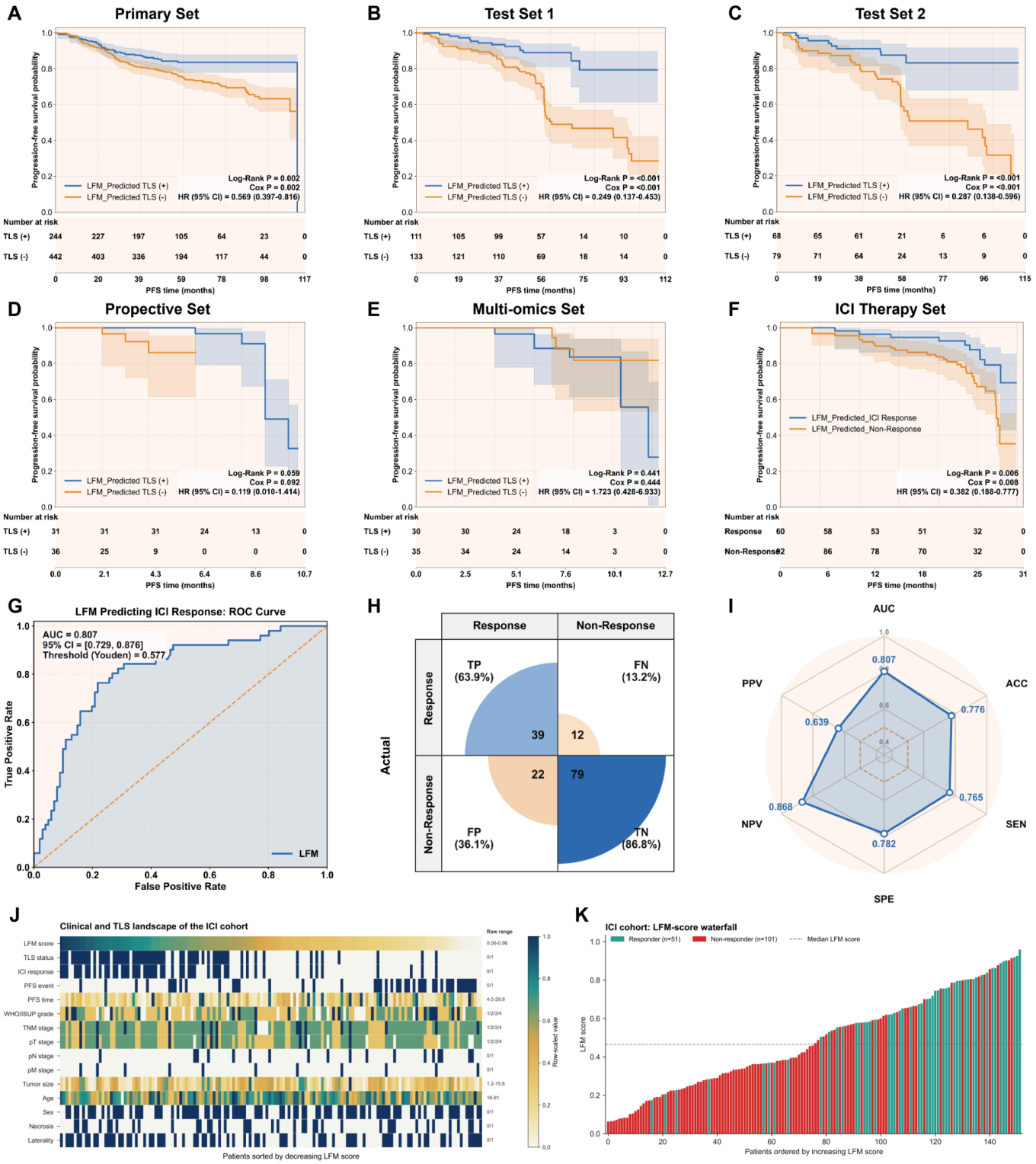
Heatmap illustrating the correlations among risk scores derived from the four optimal models, progression-free survival (PFS), and clinico-pathological characteristics in test set 1.




## Figure 4

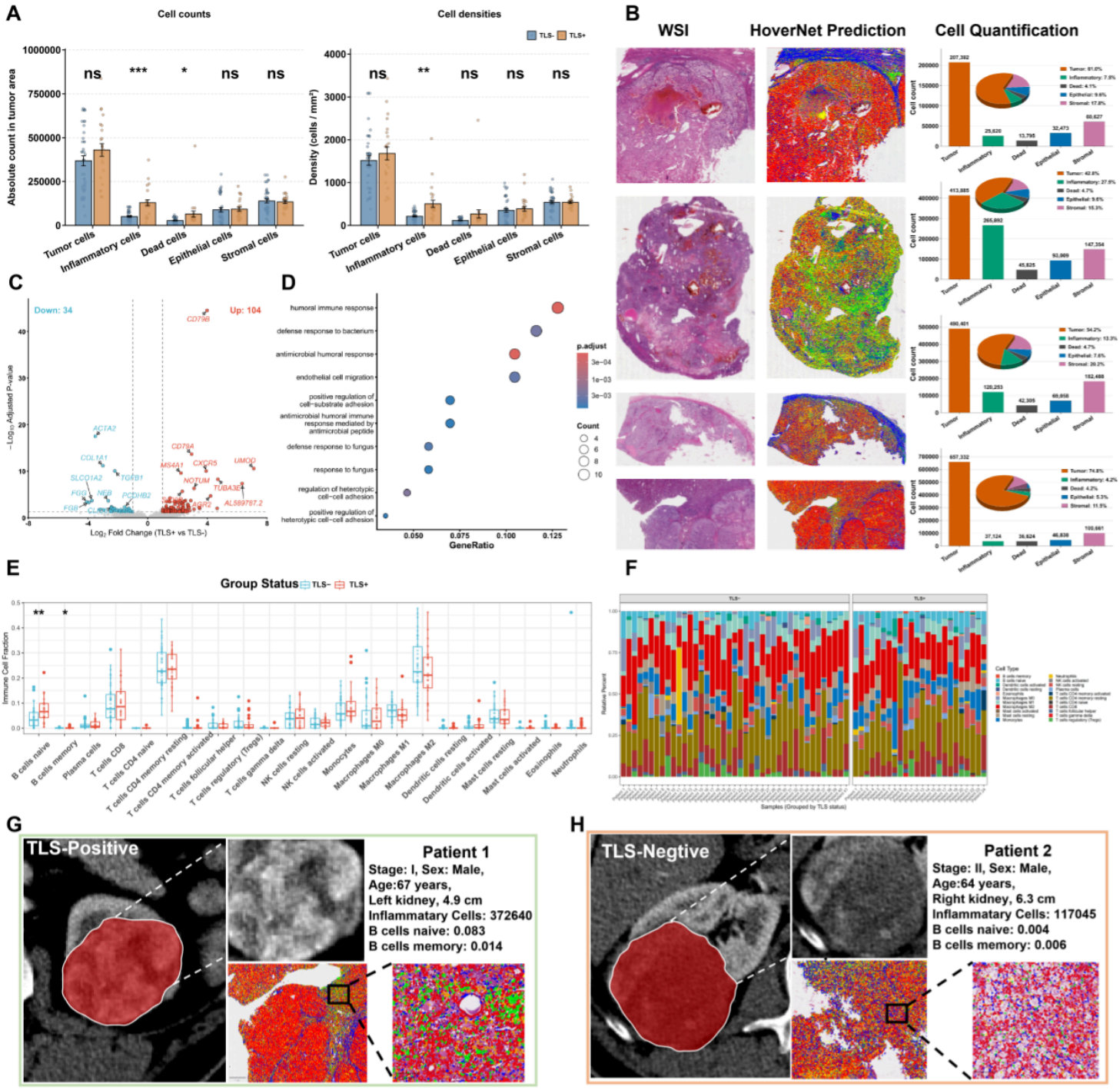
**Construction and validation of multi-modal fusion models.** Area under the curve (AUC) values of the **(A)**early fusion (pre-fusion) and **(B)** late fusion (post-fusion) models in the primary set and test set 1.

Calibration curves **(C, E)** and decision curve analyses **(D, F)** evaluating the optimal early and late fusion models in test set 1. **(G)** Comparison of the predictive performance between the optimal early and late fusion models. **(H)** ROC curves and detailed performance metrics of the LFM for predicting TLS positive in test set 2, the multi-omics set, and the prospective set. Violin plots comparing LFM-derived risk scores between patients with distinct clinical outcomes in **(I)**the test set 2, **(J)** the multi-omics set, and **(K)** the prospective set.



## Figure 5

**Prognostic stratification and prediction of ICI therapeutic response by the LFM.** Kaplan-Meier curves for progression-free survival (PFS) stratified by LFM-predicted TLS status in **(A)** the primary set, **(B)** test set 1, **(C)** test set 2, **(D)** the prospective set, and **(E)** the multi-omics set. **(F)** Kaplan-Meier survival analysis comparing clinical outcomes between LFM-predicted responders and non-responders in the ICI therapy cohort. Predictive evaluation of the LFM for ICI therapy response in advanced ccRCC patients, including **(G)** the ROC curve, **(H)** the confusion matrix, and **(I)** quantitative performance metrics. **(J)** Heatmap demonstrating clinico-pathological correlations with LFM-predicted ICI responses. **(K)** Waterfall plot depicting individual patient tumor responses to ICI therapy, stratified by LFM predictions.




## Figure 6

**Pathological and transcriptomic interpretability of the LFM. (A)** Bar charts comparing the absolute count and density of distinct cell types, quantified via HoVer-Net and QuPath, between LFM-predicted TLS-positive and TLS-negative patients. **(B)**Representative examples (n=4) of whole-slide image (WSI) cellular segmentation and classification utilizing HoVer-Net. **(C)** Volcano plot illustrating differentially expressed genes (DEGs) derived from bulk RNA sequencing of tumor tissues (n=59). **(D)** Functional pathway enrichment analysis of the identified DEGs. **(E)** Stacked bar charts showing the relative proportions of infiltrating immune cells per patient, derived from CIBERSORTx deconvolution. **(F)** Box

plots comparing immune cell fractions between the LFM-predicted TLS-positive and TLS-negative groups. Detailed case presentations of two patients accurately predicted by the LFM as **(G)** TLS-positive and **(H)**TLS-negative. Panels integrate baseline clinical profiles, HoVer-Net–derived cell counts, transcriptomic immune cell proportions, matched preoperative CECT images, and WSI heatmaps.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx