Transfer learning radiomic model predicts intratumoral tertiary lymphoid structures in hepatocellular carcinoma: a multicenter study

Shichao Long,[1,2] Mengsi Li ![ORCID] ,[1] Juan Chen,[1] Linhui Zhong,[1] Ganmian Dai,[3] Deng Pan,[4] Wenguang Liu,[1] Feng Yi,[2] Yue Ruan,[1] Bocheng Zou,[1] Xiong Chen,[5] Kai Fu ![ORCID] ,[2,6,7] Wenzheng Li ![ORCID] [1]

Correspondence to
Professor Wenzheng Li;
wenzheng727@csu.edu.cn

Professor Kai Fu;
fu_kai@csu.edu.cn

ABSTRACT

Background Intratumoral tertiary lymphoid structures (iTLS) in hepatocellular carcinoma (HCC) are associated with improved survival and may influence treatment decisions. However, their non-invasive detection remains challenging in HCC. We aim to develop a non-invasive model using baseline contrast-enhanced MRI to predict the iTLS status.

Methods A total of 660 patients with HCC who underwent surgery were retrospectively recruited from four centers between October 2015 and January 2023 and divided into training, internal test, and external validation sets. After features dimensionality and selection, corresponding features were used to construct transfer learning radiomic (TLR) models for diagnosing iTLS, and model interpretability was explored with pathway analysis in The Cancer Genome Atlas-Liver HCC. The performances of models were assessed using the area under the receiver operating characteristic curve (AUC). The log-rank test was used to evaluate the prognostic value of the TLR model. The combination therapy set of 101 patients with advanced HCC treated with first-line anti-programmed death 1 or ligand 1 plus antiangiogenic treatment between January 2021 and January 2024 was used to investigate the value of the TLR model for evaluating the treatment response.

Results The presence of iTLS was identified in 46.0% (n=308) patients. The TLR model demonstrated excellent performance in predicting the presence of iTLS in training (AUC=0.91, 95% CI: 0.87, 0.94), internal test (AUC=0.85, 95% CI: 0.77, 0.93) and external validation set (AUC=0.85, 95% CI: 0.81, 0.90). The TLR model-predicted iTLS group has favorable overall survival (HR=0.66; 95% CI: 0.48, 0.90; p=0.007) and relapse-free survival (HR=0.64; 95% CI: 0.48, 0.85; p=0.001) in the external validation set. The model-predicted iTLS status was associated with inflammatory response and specific tumor-associated signaling activation (all p<0.001). The proportion of treatment responders was significantly higher in the model-predicted group with iTLS than in the group without iTLS (36% vs 13.73%, p=0.009).

Conclusion The TLR model has indicated accurate prediction of iTLS status, which may assist in the risk stratification for patients with HCC in clinical practice.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Previous studies have demonstrated that the use of CT-based radiomic models to non-invasively predict intratumoral tertiary lymphoid structures (iTLS) in hepatocellular carcinoma (HCC). However, these studies were limited by small sample sizes and single or dual-center designs. To date, no research has explored the potential of magnetic resonance imaging-based deep learning radiomic models in predicting iTLS status.

WHAT THIS STUDY ADDS

⇒ This multicenter research designed a multiphase MRI-based model to estimate the iTLS status in patients with HCC who underwent surgery. The integrated model, which combined deep learning-derived and radiomic-derived scores, attained consistent area under the curve values of 0.85 in both internal and external validation data sets. In addition, the study highlighted the clinical relevance of the predictive biomarker, showing that higher model scores correlated with favorable prognosis and responsiveness to combination therapy-based treatments across four centers with 101 patients with advanced HCC.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The developed deep learning radiomics-based framework demonstrates significant potential in predicting iTLS in patients with HCC, offering valuable insights for improved risk stratification and tailored therapeutic strategies.

INTRODUCTION

In recent times, the CARES-310 project has demonstrated that immune checkpoint inhibitors (ICIs) combined with anti-angiogenic tyrosine-kinase inhibitors provide significant long-term survival benefits compared with sorafenib monotherapy in patients with hepatocellular carcinoma (HCC).[1] Nevertheless, anti-programmed death 1 or ligand 1

(PD-(L)1) plus antiangiogenic treatment (ie, combination therapy) has only benefited 11–30% of patients with HCC.[1–3] Biomarkers, including immune-associated molecules, circulating cancer indicators, and multigene panels associated with tumor diversity are widely used to assess clinical outcomes and treatment effectiveness. However, these indicators often face limitations in terms of accuracy and generalizability.[4]

Tertiary lymphoid structures (TLS) are ectopic lymphoid aggregates found in chronically inflamed tissues and cancers.[5] Studies have shown that TLS promotes T-cell activation, enhances antibody presentation, and facilitates a humoral antitumor response mediated by B cells.[6 7] Recently, intratumor TLS (iTLS) has emerged as a promising biomarker for predicting the efficacy of combination therapy and improving clinical outcomes in various cancers, including melanoma, lung cancer, and HCC.[8–11] Histopathological examination is the gold standard for diagnosing iTLS.[9] Nevertheless, invasive biopsy may induce pain, bleeding and needle tract seeding; therefore, it is usually not applicable for pretreatment evaluation for 90% patients with HCC.[12] Hence, the development of non-invasive tools for assessing iTLS is critical for precise treatment management and evaluating clinical outcomes in patients with HCC.

Combined deep learning radiomic is an emerging medical imaging technique that uses big data and machine learning to quantitatively analyze image features, assessing disease progression, treatment response, and clinical outcomes.[13 14] Compared with traditional radiomic, combined application has shown superior performance in predicting the biological behavior of HCC and the clinical benefits for patients.[15–17] In HCC, two studies have proposed a CT-based classifier for non-invasive iTLS evaluation.[18 19] However, those classifiers lacked independent external and larger cohort validations. Additionally, multiphase MRI has excellent clinical benefits compared with CT, including avoiding radiation exposure, providing superior soft-tissue resolution and diagnostic accuracy for liver disease.[12 20]

Hence, this study aimed to develop a deep learning-based radiomic model using baseline multiphase MRI data for the non-invasive prediction of iTLS status and to explore its potential clinical applications.

METHODS
Study patients
From October 2015 to January 2023, consecutive patients who had undergone hepatectomy were retrospectively enrolled for model construction and verification. The training set (n=307) and internal test set (n=76) were derived from Institute 1, while the external validation set came from Institutes 2, 3, and 4 (n=277). Inclusion criteria included: (1) preoperative contrast-enhanced MRI data within 30 days and (2) histologically confirmed HCC and available whole-slide image of H&E staining. Exclusion criteria were: (1) insufficient MRI or histological image

quality and (2) initial treatment with alternative therapies (eg, liver transplantation, transarterial chemoembolization, radiotherapy, and systemic therapy). Patients were followed with ultrasound, dynamic CT, or MRI at 1-month post-surgery, every 3 months for the first 2 years, and then every 6 months thereafter until death or the last follow-up date of June 30, 2024. Relapse-free survival (RFS) was defined as the time from surgery to first recurrence, metastasis, or death from any cause. Overall survival (OS) was defined as the time from surgery to any-cause death.

The combination therapy set included patients with advanced HCC who underwent combination therapy with anti-PD-(L)1 and antiangiogenic agents as first-line treatment between January 2021 and January 2024 at Institutes 1 (n=69), 2 (n=13), 3 (n=10), and 4 (n=9). Patients were required to have at least a 2-month interval between prior therapies (eg, transarterial chemoembolization, radiofrequency ablation, hepatectomy) and the initiation of combination therapy and have at least one tumor response assessment after initiating combination therapy. Baseline MRI data and objective response rate assessments were obtained via radiologic examination every 2 months (±2 weeks) according to the Response Evaluation Criteria in Solid Tumors (V.1.1).[21] Patients with partial or complete response were regarded as responders and progress or stable disease as non-responders.[22] Progression-free survival (PFS) was the interval between treatment initiation and the occurrence of disease progression or death from any cause. OS was also defined as the date of first combination therapy to any-cause death. Patients were censored at the time of withdrawal from treatment or the last follow-up. Detailed information on drugs is provided in online supplemental method S1.

Baseline characteristics
Baseline data included age, sex, cirrhosis, Barcelona Clinic Liver Cancer stage, alpha-fetoprotein, hepatitis B surface antigen (HBsAg), hepatitis C virus antibody (HCVAb) status, aspartate transaminase (AST), alanine aminotransferase (ALT), and platelet count. Radiographic features associated with iTLS status,[18] along with imaging features defined by the Liver Imaging Reporting and Data System (V.2018),[23] were reviewed in all patients by reader 1 with 9 years of abdominal imaging experience who was blinded to the clinical and pathologic findings. Reader 2 with 20 years of abdominal imaging experience analyzed imaging features independently in 40 randomly selected patients for testing interobserver variability. Details of radiographic features are elucidated in online supplemental method S2.

Characterization of iTLS
The aggregated ectopic lymphoid cells (especially B and T cells) were identified as TLS, characterized by their lack of integrated organized structures such as capsules. H&E-stained whole-slide images were analyzed for iTLS using an automated pipeline (https://github.com/YuMeng-W/TumSeg-main) while blinded to clinical and MRI data.[24]

Table 1 Clinical and radiographic features of patients

Characteristics	Training set (n=307)	Internal test set (n=76)	External validation set (n=277)	Combination therapy set (n=101)
Clinical features				
Age (y)*	52 [45–62]	52 [47–61]	57 [49–65]	58 [51–66]
Sex (male)	233 (75.9)	58 (76.3)	244 (88.1)	85 (84.2)
HBsAg/HCVAb status (positive)	224 (73.0)	52 (68.4)	240 (86.6)	76 (75.2)
BCLC (0-A)	205 (66.8)	57 (75.0)	192 (69.3)	19 (18.8)
Cirrhosis (present)	250 (81.4)	63 (82.9)	231 (83.4)	85 (84.2)
AFP (>400 ng/mL)	100 (32.6)	35 (46.1)	88 (31.8)	35 (34.7)
ALT (UI/l)*	33 [23–48]	34 [24–47]	45 [28–77]	39.9 [29–57]
AST (UI/l)*	37 [26–51]	39 [29–60]	34 [25–50]	39.6 [30–70]
PLT (×10⁹/L)*	144 [112–200]	145 [107–181]	161 [128–214]	186 [135–231]
Radiographic features				
Max diameter (cm)*	5.0 [3.1–7.5]	6.4 [4.0–9.2]	5.16 [3.2–7.6]	7.14 [5.63–10.7]
Number (≥2)	60 (19.5)	15 (19.7)	74 (26.7)	23 (22.8)
Nonrim APHE (present)	238 (77.5)	55 (72.4)	217 (78.3)	92 (91.1)
Nonperipheral washout (present)	275 (89.6)	70 (92.1)	250 (90.3)	91 (90.1)
Rim APHE (present)	54 (17.6)	19 (25.0)	53 (19.1)	5 (4.95)
Enhancing capsule (present)	225 (73.3)	66 (86.8)	189 (68.2)	63 (62.4)
Corona enhancement (present)	56 (18.2)	18 (23.7)	69 (24.9)	33 (32.7)
Intratumor hemorrhage (present)	92 (30.0)	27 (35.5)	67 (24.2)	20 (19.8)
Necrosis (present)	171 (55.7)	49 (64.5)	189 (68.2)	74 (73.3)
Intratumor arteries (present)	109 (35.5)	33 (43.4)	120 (43.3)	46 (45.5)
Unsmooth margin (present)	134 (43.6)	34 (44.7)	145 (52.3)	35 (34.7)
iTLS				
iTLS status (present):	143 (46.6)	29 (38.2)	147 (53.1)	NA

Note: Unless indicated otherwise, data are the number of patients, and data in parentheses are percentages.
*Data are medians, with IQRs in square brackets.
AFP, alpha-fetoprotein; ALT, alanine aminotransferase; APHE, arterial phase hyperenhancement; AST, aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancer; HBsAg, hepatitis B surface antigen; HCVAb, hepatitis C virus antibody; iTLS, intratumor tertiary lymphoid structure; NA, not applicable; PLT, platelet count.

A random subset of 50 images was chosen to assess the agreement between the automated algorithm and the manual evaluation conducted by a senior pathologist with 10 years of experience in liver pathology. Briefly, the area of the pre-segmented intratumoral region was calculated, and lymphoid aggregates with a diameter or major axis of the ellipse exceeding 150 μm were extracted. Following the previous report,[25] patients were categorized as the presence of iTLS (iTLS+) and the absence of iTLS (iTLS−).

MRI technique and imaging preprocessing

MRI data were required using 3.0-T or 1.5-T scanners (General Electric SIGNA HDxt and Discovery MR750; Siemens Verio and Prisma). MRI sequences and parameters are detailed in online supplemental table S1 and online supplemental method S3. To mitigate bias introduced by the magnetic field strength, N4 bias field correction was applied to adjust the density distribution and reduce the differences between scanners. Subsequently,

MRI signal intensity normalization was performed to ensure the comparison between patient images. Finally, in order to minimize the pixel spacing in MRI imaging, we adjusted voxel resampling to 1 mm. Detailed methods of image preprocessing are provided in online supplemental method S3.

Model development and performance

Image segmentation and features extraction

Considering the generalized application of the models,[23 26 27] the region-of-interests for lesion boundaries were manually contoured on T2-weighted (T2WI) and contrast-enhanced T1-weighted images of arterial phase (AP) and portal venous phase (PVP) by two abdominal radiologists (reader 3 and reader 4, with at least 6 years of experience in liver imaging). To ensure reliability, readers 3 and 4 repeated manual delineations for 40 randomly selected patients.[28] Details are in online supplemental method S4. The region-of-interests were used to extract radiomic features using

PyRadiomics software (V.2.11, https://github.com/Radiomics/pyradiomics), which adheres to the Image Biomarker Standardization Initiative guidelines.[29] Transfer learning-based (TL) features were extracted from the ResNet50 network, pretrained on ImageNet and fine-tuned with lesion boundaries in the training data set.[30] Detailed parameters of feature extraction are provided in online supplemental methods S5 and S6.

Feature selection

Feature selection involved interclass and intraclass correlation coefficients (ICC), Mann-Whitney U test with Bonferroni correction ($p<0.05$), minimum redundancy maximum relevance, and least absolute shrinkage and selection operator (online supplemental method S7).

Model development

After z-score normalization of features, two single-modality models were created using logistic regression with selected TL and radiomic features, respectively, to predict iTLS status. A previous study constructed a clinical-radiological (CR) predictive model for identifying the presence of iTLS, combining variables such as intratumor arteries (yes or no) and hemorrhage (yes or no), AST levels (<33.2 UI/l/≥33.2 UI/L), HBsAg status or HCVAb status (yes or no), and platelet count ($<186.5\times10^9$/L/$\geq186.5\times10^9$/L), with cut-offs stratified by the maximum Youden Index of the CR model.[18] Consistent with this study, we have also employed the same cut-offs for these clinical variables. The output model score was defined as the predictive probability generated using the predict function of the stats R package. Univariate logistic regression analysis was performed and identified the variables with a p value<0.1. Subsequently, those factors were selected using multivariate stepwise backward analyses to select optimal combined variables according to the lowest Akaike information criterion (AIC) value and develop the combined TL radiomic (TLR) model (online supplemental table S2). The patients were categorized into predicted iTLS+ and iTLS– groups based on the Youden Index, which was used to calculate the optimal cut-off using the bootstrapping algorithm (1,000 iterations) in the training set.

Model performance and comparison

Model performance was evaluated using the area under the receiver operating characteristic curve (AUC). DeLong test and integrated discrimination improvement (IDI) were used to compare predictive performance between those models. The likelihood ratio test assessed the overall goodness-of-fit between nested models. Calibration curves and the Hosmer-Lemeshow test were used to evaluate the goodness-of-fit between model predictions and actual iTLS status. Decision curve analysis was used to estimate the clinical utility and net benefit of each model.
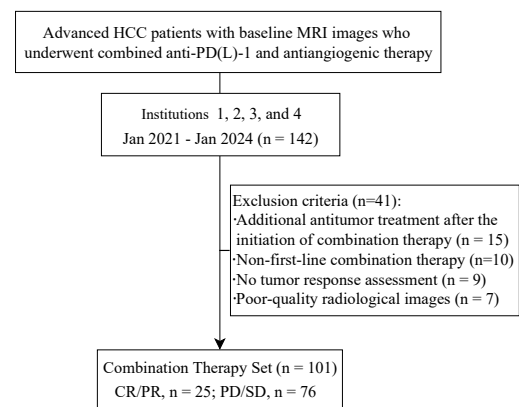
A
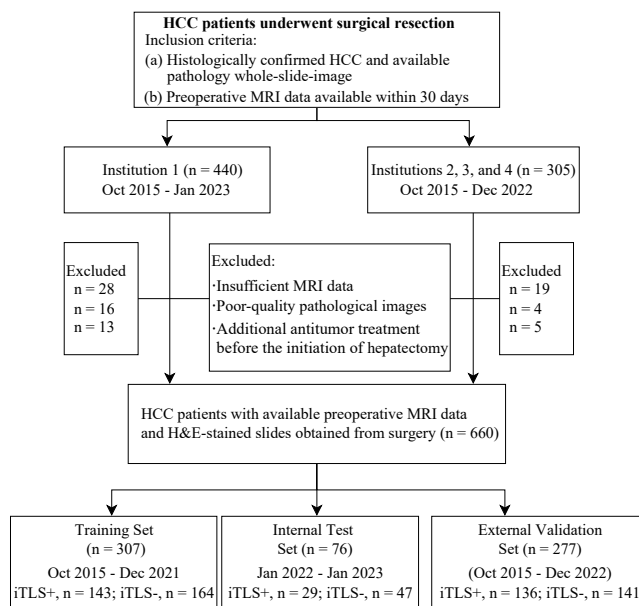
B

Figure 1 Recruitment criteria. (A–B) The flow diagram illustrates the enrollment process of patients with HCC from four centers, where patients are categorized into those receiving surgical treatment for model building, testing, and validation; and those receiving combination therapy for treatment response assessment within the model. CR, complete responses; HCC, hepatocellular carcinoma; iTLS, intratumor tertiary lymphoid structure; PD, progress disease; anti-PD(L)-1, anti-programmed death 1 or ligand 1; PR, partial responses; SD, stable disease; TCGA, The Cancer Genome Atlas; TCIA, The Cancer Imaging Archive.

Table 2 Models performance for predicting the intratumor tertiary lymphoid structures status

	AUC (95% CI)	Sensitivity	Specificity	Accuracy	P value*
TLR model					
Training set	0.91 (0.87, 0.94)	85 [122/143]	84 [137/164]	84 [259/307]	<0.001
Internal test set	0.85 (0.77, 0.93)	97 [28/29]	68 [32/47]	79 [60/76]	<0.001
External validation set	0.85 (0.81, 0.90)	83 [113/136]	76 [107/141]	79 [220/277]	<0.001
TL model					
Training set	0.88 (0.84, 0.92)	78 [111/143]	88 [144/164]	83 [255/307]	<0.001
Internal test set	0.78 (0.67, 0.89)	79 [23/29]	81 [38/47]	79 [61/76]	<0.001
External validation set	0.78 (0.72, 0.83)	82 [112/136]	67 [94/141]	74 [206/277]	0.002
Radiomic model					
Training set	0.72 (0.67, 0.78)	78 [111/143]	59 [97/164]	67 [208/307]	0.006
Internal test set	0.71 (0.58, 0.83)	59 [17/29]	85 [40/47]	75 [57/76]	0.008
External validation set	0.71 (0.65, 0.77)	61 [83/136]	76 [107/141]	68 [190/277]	0.01
Clinical-radiological model					
Training set	0.61 (0.54, 0.67)	71 [101/143]	54 [88/164]	61 [189/307]	ref
Internal test set	0.50 (0.37, 0.64)	38 [11/29]	79 [37/47]	63 [48/76]	ref
External validation set	0.60 (0.53, 0.66)	62 [84/136]	60 [84/141]	60 [168/277]	ref

Note: Unless otherwise specified, data are percentages, with proportions of patients (numerator/denominator) in square brackets.
*P value was calculated using the DeLong test with the clinical-radiological model as the reference (ref) in the corresponding data, and data in parentheses are 95% CIs. The sensitivity, specificity, and accuracy of the models in the three data sets were calculated using the cut-off values that maximize the Youden Index in the training set.
AUC, area under the receiver operating characteristic curve; TL, transfer learning; TLR, TL radiomic.

Model interpretability

RNA sequencing, histological imaging, baseline multiphase MRI, and manual labels data from 14 patients with HCC were collected from The Cancer Imaging Archive (TCIA) for evaluating the model interpretability. Briefly, pathway enrichment was performed using the differential expression genes between model-based groups. Pearson correlation analysis was conducted on the relationship between model-derived score and significant pathway score using the single sample gene set enrichment algorithm. Details are presented in online supplemental method S8.

Statistical analysis

Statistical analysis was performed using R software (V.4.3.1; https://www.r-project.org) and Python (V.3.7; https://www.python.org). Quantitative data are presented as mean (SD) or median (IQR), and qualitative data are reported as numbers and percentages. Continuous variables were
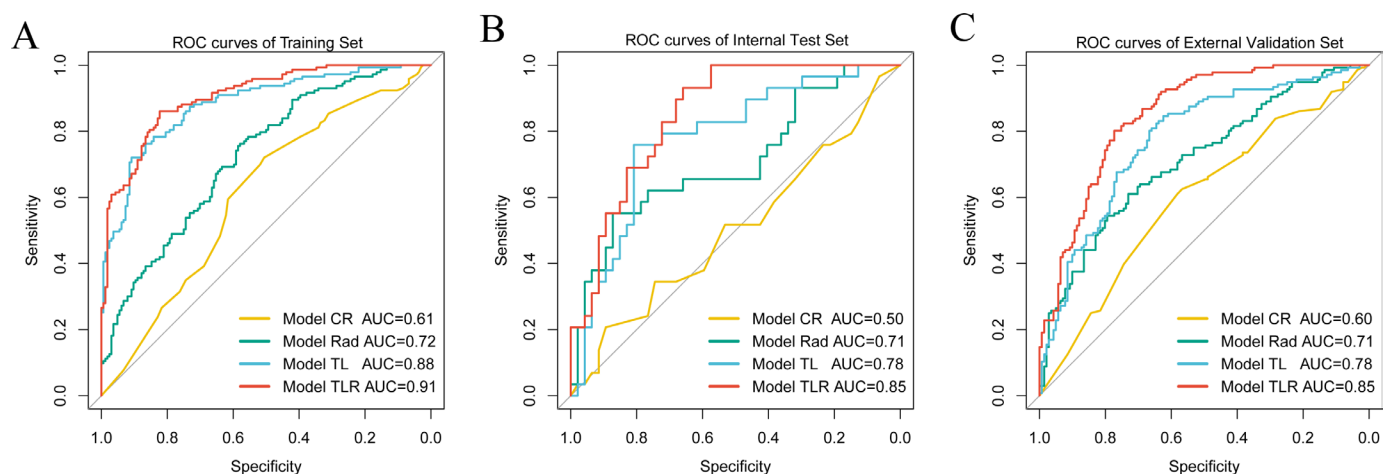

Figure 2 Diagnostic performance of models. (A–C) Areas under the receiver operating characteristic curve (AUCs) of different models for predicting intratumor tertiary lymphoid structures status in the training set (n=307), internal test set (n=76), and external validation set (n=277). Yellow line=clinical-radiographic (CR) model, blue line=transfer learning-based (TL) model, green line=radiomic (Rad) model, red line=combined transfer learning-based radiomic (TLR) model. ROC, receiver operating characteristic curve.
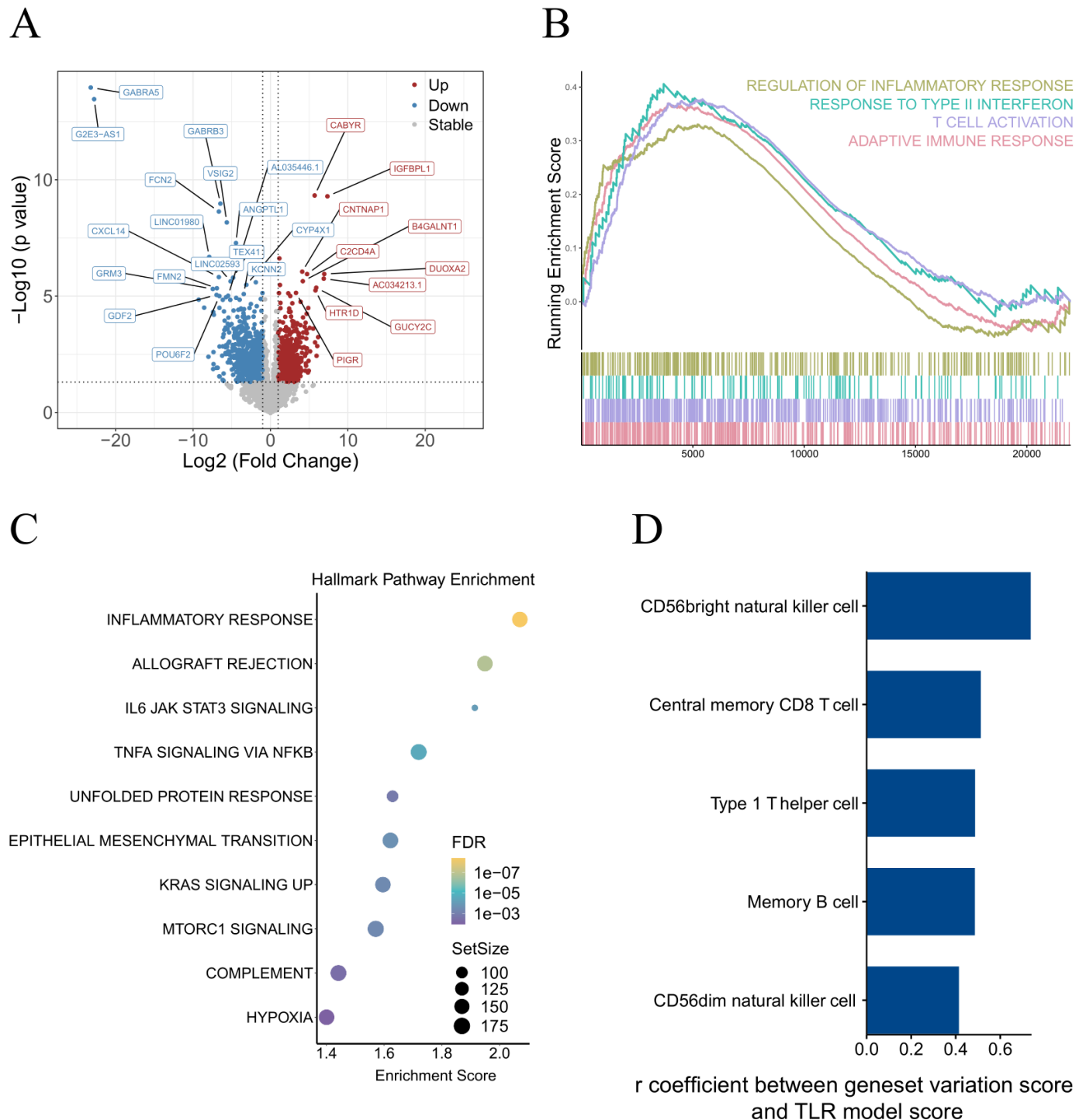
Figure 3 Biological patterns via transcriptome analysis of TLR model. (A) Volcano plot shows the differentially expressed genes between TLR model-predicted iTLS+ group and predicted iTLS− group. (B) Gene Set Enrichment Analysis shows the pathways associated with the TLR-related differentially expressed genes from the Gene Ontology Biological Function database. (C) The bubble plot shows the result of hallmark pathway analysis. FDR=adjust p value using false discovery rate, SetSize=the count of gene sets. (D) Bar plots show the top-five representative correlated pathways with the TLR model score based on the gene set variation analysis scores of immune signatures. iTLS, intratumor tertiary lymphoid structures; TLR, transfer learning radiomic.

compared using the Student's t-test or Mann-Whitney U test. Categorical variables were compared using the χ^2 test. The kappa analysis and ICC were used to evaluate interobserver agreement for qualitative and quantitative analyses, respectively. Univariate and multivariate regression analyses

identified risk factors using ORs, HRs, and 95% CIs. The sample size was based on an expected prevalence of 31.7% to ensure 90% power and accommodate a maximum of 32 predictors, necessitating a minimum of 333 participants. The AUC, sensitivity, specificity, and accuracy of the model
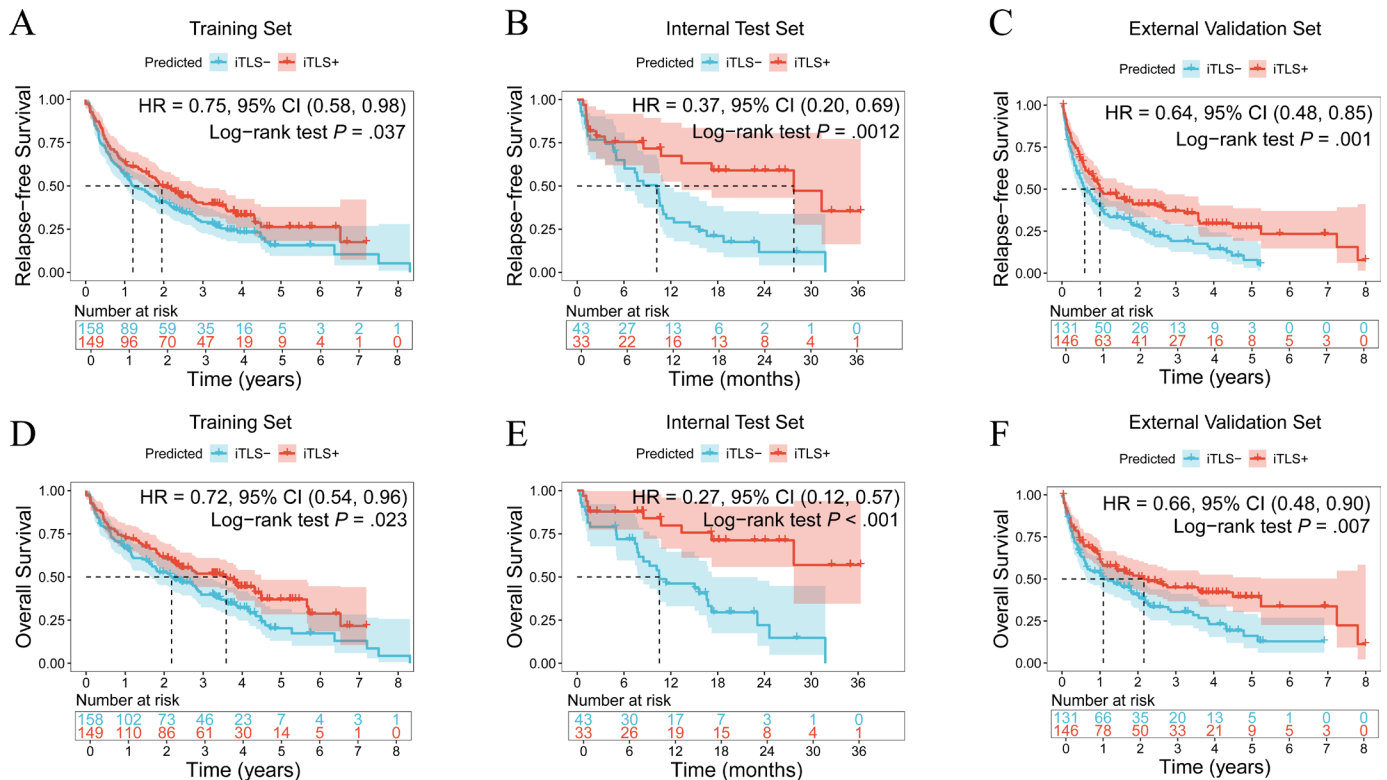
Figure 4 Kaplan-Meier curves of relapse-free survival and overall survival between different iTLS status across three sets. (A–F) Survival differences for the relapse-free survival (A–C) and overall survival (D–F) between transfer learning radiomic model-predicted iTLS+ and iTLS– groups in the training, internal test, and external validation sets. iTLS, intratumoral tertiary lymphoid structure.

were evaluated using the cut-off value that maximized the Youden Index, which was determined from the training set. Kaplan-Meier (K-M) analysis with log-rank test was used to compare the prognosis of different groups. The significance threshold was set at p<0.05 (two-sided) following the Holm-Bonferroni correction.

RESULTS

Study population

The flowchart of this study is shown in online supplemental figure S1. Out of 745 patients who underwent hepatectomy, 660 patients (308 iTLS+ and 352 iTLS–) were eligible for model development and validation. High consistency (κ statistic=0.906, 95% CI: (0.84, 0.95), p<0.001) was observed between the assessments of senior pathologist and automated algorithm. Among the 307 patients in the training cohort, 143 (46.6%) were histologically confirmed to have iTLS. In the internal test set, 29 (38.2%) of 76 patients had iTLS, while in the external validation set, 136 (49%) of 277 patients were confirmed to have iTLS (table 1) (figure 1A). The combination therapy set included 101 patients with advanced HCC: 25 (24.7%) had a response (partial response, n=21; complete response, n=4), while 76 (75.3%) had a non-response (stable disease: n=58, progressive disease: n=18) (figure 1B). The TCIA set included 14 patients, 5 (35.7%) with iTLS and 9 (64.3%) without iTLS (online

supplemental figure S2). The within-sets differences for variables were not significant between the iTLS+ and iTLS– groups, except for the maximum tumor diameter (p<0.001), intratumor hemorrhage (p=0.02), and intratumor necrosis (p=0.004) in the external validation set, and intratumor necrosis (p=0.03) in the internal test set (online supplemental table S3). Additionally, inter-reader agreement of radiographic features (kappa values, 0.62–0.95) between radiologists is provided in online supplemental table S4.

Model development

After feature selection, 19 TL (7 in T2WI, 4 in AP, and 8 in PVP phases) and 13 radiomic features (4 in T2WI, 5 in AP, and 4 in PVP phases) were used to construct the TL model and Rad models, respectively (online supplemental figures S3 and S4). The CR model-derived score was calculated based on a previous study.[18] The univariate logistic analysis demonstrated that ALT, intratumor hemorrhage, intratumor necrosis, Rad and TL model scores were significantly associated with iTLS+ in the training set (online supplemental table S5). After stepwise backward multivariate analysis, the TL (OR=5.91 (95% CI: 4.00, 8.75); p<0.001) and Rad (OR=2.56 (95% CI: 1.73, 3.78); p<0.001) model scores were selected to construct the TLR model, which achieved the lowest AIC value (online supplemental tables S6 and S7).
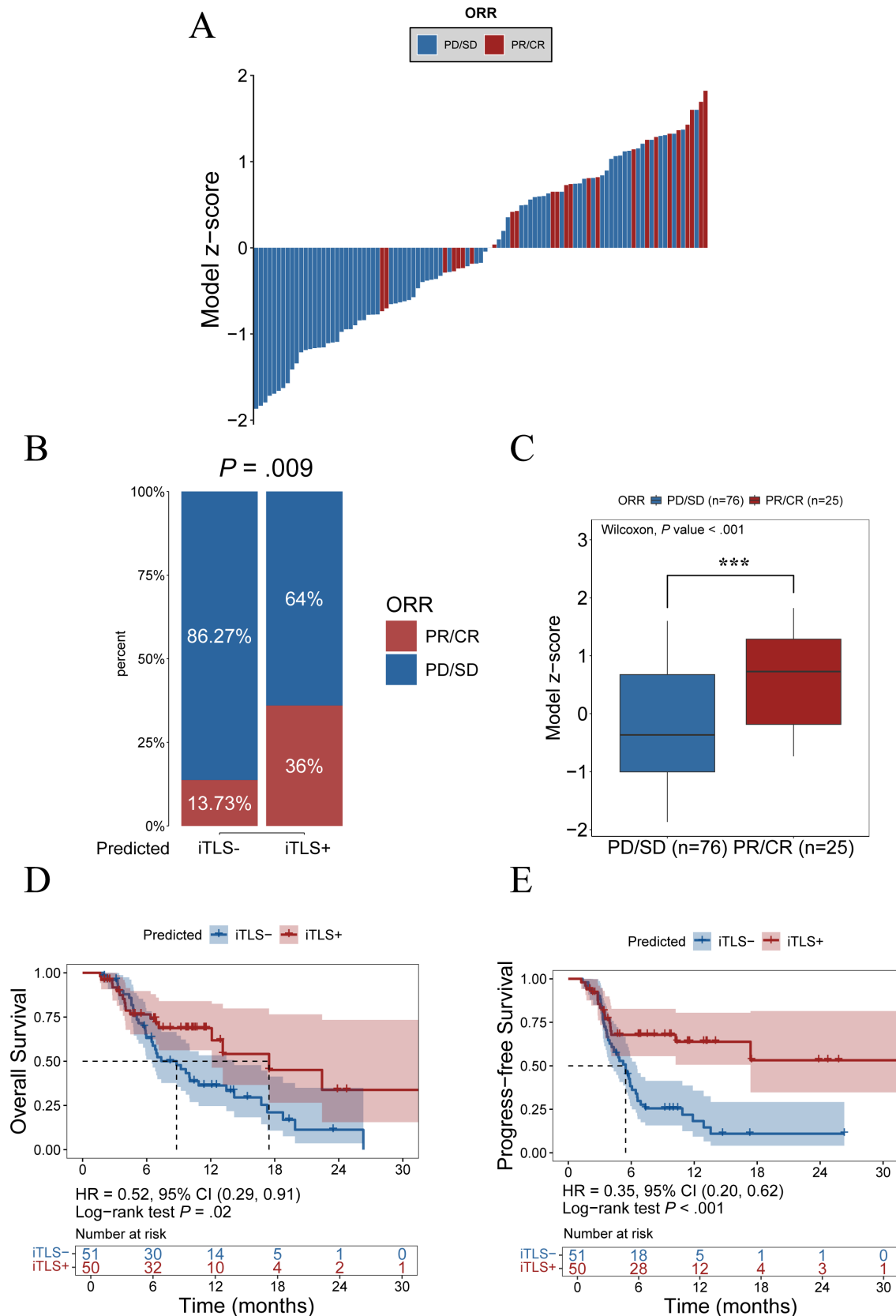
Figure 5 Prognostic value of the TLR model in combination treatment set. (A) Waterfall plot showing the model-score distribution of patients achieving partial/complete responses (PR/CR) and progress/stable disease (PD/SD). (B) The proportion of combination-therapy response subpopulations in the iTLS+ group compared with TLS− groups. (C) Distribution of model-derived score between PD/SD and PR/CR groups. P value was calculated using the Mann-Whitney U test. (D–E) The Kaplan-Meier curve of overall survival and progress-free survival according to TLR model score. ORR, objective response rate, z-score, standard TLR model score. iTLS, intratumoral tertiary lymphoid structure; TLR, transfer learning radiomic.

Comparison of predictive performance between models

Significant differences in the four model scores (range from −2.9 to 2.9) between the iTLS+ and iTLS− groups in all sets are shown in online supplemental figure S5 (all p<0.01), except for the CR model score (p=0.95) in the internal test set. The AUCs of the TL model and the Rad model for predicting iTLS status were 0.88 and 0.72 in the training set, 0.78 and 0.71 in the internal test set, and 0.78 and 0.71 in the external validation set, respectively (table 2 and figure 2). The TLR model showed excellent discrimination for diagnosing iTLS with AUCs of 0.91 (95% CI: 0.87, 0.94) in the training set, 0.85 (95% CI: 0.77, 0.95) in the internal test set, and 0.85 (95% CI: 0.81, 0.90) in the external validation set (table 2 and figure 2). These AUCs of the TL model, the Rad model and the TLR model were significantly higher than the CR model in identifying iTLS status across three sets (all p<0.05) (table 2). The TLR model has better performance for predicting iTLS status than Rad (all DeLong test p<0.001; all IDI p<0.05) and TL models (all DeLong test p<0.01; all IDI p<0.05) across three sets (online supplemental table S8), and enables optimization to achieve higher accuracy across all data sets compared with Rad and TL models (likelihood ratio test p<0.001) (online supplemental table S9). No significant differences were observed in TLR model performance among the training, internal test, and external validation sets: training versus internal test (p=0.23), training versus external (p=0.05), external validation versus internal test (p=0.99), which indicated the good generalizability and stability of the TLR model. Calibration curves for the TLR model-predicted iTLS status showed no significant deviation from the actual iTLS status in all data sets (p=0.95 in training, p=0.15 in internal test, and p=0.49 in external validation sets) (online supplemental figure S6A–C). Decision curve analysis indicated a preferable net benefit of the TLR model over other models especially for the CR model across the range of threshold probabilities (online supplemental figure S6D–F). Online supplemental figure S7 shows representative iTLS+ and iTLS− H&E-stained images with TLR model-derived score and corresponding preoperative multiphase MRI images.

TLR model interpretability

The baseline characteristics of the TCIA cohort are presented in online supplemental table S10. The 805 upregulated and 1014 downregulated differential expression genes were identified between the predicted iTLS+ group versus predicted iTLS− group (figure 3A). Gene Set Enrichment Analysis demonstrated associations between the predicted iTLS+ group and inflammatory and adaptive immune responses (figure 3B). Enrichment analysis revealed that differentially expressed genes were mainly associated with immune pathways and specific tumor signaling, including IL-6 JAK STAT3 SIGNALING, TNFA SIGNALING VIA NFKB, and MTORC1

SIGNALING (figure 3C, all p<0.001). Gene set variation analysis showed that TLS-associated gene sets were positively correlated (r=0.42–0.74, all p<0.001) with the TLR model score (figure 3D).

Prognostic value of the TLR model in the surgery set

The median follow-up duration was 17 months (IQR, 5.6–35) among patients with HCC who underwent surgery. The patients with histologically confirmed iTLS have longer median survival time of RFS and OS across three data sets (online supplemental figure S8A–F, all p<0.01). The K-M analysis indicated that the predicted iTLS+ group had longer RFS (figure 4A–C) and OS (figure 4D–F) compared with the predicted iTLS− group in the training set (RFS of HR=0.75 (95% CI: 0.58, 0.98), p=0.037; OS of HR=0.72 (95% CI: 0.54, 0.96), p=0.023), internal test set (RFS of HR=0.37 (95% CI: 0.20, 0.69), p=0.0012; OS of HR=0.27 (95% CI: 0.12, 0.57), p<0.001), and external validation set (RFS of HR=0.64 (95% CI: 0.48, 0.85), p=0.001; OS of HR=0.66 (95% CI: 0.48, 0.90), p=0.007).

Prognostic value of the TLR model in combination treatment set

The group with an increasing TLR model score exhibited a favorable response rate in both continuous and categorical terms (figure 5A,B). The response group (n=25, mean±SD: 0.62±0.77) had higher TLR model scores than the non-response group (n=76, mean±SD: −0.20±0.98) (p=0.001, figure 5C). The predicted iTLS+ group had longer OS (HR=0.52 (95% CI: 0.29, 0.91), p=0.02) and PFS (HR=0.35 (95% CI: 0.20, 0.62), p<0.001) compared with the predicted iTLS− group (figure 5D,E).

DISCUSSION

In HCC, iTLS serve as a promising indicator for identifying the patients with favorable prognosis. In this study, we developed a TLR model based on baseline multiphase MRI to non-invasively predict iTLS status and investigated its clinical utility in HCC. The TLR model demonstrated robust performance across multiple institutions and reflected biological patterns associated with the inflammatory tumor immune microenvironment. Furthermore, the TLR model-based stratification effectively evaluated prognostic outcomes and treatment responses in patients with advanced HCC.

The presence of iTLS significantly influences prognosis and response to ICI-based therapy in patients with HCC.[11 19] In preoperative immune checkpoint blockade trials, TLS is composed of CD20+ B cells surrounded by T cells, forming lymphoid follicles of varying maturity.[6 31 32] Similar, in HCC, an increase in CD4+ and CD8+ T cells expressing TCF7 within TLS indicates the maturation of T cells within TLS.[11 31] As a common marker within tumors, TLS is associated with good prognosis, a correlation that is independent of CD8+ T cell infiltration and PD-L1 expression.[33] Previous

studies,[11 18 19 25 34] mainly conducted in patients with resected HCC, have found that the presence of iTLS is significantly correlated with favorable RFS, but the relationship with OS remains controversial. We hypothesize that this controversy may arise from variations in treatment modalities and etiology of HCC. To minimize the impact of inconsistent treatment methods, only patients who underwent surgical resection were included in our analysis. Moreover, viral-infection tumor was observed with higher T-cell infiltration, which may influence the formation of TLS and prognosis in viral-associated tumors.[19 35–37] Interestingly, the TLR model-based group of 14 patients from the TCIA cohort showed consistent biological functionality, suggesting that our model may not be influenced by virus infection. However, this finding still requires validation with larger and more diverse cohorts and in HCC cohorts with different etiologies in subsequent studies.

HCC is a type of tumor that can be directly diagnosed through serum markers and typical imaging features.[12 20] According to the existing classification system, the risk-to-benefit ratio of percutaneous biopsy is not proportional, making it difficult to obtain tissue samples before treatment, which leads to the inability to acquire important pathological information before treatment.[12] A previous study proposed a CT-based clinical radiographic model for diagnosing iTLS status with an average AUC of 0.75, demonstrating good predictive ability.[18] However, the model's generalizability was limited by an optimal cut-off value specific to a single center, resulting in an average AUC of 0.58 in our cohorts. In recent studies, a radiomic-based model could non-invasively predict the iTLS status in different solid tumors, including intrahepatic cholangiocarcinoma,[38 39] breast cancer,[40] lung adenocarcinoma[41] and HCC.[18 19] Li *et al* and his team develop a CT-based radiomic model to predict iTLS status with an AUC of 0.902 in 246 patients with HCC at dual-center cohorts.[19] An MRI-based model offers diagnostic accuracy of liver disease (eg, avoids radiation exposure and superior accuracy)[20] and stability in radiomic feature extraction.[42 43] Therefore, diagnostic MRI sequences (including arterial, portal venous, and T2WI phases) were used to develop an optimal deep-learning radiomic model. In this study, the multiphase MRI-based TLR model has preferable performance (AUC=0.85) and generalizability than single-modality models among 660 patients with HCC across four independent multicenter validation cohorts.

Model interpretability remains a significant challenge in radiomics research.[44] Previous studies have demonstrated that a combined clinical and radiomic model can predict the pathological features of patients with HCC and clarify the tumor immune microenvironment and radiomic risk score by combining CT images, bulk RNA sequencing, and single-cell RNA sequencing.[45 46] Our study revealed that the TLR model score correlates with the anticancer microenvironment, inflammatory

response and specific cancer-associated pathways, including memory lymphoid cell, Th1 helper, IL6-JAK-STAT3, and mTORC1 signaling. STAT3 inhibits the accumulation of $CD8^+$ T cells in cancer cells by down-regulating the IFN-γ/CXCR3/CXCL10 axis, thereby helping tumor cells evade the immune system.[47] A previous study demonstrated that the crosstalk of the STAT3 and mTORC1 pathway is linked to enhanced polymorphonuclear myeloid-derived suppressor cell recruitment and tumorigenicity.[48] These correlations align with recent studies on patients with HCC, emphasizing the relevance of tumor immunology in understanding MRI-based model.[11] Histologically confirmed iTLS is associated with favorable prognosis and response to ICI-based treatment.[11 18 25] We observed that a higher TLR model score was associated with improved RFS and OS times in patients with HCC who underwent surgery, and PFS in patients with advanced HCC. We also demonstrated that the TLR-based predicted iTLS+ group has a high proportion (36%) in responders with advanced HCC who have undergone combination therapy compared to the predicted iTLS– group. Overall, this study elucidates the radiological associations between multiphase MRI-based radiomic, the tumor immune microenvironment, and clinical outcomes following combination therapy, supporting the use of non-invasive indicators in patients with advanced HCC. Therefore, the MRI-based non-invasive model can potentially improve treatment modality selection in patients with HCC.

This study has several limitations. First, the manual delineation of tumor lesions is time-consuming and labor-intensive, highlighting the need for semi-automated or automated segmentation systems to facilitate integration into clinical practice. Second, the use of a retrospective cohort introduces the potential for selection bias (eg, most patients with hepatitis virus-related HCC), which may limit the model's applicability to patients with HCC with non-alcoholic fatty liver disease. Prospective studies are required to explore predictive feature differences among patients with HCC with various etiologies. Third, the small sample size for combination therapy response and RNA expression analysis may impact the generalizability of those results. Larger patient populations are required to validate these findings. Finally, the use of individualized but non-standardized combination therapies, determined through multidisciplinary discussions, necessitates additional validation in prospective cohorts for validation. While variability in anti-PD-(L)1 antibodies and combination therapies reflect real-world practices, it introduces internal variability that may affect the robustness of the findings.

In conclusion, the TLRs model based on baseline MRI data effectively predicts iTLS status in HCC. The model's stability and generalizability in the external validation sets demonstrate its potential value for the therapeutic management of patients with HCC.

Author affiliations

[1]Department of Radiology, and National Clinical Research Center for Geriatric Disorders, Xiangya Hospital Central South University, Changsha, Hunan, China
[2]Institute of Molecular Precision Medicine and Hunan Key Laboratory of Molecular Precision Medicine, Xiangya Hospital Central South University Department of General Surgery, Changsha, Hunan, China
[3]Department of Radiology, The Second Affiliated Hospital of Hainan Medical University, Haikou, Hainan, China
[4]Department of Nuclear Medicine, Hainan Cancer Hospital, Haikou, Hainan, China
[5]Department of Oncology, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fuzhou, Fujian, China
[6]Hunan Key Laboratory of Molecular Precision Medicine, Department of General Surgery, Xiangya Hospital & Center for Medical Genetics, School of Life Sciences, Central South University, Changsha, Hunan, China
[7]MOE Key Lab of Rare Pediatric Diseases & Hunan Key Laboratory of Medical Genetics of the School of Life Sciences, Central South University, Changsha, Hunan, China

ORCID iDs

Mengsi Li http://orcid.org/0000-0002-7486-2442
Kai Fu http://orcid.org/0000-0001-8571-5352
Wenzheng Li http://orcid.org/0009-0001-2279-5663

REFERENCES

1 Qin S, Chan SL, Gu S, *et al*. Camrelizumab plus rivoceranib versus sorafenib as first-line therapy for unresectable hepatocellular carcinoma (CARES-310): a randomised, open-label, international phase 3 study. *Lancet* 2023;402:1133–46.
2 Finn RS, Qin S, Ikeda M, *et al*. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N Engl J Med* 2020;382:1894–905.
3 Kelley RK, Rimassa L, Cheng A-L, *et al*. Cabozantinib plus atezolizumab versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol* 2022;23:995–1008.
4 Greten TF, Villanueva A, Korangy F, *et al*. Biomarkers for immunotherapy of hepatocellular carcinoma. *Nat Rev Clin Oncol* 2023;20:780–98.
5 Li H, Ding J-Y, Zhang M-J, *et al*. Tertiary lymphoid structures and cytokines interconnections: The implication in cancer immunotherapy. *Cancer Lett* 2023;568:216293.
6 Helmink BA, Reddy SM, Gao J, *et al*. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature New Biol* 2020;577:549–55.
7 Vanhersecke L, Brunet M, Guégan J-P, *et al*. Mature tertiary lymphoid structures predict immune checkpoint inhibitor efficacy in solid tumors independently of PD-L1 expression. *Nat Cancer* 2021;2:794–802.
8 Ding G-Y, Ma J-Q, Yun J-P, *et al*. Distribution and density of tertiary lymphoid structures predict clinical outcome in intrahepatic cholangiocarcinoma. *J Hepatol* 2022;76:608–18.
9 Schumacher TN, Thommen DS. Tertiary lymphoid structures in cancer. *Science* 2022;375:eabf9419.
10 Zhong K, Xu Y, Cheng Y, *et al*. Case report: Primary hepatocellular carcinoma with portal vein tumor thrombus characterized by active tumor immune microenvironment achieving a complete response following treatment of combined immunotherapy. *Front Immunol* 2022;13:999763.
11 Shu DH, Ho WJ, Kagohara LT, *et al*. Immunotherapy response induces divergent tertiary lymphoid structure morphologies in hepatocellular carcinoma. *Nat Immunol* 2024;25:2110–23.
12 Tapper EB, Lok A-F. Use of Liver Imaging and Biopsy in Clinical Practice. *N Engl J Med* 2017;377:756–68.
13 Lambin P, Leijenaar RTH, Deist TM, *et al*. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–62.
14 Bera K, Braman N, Gupta A, *et al*. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol* 2022;19:132–46.
15 Nitski O, Azhie A, Qazi-Arisar FA, *et al*. Long-term mortality risk stratification of liver transplant recipients: real-time application of deep learning algorithms on longitudinal data. *Lancet Digit Health* 2021;3:e295–305.
16 Li M, Fan Y, You H, *et al*. Dual-Energy CT Deep Learning Radiomics to Predict Macrotrabecular-Massive Hepatocellular Carcinoma. *Radiology* 2023;308:e230255.
17 Chen M, Kong C, Qiao E, *et al*. Multi-algorithms analysis for pre-treatment prediction of response to transarterial chemoembolization in hepatocellular carcinoma on multiphase MRI. *Insights Imaging* 2023;14:38.
18 Li P, Liang Y, Zeng B, *et al*. Preoperative prediction of intra-tumoral tertiary lymphoid structures based on CT in hepatocellular cancer. *Eur J Radiol* 2022;151:110309.
19 Li J, Zhang L, Xing H, *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)* 2024;11(21):e2309348.
20 Di Martino M, De Filippis G, De Santis A, *et al*. Hepatocellular carcinoma in cirrhotic patients: prospective comparison of US, CT and MR imaging. *Eur Radiol* 2013;23:887–96.
21 Lim M, Muquith M, Miramontes B, *et al*. Surrogate and modified endpoints for immunotherapy in advanced hepatocellular carcinoma. *Hepatology* 2023;78:1755–62.
22 Xu B, Dong S-Y, Bai X-L, *et al*. Tumor Radiomic Features on Pretreatment MRI to Predict Response to Lenvatinib plus an Anti-

PD-1 Antibody in Advanced Hepatocellular Carcinoma: A Multicenter Study. *Liver Cancer* 2023;12:262–76.

23 American College of Radiology. Liver reporting & data system. n.d. Available: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS

24 Wang Y, Lin H, Yao N, *et al*. Computerized tertiary lymphoid structures density on H&E-images is a prognostic biomarker in resectable lung adenocarcinoma. *iScience* 2023;26:107635.

25 Calderaro J, Petitprez F, Becht E, *et al*. Intra-tumoral tertiary lymphoid structures are associated with a low risk of early recurrence of hepatocellular carcinoma. *J Hepatol* 2019;70:58–65.

26 Wang Y, Liu Z, Xu H, *et al*. MRI-based radiomics model and nomogram for predicting the outcome of locoregional treatment in patients with hepatocellular carcinoma. *BMC Med Imaging* 2023;23:67.

27 Gong X-Q, Liu N, Tao Y-Y, *et al*. Radiomics models based on multisequence MRI for predicting PD-1/PD-L1 expression in hepatocellular carcinoma. *Sci Rep* 2023;13:7710.

28 Bilal Masokano I, Pei Y, Chen J, *et al*. Development and validation of MRI-based model for the preoperative prediction of macrotrabecular hepatocellular carcinoma subtype. *Insights Imaging* 2022;13:201.

29 Whybra P, Zwanenburg A, Andrearczyk V, *et al*. The Image Biomarker Standardization Initiative: Standardized Convolutional Filters for Reproducible Radiomics and Enhanced Clinical Insights. *Radiology* 2024;310:e231319.

30 Wainberg M, Merico D, Delong A, *et al*. Deep learning in biomedicine. *Nat Biotechnol* 2018;36:829–38.

31 Cabrita R, Lauss M, Sanna A, *et al*. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature New Biol* 2020;577:561–5.

32 Sautès-Fridman C, Petitprez F, Calderaro J, *et al*. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat Rev Cancer* 2019;19:307–25.

33 Teillaud J-L, Houel A, Panouillot M, *et al*. Tertiary lymphoid structures in anticancer immunity. *Nat Rev Cancer* 2024;24:629–46.

34 Wen S, Chen Y, Hu C, *et al*. Combination of Tertiary Lymphoid Structure and Neutrophil-to-Lymphocyte Ratio Predicts Survival in Patients With Hepatocellular Carcinoma. *Front Immunol* 2021;12:788640.

35 Cheng N, Li P, Cheng H, *et al*. Prognostic Value of Tumor-Infiltrating Lymphocytes and Tertiary Lymphoid Structures in Epstein-Barr Virus-Associated and -Negative Gastric Carcinoma. *Front Immunol* 2021;12:692859.

36 Ruffin AT, Cillo AR, Tabib T, *et al*. B cell signatures and tertiary lymphoid structures contribute to outcome in head and neck squamous cell carcinoma. *Nat Commun* 2021;12:3349.

37 Lu Y, Yang A, Quan C, *et al*. A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. *Nat Commun* 2022;13:4594.

38 Xu Y, Li Z, Yang Y, *et al*. Association Between MRI Radiomics and Intratumoral Tertiary Lymphoid Structures in Intrahepatic Cholangiocarcinoma and Its Prognostic Significance. *J Magn Reson Imaging* 2024;60:715–28.

39 Xu Y, Li Z, Yang Y, *et al*. A CT-based radiomics approach to predict intra-tumoral tertiary lymphoid structures and recurrence of intrahepatic cholangiocarcinoma. *Insights Imaging* 2023;14:173.

40 Li K, Ji J, Li S, *et al*. Analysis of the Correlation and Prognostic Significance of Tertiary Lymphoid Structures in Breast Cancer: A Radiomics-Clinical Integration Approach. *Magnetic Resonance Imaging* 2024;59:1206–17.

41 Lu H, Lou H, Wengert G, *et al*. Tumor and local lymphoid tissue interaction determines prognosis in high-grade serous ovarian cancer. *Cell Rep Med* 2023;4:101092.

42 Mackin D, Fave X, Zhang L, *et al*. Measuring Computed Tomography Scanner Variability of Radiomics Features. *Invest Radiol* 2015;50:757–65.

43 Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, *et al*. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. *Radiology* 2018;288:407–15.

44 Tomaszewski MR, Gillies RJ. The Biological Meaning of Radiomic Features. *Radiology* 2021;298:505–16.

45 Feng Z, Li H, Liu Q, *et al*. CT Radiomics to Predict Macrotrabecular-Massive Subtype and Immune Status in Hepatocellular Carcinoma. *Radiology* 2023;307:e221291.

46 Xia T-Y, Zhou Z-H, Meng X-P, *et al*. Predicting Microvascular Invasion in Hepatocellular Carcinoma Using CT-based Radiomics Model. *Radiology* 2023;307:e222729.

47 Wang D, Yu W, Lian J, *et al*. Th17 cells inhibit CD8$^+$ T cell migration by systematically downregulating CXCR3 expression via IL-17A/STAT3 in advanced-stage colorectal cancer patients. *J Hematol Oncol* 2020;13:68.

48 Sun H, Yang W, Tian Y, *et al*. An inflammatory-CCRK circuitry drives mTORC1-dependent metabolic and immunosuppressive reprogramming in obesity-associated hepatocellular carcinoma. *Nat Commun* 2018;9:5214.

49 Collins GS, Moons KGM, Dhiman P, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024;385:e078378.