BMC Cancer

RESEARCH

Open Access

Predicting treatment response and prognosis of immune checkpoint inhibitors-based combination therapy in advanced hepatocellular carcinoma using a longitudinal CT-based radiomics model: a multicenter study

Jun Xu[1,2,3,4], Junjun Li[5], Tengfei Wang[1,2,4*], Xin Luo[6], Zhangxiang Zhu[7], Yimou Wang[1,2], Yong Wang[8], Zhenglin Zhang[1,2], Ruipeng Song[9], Li-Zhuang Yang[1,2,4], Hongzhi Wang[1,2,4], Stephen T. C. Wong[10,11] and Hai Li[1,2,4*]

Abstract

Background Identifying effective predictive strategies to assess the response of immune checkpoint inhibitors (ICIs)-based combination therapy in advanced hepatocellular carcinoma (HCC) is crucial. This study presents a new longitudinal CT-based radiomics model to predict treatment response and prognosis in advanced HCC patients undergoing ICIs-based combination therapy.

Methods Longitudinal CT images were collected before and during the treatment for HCC patients across three institutions from January 2019 to April 2022. A total of 1316 radiomic features were extracted from arterial and portal venous phase abdominal CT images for each patient. A model called Longitudinal Whole-liver CT-based Radiomics (LWCTR) was developed to categorize patients into responders or non-responders using radiomic features and clinical information through support vector machine (SVM) classifiers. The area under the curve (AUC) was used as the performance metric and subsequently applied for risk stratification and prognostic assessment. The Shapley Additive explanations (SHAP) method was used to calculate the Shapley value, which explains the contribution of each feature in the SVM model to the prediction.

Results This study included 395 eligible participants, with a median age of 57 years (IQR 51–66), comprising 344 males and 51 females. The LWCTR model performed well in predicting treatment response, achieving an AUC of 0.883 (95% confidence interval [CI] 0.881–0.888) in the training cohort, 0.876 (0.858–0.895) in the internal validation cohort, and 0.875 (0.860–0.887) in the external test cohort. The Rad-Nomo model, integrating the LWCTR model's prediction score (Rad-score) with the modified Response Evaluation Criteria in Solid Tumors (mRECIST), demonstrated strong prognostic performance. It achieved time-dependent AUC values of 0.902, 0.823, and 0.850 at 1, 2, and 3 years in the internal validation cohort and 0.893, 0.848, and 0.762 at the same intervals in the external test cohort.

*Correspondence:
Tengfei Wang
wangtf@cmpt.ac.cn
Hai Li
hli@cmpt.ac.cn
Full list of author information is available at the end of the article

Conclusion The proposed LWCTR model performs well in predicting treatment response and prognosis in patients with HCC receiving ICIs-based combination therapy, potentially contributing to personalized and timely treatment decisions.

Keywords Immune checkpoint inhibitor, Hepatocellular carcinoma, Radiomics, Machine learning, Computed tomography

Background

Hepatocellular carcinoma (HCC) is the fifth most common cancer globally and ranks third in terms of cancer-related deaths [1]. Immune checkpoint inhibitors(ICIs) targeting programmed death ligand 1 (PD-(L)1) inhibitors, in conjunction with tyrosine kinase inhibitors (TKIs) or vascular endothelial growth factor (VEGF) antibodies, constitute the primary therapeutic strategy for managing advanced HCC. Unfortunately, only approximately 30% of patients experience substantial benefit from this combination therapy [2–4]. The modified Response Evaluation Criteria in Solid Tumors (mRECIST) is the standard tool for assessing radiological endpoints in HCC, using therapy-induced tumor necrosis and dynamic imaging to measure arterial phase contrast changes [5–7]. It is superior to RECIST 1.1 in assessing traditional HCC treatments that focus on local and molecular therapies. However, its effectiveness is limited in capturing unconventional response patterns in immunotherapy [8–10]. Therefore, identifying novel prognostic biomarkers is crucial for precise immunotherapy in advanced HCC [11].

The progress in artificial intelligence (AI) and the development of radiomics have significantly impacted radiology [12, 13]. Radiomics is a rapidly advancing field in medical imaging that uses quantitative, high-throughput analysis of extensive image data to assess tumor characteristics comprehensively. As a non-invasive biomarker, radiomics has been successfully applied in the diagnosis [14], staging [15], prognosis [16], and efficacy prediction [17] of liver cancer. Currently, several studies have explored the potential of radiomics in predicting immune therapy response in HCC [18, 19]. These studies analyzed the correlation between immune therapy response and radiological features of tumor tissue at the same time, demonstrating positive results and potential value.

Despite the demonstrated potential of radiomics in predicting immunotherapy response in HCC, limitations and challenges persist. A primary concern is that most existing studies develop predictive models based on short-term efficacy endpoints, such as mRECIST. This focus has resulted in a relative paucity of models aimed at predicting long-term outcomes, including overall survival (OS) and progression-free survival (PFS), thereby limiting their value in long-term clinical decision-making. Furthermore, contemporary research predominantly relies on baseline CT images, while the considerable potential of longitudinal dynamic imaging data remains largely underutilized. This methodological constraint may limit the precision of efficacy predictions and, consequently, diminish their practical clinical utility [20]. Additionally, the critical roles of non-neoplastic liver tissue and the metabolic microenvironment in shaping immunotherapy response require further elucidation to advance our understanding of the underlying biological mechanisms [21–23]. To ensure the robustness and clinical applicability of radiomics models, multi-center validation is essential for rigorous performance assessment, enhancing generalizability and bolstering reliability for real-world clinical translation.

In this study, we proposed a longitudinal whole-liver CT-based radiomics (LWCTR) model to predict the response to ICIs-based combination therapy in patients with advanced HCC. The LWCTR model extracts radiomic features from pre- and post-treatment whole-liver CT images, integrating clinical data and using OS as the primary endpoint. The generalization of the model was evaluated using independent test datasets from two medical centers. Furthermore, we assessed the added prognostic value of the LWCTR model by comparing its risk stratification capability with that of the mRECIST criteria. Finally, the Shapley Additive Explanations (SHAP) algorithm was used to interpret and visualize the proposed model.

Methods

Study design and participants

This retrospective study analyzed 395 patients with advanced HCC who received ICIs-based combination therapy (ICIs + TKIs or ICIs + anti-VEGF) between January 2019 and April 2022 (Fig. 1). Patients were recruited from three medical centers and assigned to the following cohorts:

Cohort 1: Three hundred fifteen patients from the First Affiliated Hospital of the University of Science and Technology of China (center 1). Patients were randomly divided into training cohort and internal validation cohort at an 8:2 ratio.

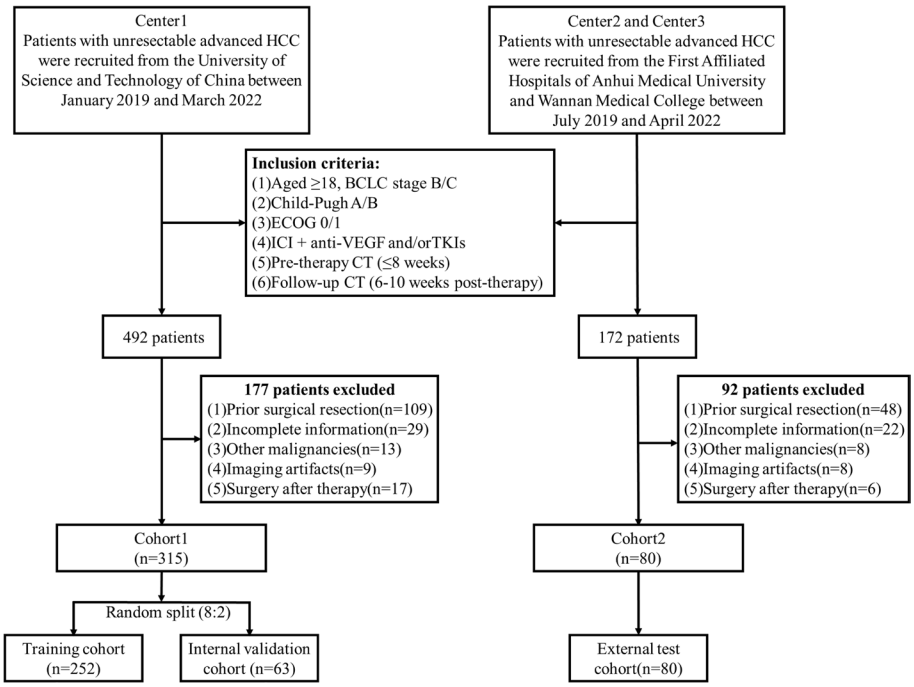

Fig. 1 Flow diagram of the patient selection process from the three participating medical centers of this study

Cohort 2: Eighty patients from the First Affiliated Hospital of Anhui Medical University (center 2) and the First Affiliated Hospital of Wannan Medical College (center 3) constituted the external test cohort.

The study adhered to the principles of the Helsinki Declaration and was approved by the Institutional Ethics Committees (approval numbers: 2023ky197, 2023–510, and 2024–35). Given the retrospective design, the requirement for informed consent was waived.

The diagnosis of HCC is based on established guidelines or criteria [24–26]. The inclusion criteria were as follows: (1) age ≥ 18 years, with histological or clinical confirmation of Barcelona Clinic Liver Cancer (BCLC) stage B or C HCC; (2) Child–Pugh class A or B liver function; (3) Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1; (4) received combined therapy of ICIs + anti-VEGF and/or TKIs; (5) had baseline contrast-enhanced abdominal CT scans within 8 weeks prior to initiating ICIs therapy and at least one follow-up CT scan within 6–10 weeks after treatment initiation. The exclusion criteria were as follows:(1) patient who had a history of surgical resection for liver cancer prior to the start of the study treatment;(2) incomplete clinical or follow-up information;(3) with a history of malignancies other than HCC;(4) imaging artifacts that could affect the accuracy of the assessments;(5) underwent conversion surgery resection after combination therapy.

Administration of ICIs, follow-up and assessments

This study utilizes real-world data and aligns with guidelines from Western countries and China, advocating for ICIs and molecular-targeted agents in HCC treatment. We categorized these drugs by shared targets and validated treatment strategies involving multiple medications, drawing inspiration from umbrella trials that assess drug effectiveness based on specific cancer mutations-a proven approach in numerous trials [27–29]. Three experienced radiologists (J.J.L., Z. Zhu, and Yo. W.), each with over six years of experience at their respective institutions, independently assessed tumor response on initial and subsequent contrast-enhanced CT scans using mRECIST criteria. Specifically, Z. Zhu and Yo. W. independently reviewed each patient's scans. In cases of disagreement, J.J.L., as the adjudicating radiologist, re-evaluated the cases along with the initial assessments to determine the final response, which was considered conclusive. Patient response defined by mRECIST was categorized based on the disease control rate (DCR), classifying patients as responders (complete response [CR], partial response [PR], or stable disease [SD]) or non-responders (progressive disease [PD]). Detailed information regarding ICI administration and assessments is outlined in Additional file 1: Appendix S1 and Table S1. Patient survival status was systematically monitored through telephone follow-ups, outpatient visits, hospital registration systems, and the death registry of

the Anhui Provincial Center for Disease Control and Prevention. Follow-up was conducted until May 31, 2023, or until the patient's demise.

In oncology clinical trials, OS is esteemed as a highly dependable and clinically relevant endpoint, defined as the time from ICI treatment initiation to death from any cause. In this study, OS was the primary endpoint, with PFS—the interval from ICI initiation to disease progression (per mRECIST criteria) or death—as the secondary endpoint. Treatment response was assessed by comparing individual survival times to the median OS of Cohort 1 (12.0 months). This criterion was validated as a reliable method by Lin et al. [30] for predicting treatment outcomes in metastatic colorectal cancer.

CT Study protocols and clinical characteristics

We analyzed arterial and portal venous phase abdominal CT images of patients scanned using multi-row CT scanners from different manufacturers. Detailed imaging protocols are provided in Additional file 1: Appendix S2. Subject demographics and baseline clinical characteristics of eligible patients, including biological variables such as age, sex, etiology, cirrhosis, Child–Pugh class, Eastern Cooperative Oncology Group performance status(ECOG), alpha-fetoprotein, macroscopic portal vein, extrahepatic spread, among others, were collected from the Hospital Information System. In the training set, the Least Absolute Shrinkage and Selection Operator (LASSO) method with tenfold cross-validation was applied to identify the clinical features most relevant to treatment response. Subsequently, these attributes were incorporated into the development of the clinical model.

Liver segmentation, radiomics feature extraction and selection

The whole liver was semi-automatically segmented in all eligible patients' baseline and first treatment response assessment CT scans (6–10 weeks). The nnU-Net model, [31] a pre-trained liver segmentation model, was initially used to automatically segment the liver region in the arterial and portal venous phase CT images. Reader 1 (J.X), an expert with eight years of abdominal imaging experience, assessed and manually refined the predicted liver mask, especially for poor covers, utilizing the ITK-SNAP software (3·8·0). Display parameters were set to a 200 Hounsfield Unit (HU) range and a 45 HU center to enhance diagnostic clarity(see Additional file 1:Figure S1). To determine the feature stability, reader 1 (JX) performed two repeat segmentations within one month based on the nnU-Net automatic segmentation. Reader 2 (J. J. L.), with ten years of abdominal imaging experience, independently delineated the segmentations of 50 randomly selected patients. Details of liver segmentation, radiomics feature extraction, and feature selection are provided in Additional file 1:Appendix S3.

Model development

We utilized support vector machine (SVM) [32] classifiers with a linear kernel, integrating baseline clinical characteristics and radiomic features extracted from multiple time points to predict patients' treatment responses, specifically their 12-month OS probability. Seven predictive models were developed: (1) Clinical Model: Based solely on baseline clinical features. (2) Radiomic-Pre model: Incorporated pre-treatment arterial and portal venous phase CT radiomic features. (3) Radiomic-Post Model: Utilized post-treatment arterial and portal venous phase CT radiomic features. (4) Radiomic-LA Model: Integrated pre- and post-treatment arterial phase CT radiomic features. (5) Radiomic-LV Model: Combined pre- and post-treatment portal venous phase CT radiomic features. (6) Radiomic-Fusion Model: Merged pre- and post-treatment arterial and portal venous phase CT radiomic features. (7) LWCTR Model: Incorporated baseline clinical features along with fused radiomic features from the Radiomic-Fusion Model. Details of the features included in each model are provided in Additional file 1: Table S3. Model performance was optimized through a rigorous approach encompassing grid search and five-fold cross-validation within the training cohort, followed by internal validation. Generalizability was assessed via subsequent external validation on an independent test cohort. Prediction scores, ranging from 0 to 1, indicated prognosis: lower scores suggested poorer outcomes, while higher scores indicated increased survival likelihood.

To enhance prognostic accuracy, we constructed three distinct prognostic models using Cox proportional hazards regression: the Rad-score model, derived from the LWCTR model prediction score (Rad-score); the mRECIST model; and the Rad-Nomo model, a nomogram integrating the Rad-score and the mRECIST model.

Model explanation and visualization

SHAP is a technique employed to interpret predictions generated by machine learning models. It is rooted in the concept of Shapley values from game theory [33–35]. Our study leveraged a SHAP kernel explainer to determine feature contributions. By adopting this methodology, we aim to understand better how individual elements affect the compound classification task. The process depicting the development and validation of the model is demonstrated in Fig. 2.

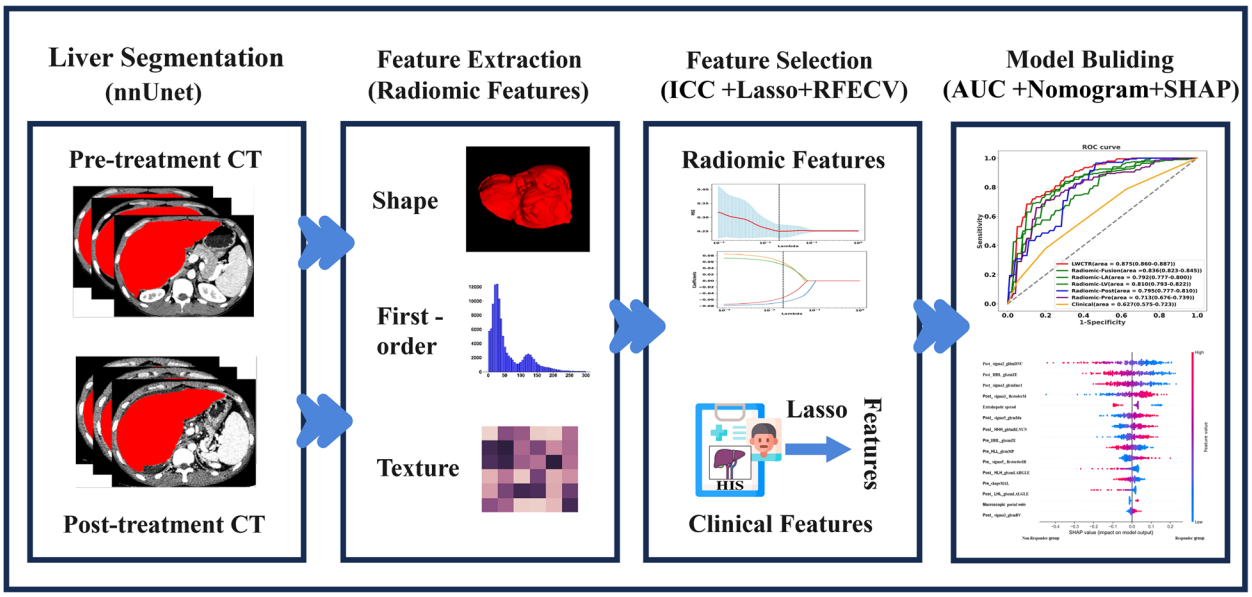

Fig. 2 An illustration of the workflow of the Longitudinal Whole-liver CT-based Radiomics Model. ICC = intraclass correlation coefficient, Lasso = the least absolute shrinkage and selection operator, RFECV = recursive feature elimination with cross-validation, AUC = area under curve, SHAP = Shapley Additive Explanations

Statistical analysis

To evaluate model performance and clinical utility, we conducted a series of statistical analyses. Receiver operating characteristic (ROC) curve analysis quantified model performance, using the area under the curve (AUC) and 95% confidence intervals (CIs) to measure predictive accuracy. Statistically significant differences in AUC values were assessed via the DeLong test. Calibration curves compared predicted probabilities to observed outcomes, while decision curve analysis (DCA) evaluated net benefit and clinical applicability across threshold probabilities [36].

Risk stratification was conducted by categorizing patients into high-risk and low-risk groups based on prediction scores from the top-performing models. The optimal threshold for classification was determined using the ROC Youden index [37]. Kaplan–Meier survival analysis was used to estimate and compare OS and PFS between risk groups, and survival differences were assessed using log-rank tests. Cox proportional hazards regression estimated hazard ratios (HR) and 95% CIs for key predictive factors. Time-dependent ROC analysis further evaluated the Rad-Nomo model's predictive performance for PFS (at 6, 12, and 18 months) and OS (at 1, 2, and 3 years).

For statistical significance testing, continuous variables were presented as medians with interquartile ranges (IQRs) and compared using either the unpaired two-sample t-test or the Mann–Whitney U test. Categorical variables were analyzed using either the chi-squared test or Fisher's exact test, depending on sample size constraints. A p-value of < 0.05 was considered statistically significant for all analyses. Statistical analysis was conducted using Python version 3·7·3 and R version 4·3·0 (https://www.r-project.org) by J.X. and T.W.

Results

Patient characteristics

395 eligible patients participated, with a median age of 57.6 years (IQR, 51.0–66.0 years; 344 males and 51 females). The leading cause of HCC was hepatitis B virus (HBV) infection, comprising 76.2% (301/395) of the patients. Among the included patients, 302 (76.5%) had liver cirrhosis. The median diameter of liver tumors was 8.4 cm (IQR: 5.0–11.0 cm), and multiple liver tumors were detected in 149 patients (37.5%) according to the up-to-seven criteria. During follow-up, mortality rates were 75.4% (190/252) in the training cohort, 68.3% (43/63) in the internal validation cohort, and 70.0% (56/80) in the external test cohort. Of the 99 deaths recorded across the internal validation and external test cohorts combined, 93 (93.9%) were attributed to HCC. Median OS was 12.0 months in both the training and external test cohorts, and 13.0 months in the internal validation cohort. Median PFS was 6.1 months, 6.6 months, and 6.7 months in the training, internal validation, and external test cohorts, respectively. Log-rank tests indicated no significant differences in OS ($p = 0.198$ and 0.336) and PFS ($p = 0.209$ and 0.257) between the

training cohort and the other cohorts (see Additional File 1: Figure S2). In the training cohort, 192 patients (76.2%) received ICIs + TKIs, while 60 (23.8%) received ICIs + anti-VEGF therapy. In the internal validation and external test cohorts, 85 patients (59.4%) received ICIs + TKIs, while 58 (40.6%) received ICIs + anti-VEGF therapy. Baseline characteristics were comparable across all cohorts (see Table 1).

Radiomics features and clinical characteristics selection

In the training set, four sets of radiological features were extracted from whole liver segmentation using pre- and post-treatment CT images (arterial and portal venous phases). Each set comprised 1,316 features, including 1,050 texture features, 252 first-order features, and 14 shape features (see Additional file 1: Table S2). Ultimately, we selected five features from baseline arterial phase CT, four features from baseline portal venous phase CT, and nine features from post-treatment arterial and portal venous phase CT images, as shown in Table 2. Clinical factors such as portal vein invasion and extrahepatic spread were identified as significant influencing factors. Additional file 1: Figure S3 displays the LASSO coefficients for feature selection using tenfold cross-validation at different time points. Subsequently, Spearman correlation coefficients for all feature pairs were computed (see Additional file 1: Figure S4), without excluding any feature pairs.

Performance of LWCTR model in assessing ICIs based combination therapy response

To evaluate the LWCTR model for predicting treatment response in ICIs-based combination therapies, we systematically assessed its performance by comparing it with the mRECIST criteria and six SVM-based models. As shown in Fig. 3A–C, the LWCTR model demonstrated significant advantages across all datasets, achieving AUC of 0.883 (95% CI: 0.881–0.888) in the training cohort, 0.876 (95% CI: 0.858–0.895) in the internal validation cohort, and 0.875 (95% CI: 0.860–0.887) in the external test cohort. The model consistently outperformed all comparison groups ($p < 0.05$, DeLong test), with detailed performance metrics provided in Table S4. Notably, compared to the mRECIST criteria in the external test cohort (AUC = 0.649, 95% CI: 0.568–0.734; Fig. 3C, Table 3), the LWCTR model exhibited a 34.8% improvement in predictive performance (ΔAUC = 0.226), highlighting its potential clinical applicability. In subgroup analyses of different treatment strategies, the LWCTR model maintained robust predictive accuracy, achieving AUC of 0.869 (95% CI: 0.856–0.881) in the ICIs + TKIs cohort ($n = 85$) and 0.905 (95% CI: 0.887–0.932) in the ICIs + anti-VEGF cohort ($n = 58$). These values closely aligned with those of the overall validation cohort (test & validation cohort, $n = 143$, AUC = 0.882, 95% CI: 0.870–0.896), with a ΔAUC $\leq$ 0.023, indicating strong generalizability across treatment regimens. Furthermore, to specifically evaluate the model's performance on cancer-specific survival (CSS) endpoints, a CSS-adjusted cohort was constructed, comprising 93.9% of deceased patients from the combined internal validation and external test cohorts, whose mortality was definitively attributed to HCC. In the CSS-adjusted cohort ($n = 137$), the LWCTR model maintained an AUC of 0.883 (95% CI: 0.870–0.897), demonstrating strong consistency with the AUC of 0.882 based on OS in the overall validation cohort (Fig. 3D). The LWCTR model's calibration curves aligned predicted and observed treatment responses, with all Brier scores < 0.25 (see Additional file 1: Figure S5). Furthermore, the Decision Curve Analysis (DCA) illustrated that the LWCTR model offered superior overall benefits across various threshold probabilities compared to other radiomics methods (see Additional file 1: Figure S6).

Risk stratification and prognostic accuracy

In the training cohort, the LWCTR model's prediction score (Rad-score) achieved an AUC of 0.883 for distinguishing treatment responders from non-responders, with the optimal threshold set at 'Rad-score $\geq$ 0.52' (see Additional file 1: Figure S7). Employing this threshold in both internal validation and external test cohorts revealed substantial differences in OS when classifying patients into rad response and rad non-response groups (Fig. 4A, B). Kaplan–Meier analysis indicated that although mRECIST can stratify patients by risk, its ability to identify patients in their early treatment stages is limited. Only ten patients were categorized as non-responders in the internal validation cohort, while in the external test cohort, only 16 patients were identified as non-responders (Fig. 4C, D). Per the Rad-score criteria, a higher percentage of non-responders was identified in the internal validation cohort (46.0% vs 15.9%) and the external test cohort (51.3% vs 20.0%) compared to mRECIST. Additionally, compared to the mRECIST criteria, the Rad-score criteria exhibited superior performance, showing a more significant delta median OS, with 19.0 vs. 9.0 (Δ month) and 17.0 vs. 8.0 (Δ month) in the internal validation and external test cohort, respectively. Additionally, the Rad-score criteria demonstrated robust risk stratification for PFS in both the internal validation and external test cohorts (Fig. 4E, F).

Figure 5A and B present the Rad-Nomo model's survival estimation performance at 1, 2, and 3 years within the internal validation and external test cohorts. (Nomograms for the other two models are available in Additional file 1: Figure S8). Time-dependent ROC curves

Table 1 Characteristics of the training, internal validation, and external test cohorts

Characteristics	Training cohort ($n=252$)	Internal validation cohort ($n=63$)	p value *	External test cohort ($n=80$)	p value *
Age, years	57 [51–66]	57 [50·5–66·5]	0.508	56 [52–65]	0.584
Sex (%)			0.825		0.566
Male	215 (85.3)	54 (85.7)		75 (93.7)	
Female	37 (14.7)	9 (14.3)		5 (6.3)	
Etiology			0.384		0.566
Hepatitis B virus	190 (75.4)	52 (83.9)		59 (73.7)	
Others	62 (24.6)	10 (16.1)		21 (26.3)	
Cirrhosis			1.000		1.000
Yes	195 (77.4)	47 (74.6)		60 (75.0)	
No	57 (22.6)	16 (25.4)		20 (25.0)	
Child–Pugh class			0.383		0.727
A	224 (88.9)	48 (76.2)		63 (78.8)	
B	28 (11.1)	15 (23.8)		17 (21.2)	
ECOG PS			1.000		0.233
0	161 (63.9)	41 (65.1)		52 (65.0)	
1	91 (36.1)	22 (34.9)		28 (35.0)	
BCLC stage			1.000		0.593
B	43 (17.1)	9 (14.3)		9 (11.2)	
C	209 (82.9)	54 (85.7)		71 (88.8)	
AFP (ng/ml)			0.583		0.904
< 20	69 (27.4)	18 (28.6)		20 (25.0)	
20–400	79 (31.3)	20 (31.7)		24 (30.0)	
> 400	104 (41.3)	25 (39.7)		36 (45.0)	
Up-to-seven criteria			0.853		0.96
≤ 7	158 (62.7)	37 (58.7)		59 (63.7)	
> 7	94 (37.3)	26 (41.3)		29 (36.3)	
Tumor size (cm)	8.55 [5.6–11.4]	7.5 [4.5–12.0]	0.134	7.0 [5.0–11.0]	0.119
Macroscopic portal vein invasion			0.579		0.453
Absent	75 (29.8)	15 (23.8)		22 (27.5)	
Present	177 (70.2)	48 (76.2)		58 (72.5)	
Extrahepatic spread			0.293		0.796
Absent	150 (59.5)	42 (66.7)		52 (65.0)	
Present	102 (40.5)	21 (33.3)		28 (35.0)	
Types of ICIs (n, %)					
Camrelizumab	174 (69.0)	34 (54.0)	1.000	42 (52.5)	0.436
Sintilimab	42 (16.7)	15 (23.8)	1.000	25 (31.3)	0.463
Tislelizumab	12 (4.8)	6 (9.5)	1.000	7 (8.8)	1.000
Toripalimab	14 (5.6)	3 (4.8)	1.000	1 (1.3)	1.000
Pembrolizumab	4 (1.6)	2 (3.2)	1.000	2 (2.5)	1.000
Atezolizumab	6 (2.4)	3 (4.8)	1.000	3 (3.8)	1.000
Prior HCC treatment (n, %)					
TACE	122 (48.4)	29 (46.0)	1.000	31 (38.8)	0.440
Ablative therapy	12 (4.8)	3 (4.8)	1.000	3 (3.8)	0.662
Radiotherapy	14 (5.6)	4 (6.3)	1.000	4 (5.0)	1.000
Portal vein stent	1 (0.4)	1 (1.6)	1.000	1 (1.3)	1.000
TKI therapy	72 (28.6)	17 (27.0)	0.350	21 (26.2)	1.000
Duration of follow up (months)	25 [19–29]	20 [17–29]	0.615	26 [20–30]	0.398
Median OS time(months)	12 [7–24]	13 [9–28]	0.198	12 [8–27]	0.209
Median PFS time(months)	6.1 [3.3–14.0]	6.6 [3.6–20.3]	0.336	6.7 [14.0–18.9]	0.257

Table 1 (continued)

Data are median (interquartile range) or n (%). ECOG PS Eastern Cooperative Oncology Group performance status, *BCLC* Barcelona Clinic Liver Cancer, *AFP* alpha-fetoprotein, *ICIs* immune checkpoint inhibitors, *HCC* hepatocellular carcinoma, *TACE* trans-arterial chemoembolization, *TKI* tyrosine kinase inhibitor, *OS* overall survival, *PFS* progression-free survival

* Indicates the comparison to the training cohort

Table 2 Radiomic features selected in the pre-treatment and post-treatment CT images

CT images	Number	Feature name	Abbreviation
Pre _treatment_A	5	original_gldm_GrayLevelNonUniformity	Pre_A_ gldmGLNU
		wavelet-HHH_gldm_SmallDependenceLowGrayLevelEmphasis	Pre_A_ HHH_gldmSDLGLE
		original_shape_MinorAxisLength	Pre_A_shapeMAL
		wavelet-LLH_gldm_DependenceNonUniformity	Pre_A_LLH_ gldmDNU
		wavelet-HLL_firstorder_Median	Pre_A_HLL_firstorderM
Pre _treatment_V	4	original_shape_MinorAxisLength	Pre_V_shapeMAL
		log-sigma-5-mm-3D_firstorder_InterquartileRange	Pre_ V_sigma5_ firstorderIR
		wavelet-HLL_glcm_MaximumProbability	Pre_V_HLL_glcmMP
		wavelet-HHL_glszm_ZoneEntropy	Pre_V_HHL_glszmZE
Post _treatment_A	9	wavelet-LLH_firstorder_RootMeanSquared	Post_ A_ LLH_ firstorderRMS
		wavelet-HHH_glrlm_ShortRunLowGrayLevelEmphasis	Post_ A_HHH_ glrlmSRLGLE
		log-sigma-4-mm-3D_gldm_DependenceNonUniformityNormalized	Post_ A_ sigma4_ gldmDNUN
		wavelet-LLH_glszm_ZoneVariance	Post_ A_ LLH_ glszmZV
		wavelet-LLL_glszm_GrayLevelNonUniformity	Post_ A_LLL_glsmGLNU
		wavelet-HLL_glszm_LargeAreaLowGrayLevelEmphasis	Post_ A_HLL_glszmLALGLE
		log-sigma-2-mm-3D_gldm_DependenceNonUniformity	Post_ A_ sigma2_gldmDNU
		log-sigma-1-mm-3D_ngtdm_Coarseness	Post_ A_ sigma1_ ngtdmC
		log-sigma-3-mm-3D_glszm_LargeAreaEmphasis	Post_ A_ sigma3_ glszmLAE
Post _treatment_V	9	log-sigma-2-mm-3D_gldm_DependenceNonUniformity	Post_ V_sigma2_gldmDNU
		log-sigma-3-mm-3D_firstorder_Median	Post_V_ sigma3_ firstoderM
		log-sigma-3-mm-3D_glcm_Imc1	Post_V_ sigma3_glcmImc1
		log-sigma-3-mm-3D_glrlm_RunVariance	Post_V_ sigma3_glcmRV
		log-sigma-5-mm-3D_glcm_Idn	Post_V_ sigma5_glcmIdn
		wavelet-LHL_glszm_LargeAreaLowGrayLevelEmphasis	Post_V_ LHL_glszmLALGLE
		wavelet-HLH_glszm_LargeAreaHighGrayLevelEmphasis	Post_V_ HLH_glszmLAHGLE
		wavelet-HHL_glszm_ZoneEntropy	Post_V_ HHL_glszmZE
		wavelet-HHH_glrlm_RunLengthNonUniformityNormalized	Post_V_ HHH_glrlmRLNUN

A arterial phase images, *V* portal venou s phase images, *3D* 3 dimensional, *glcm* gray-level co-occurrence matrix, *glszm* gray-level size zone matrix, *gldm* gray level dependence matrix, *glrlm* gray level run length matrix, *ngtdm* neighbouring gray tone difference matrix

demonstrated that the Rad-Nomo model provided the best prediction of patient mortality at 1-year, 2-year, and 3-year OS in the internal validation and external test cohorts (Fig. 5C, D), and for PFS at 6, 12, and 18 months (Figure S9).

Explanation and visualization of LWCTR model

To gain deeper insight into the interpretability of the LWCTR model, we conducted an in-depth analysis of the SHAP values associated with each feature. As shown in Fig. 6, the three most significant features for predicting treatment response, as determined by the SHAP, were Post_V_sigma2_gldmDNU, Post_A_HHH_glrlmSRLGLE, and Post_V_sigma3_firstoderM. The first two features are texture features that offer a quantitative description of the liver's internal structure, heterogeneity, and complexity, encompassing the tumor. The third feature is a first-order feature that characterizes the median grayscale values within the volume.

A force plot (Fig. 7) demonstrates patient-specific prediction probabilities for four representative HCC patients. Features potentially improving treatment response are highlighted in red, while those potentially influencing response negatively are displayed in blue. Notably, eight of the 395 patients in our multicenter

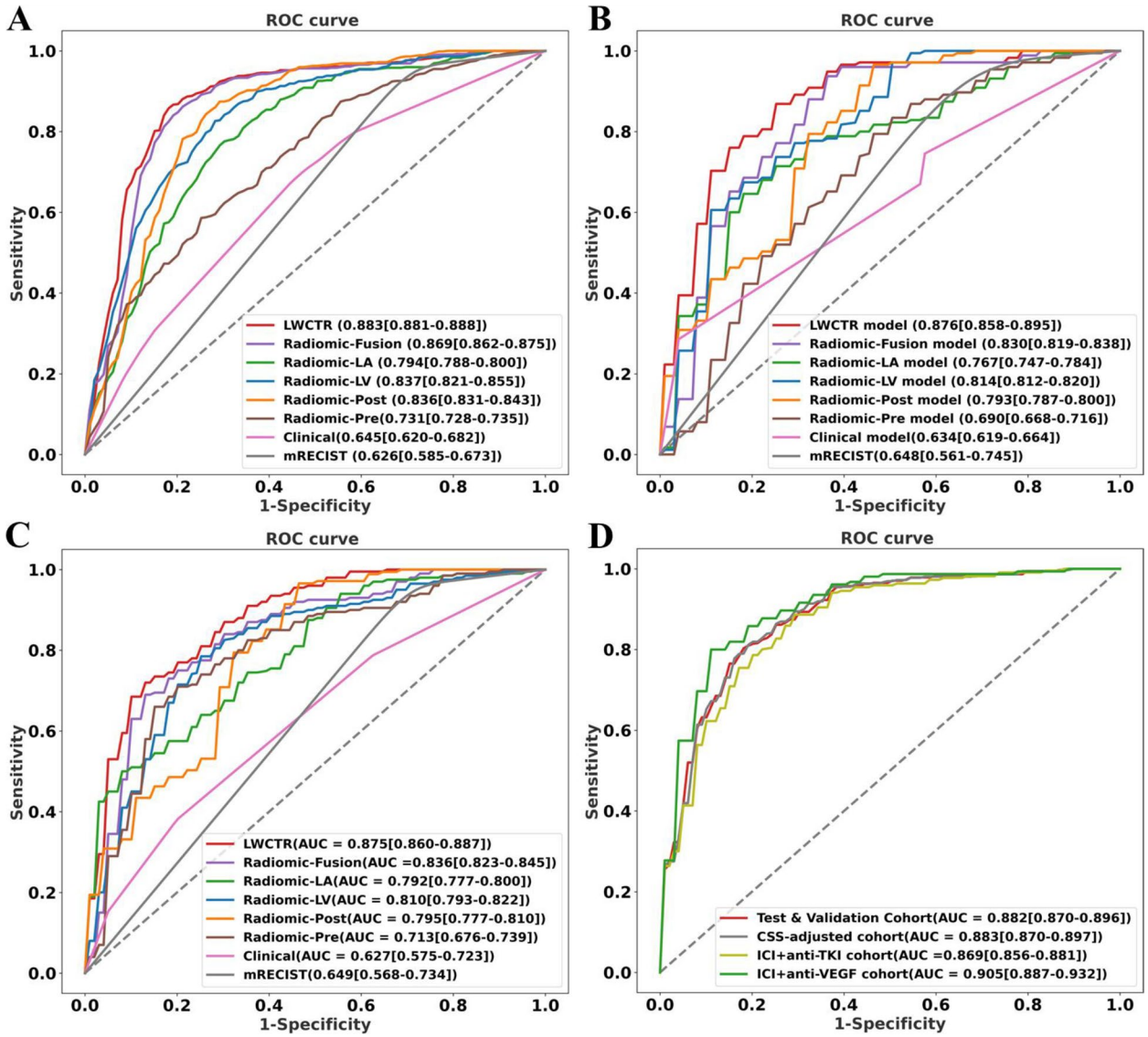

Fig. 3 Performance evaluation of the LWCTR model compared with mRECIST criteria and other predictive models using ROC analysis. **A**–**C** Receiver operating characteristic (ROC) curves for the training cohort (**A**), internal validation cohort (**B**), and external test cohort (**C**). **D** ROC curves showing the performance of the LWCTR model across different subgroups

Table 3 Response assessment according to mRECIST criteria

Tumor response	Training cohort ($n=252$)	Internal validation cohort ($n=63$)	External test cohort ($n=80$)
CR	5	1	1
PR	49	13	18
SD	157	39	45
PD	41	10	16
ORR (CR + PR)	21.4%	22.2%	23.8%
DCR (CR + PR + SD)	83.7%	84.1%	80.0%

CR Complete response, *PR* Partial response, *SD* Stable disease, *PD* Progressive disease, *ORR* Objective response rate, *DCR* Disease control rate

study experienced pseudoprogression, with our model accurately identifying six (75%).

Discussion

This study introduces the LWCTR model, a novel approach for predicting the response to ICI-based combination therapy in advanced HCC. The LWCTR model integrates multimodal features, including clinical data and radiomic characteristics from pre- and post-treatment whole-liver CT images, enabling a comprehensive assessment of tumor phenotype and treatment response. Compared to mRECIST and other predictive models,

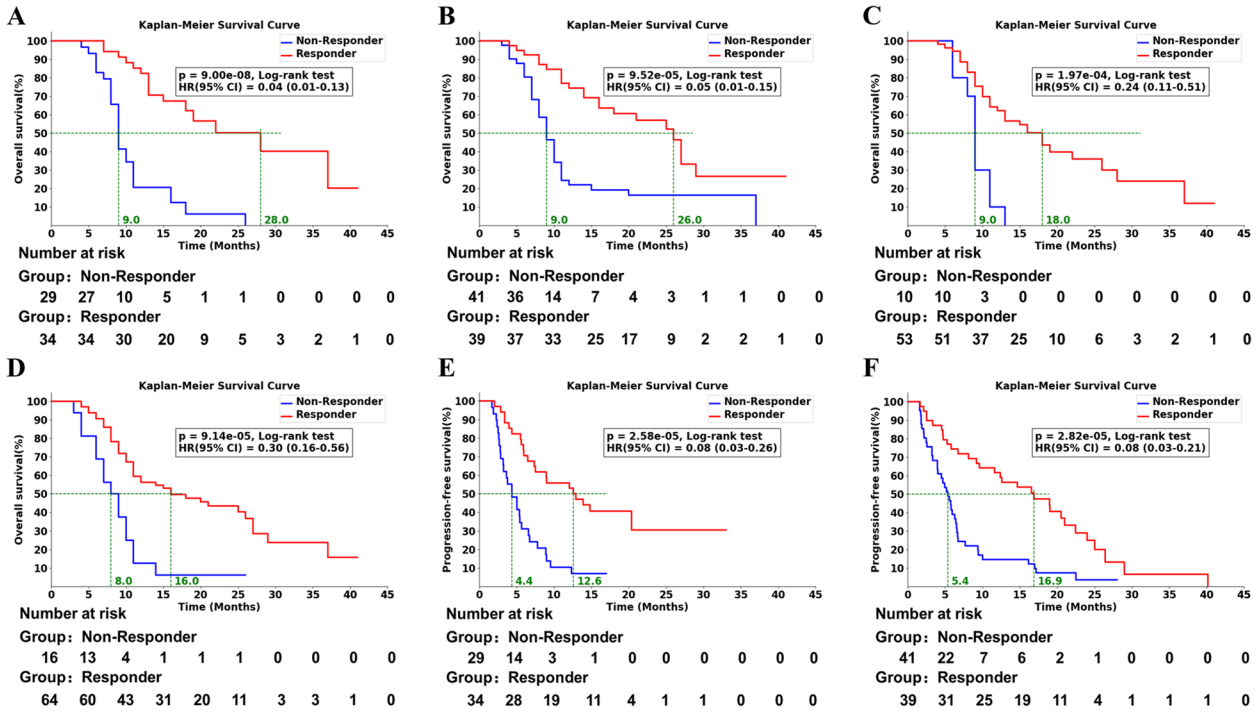

Fig. 4 Survival analysis comparing Rad-score and mRECIST criteria. Kaplan–Meier curves depict overall survival (OS) and progression-free survival (PFS) stratified by Rad-score and mRECIST criteria. Hazard ratios (HR) with 95% confidence intervals (CIs) and log-rank p-values are provided. **A**, **B** OS curves for the internal validation cohort (**A**) and external test cohort (**B**) based on Rad-score. **C**, **D** OS curves for the internal validation cohort (**C**) and external test cohort (**D**) based on mRECIST. **E**, **F** PFS curves for the internal validation cohort (**E**) and external test cohort (**F**) based on Rad-score. Responders are defined as Rad-score ≥ 0.52 or disease control (CR/PR/SD) by mRECIST

the LWCTR model demonstrates superior predictive performance across multiple validation cohorts while eliminating the need for invasive procedures. A Rad-score threshold of ≥ 0.52 indicates a higher likelihood of prolonged survival. Moreover, incorporating the LWCTR prediction score with mRECIST in the Rad-Nomo model further enhances prognostic accuracy.

This study is motivated by the challenges in assessing HCC responses to immunotherapy using conventional criteria such as mRECIST. A key limitation is the occurrence of paradoxical treatment effects, such as pseudoprogression, which may lead to erroneous clinical interpretations and premature discontinuation of effective therapies [8, 9]. Additionally, mRECIST relies primarily on lesion morphology and CT-based radiographic progression as diagnostic benchmarks, potentially underestimating true biological responses and delaying timely therapeutic decision-making [38]. Radiomics, by integrating comprehensive imaging features, offers a promising approach for more accurate tumor evaluation. While radiomics has shown promise in predicting treatment response and prognosis for HCC patients receiving ICIs, current approaches have limitations. Many studies focus on predicting short-term efficacy using models based on mRECIST or RECIST rather than OS, the primary endpoint in advanced HCC clinical trials. Recent evidence indicates that mRECIST has limited predictive value for OS, and PFS assessed by RECIST/mRECIST demonstrates only a moderate correlation with OS (r = 0.30–0.45) [11]. For instance, Hua et al. [39] developed a pretreatment CT-based radiomics-clinical model using mRECIST criteria in 150 patients with HCC to predict response to lenvatinib-ICIs combination therapy. Although the model demonstrated strong predictive performance for treatment response (AUC = 0.892), it did not achieve statistically significant prognostic stratification for OS (p = 0.062), indicating its limitations in predicting long-term survival. Similarly, Hin et al. [40] constructed the Liver Cancer Radiomics Signature (LCRS) using unsupervised clustering of CT radiomic features from multicenter cohorts. In 137 patients with unresectable HCC receiving ICIs-based combination therapy, the LCRS effectively predicted treatment response per RECIST criteria (AUC = 0.806) and stratified PFS risk. However, the study did not further explore the predictive value of LCRS for OS, leaving its potential role in long-term outcome prediction unexplored.

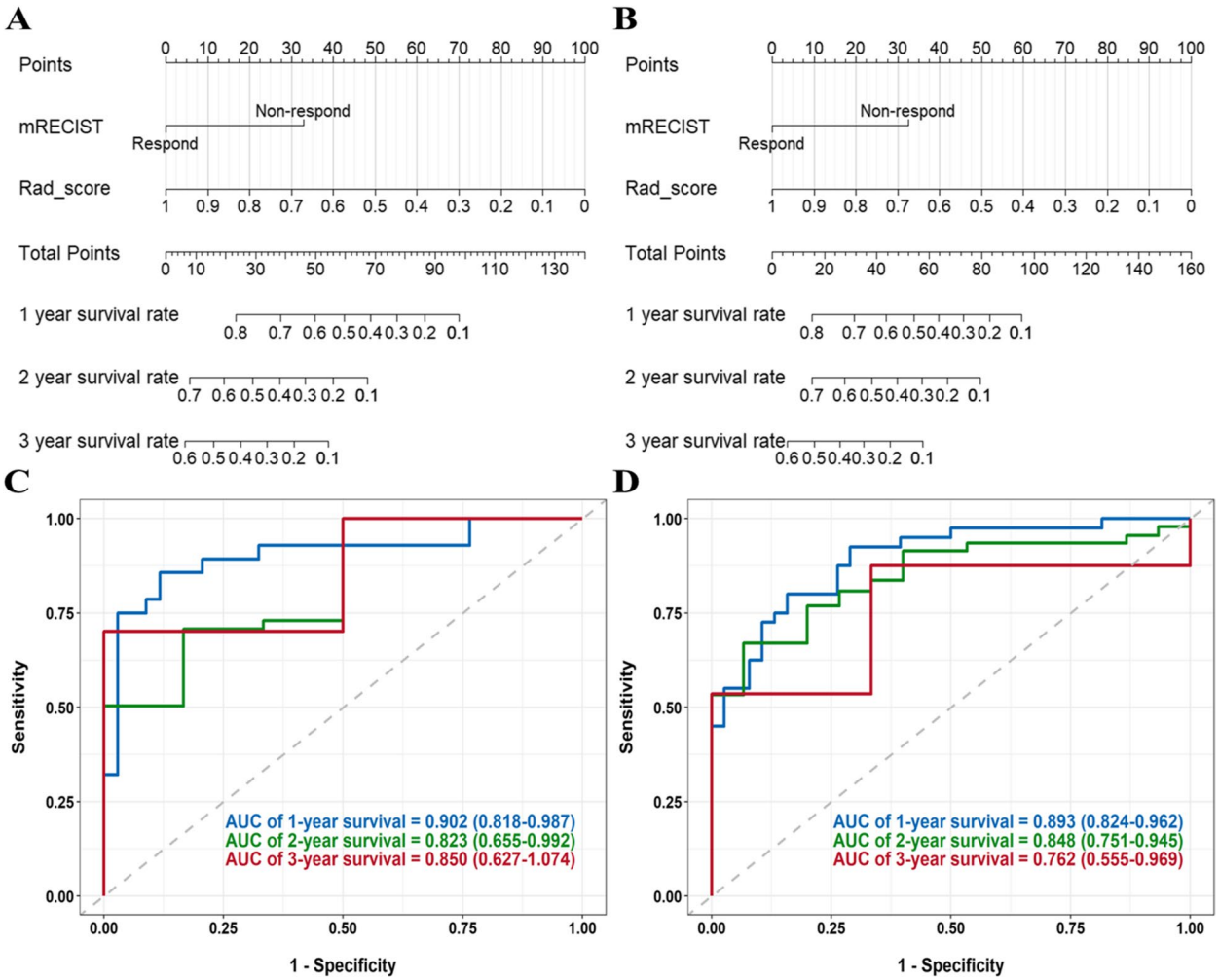

Fig. 5 Rad-Nomo nomogram and the time-dependent ROC curves in both the internal (**A**, **C**) and external (**B**, **D**) validation cohorts

In contrast to previous radiomics frameworks that primarily focused on surrogate endpoints, Our study developed the LWCTR model with OS as the primary endpoint, demonstrating robust performance in predicting early treatment responses and showing potential for detecting pseudoprogression. Using the Youden index, we determined the optimal threshold and established Rad-score criteria to classify patients as responders or non-responders, facilitating personalized follow-up. As a complement to the mRECIST model, the LWCTR model improves early response assessment and long-term prognosis estimation. Their integration in the Rad-Nomo model further enhances predictive performance, providing valuable insights beyond mRECIST alone. For example, in cases of ambiguous mRECIST findings, such as suspected pseudoprogression, the LWCTR model can aid in discerning accurate treatment responses. This additional information may prevent premature cessation of effective therapies, support the continuation of beneficial treatments, or prompt timely adjustments to treatment strategies. Moreover, integrating both assessments allows for more tailored follow-up protocols. Patients identified as high-risk through the integrated assessment might benefit from more frequent monitoring, while those with clearly positive responses could have their follow-up streamlined to support more individualized treatment decisions and optimize patient management.

In phase III clinical trials for HCC, OS is widely adopted as the primary endpoint for evaluating treatment efficacy in patients with unresectable advanced HCC due to its clinical reliability and broad applicability [4, 41–43]. Although CSS offers a theoretical advantage as a prognostic measure by excluding non-tumor-related mortality, OS maintains precedence in advanced HCC research. This sustained preference reflects the clinical reality that, within this patient population, the majority of deaths are directly attributable to HCC-related mortality, with a low proportion of non-HCC-related mortality.

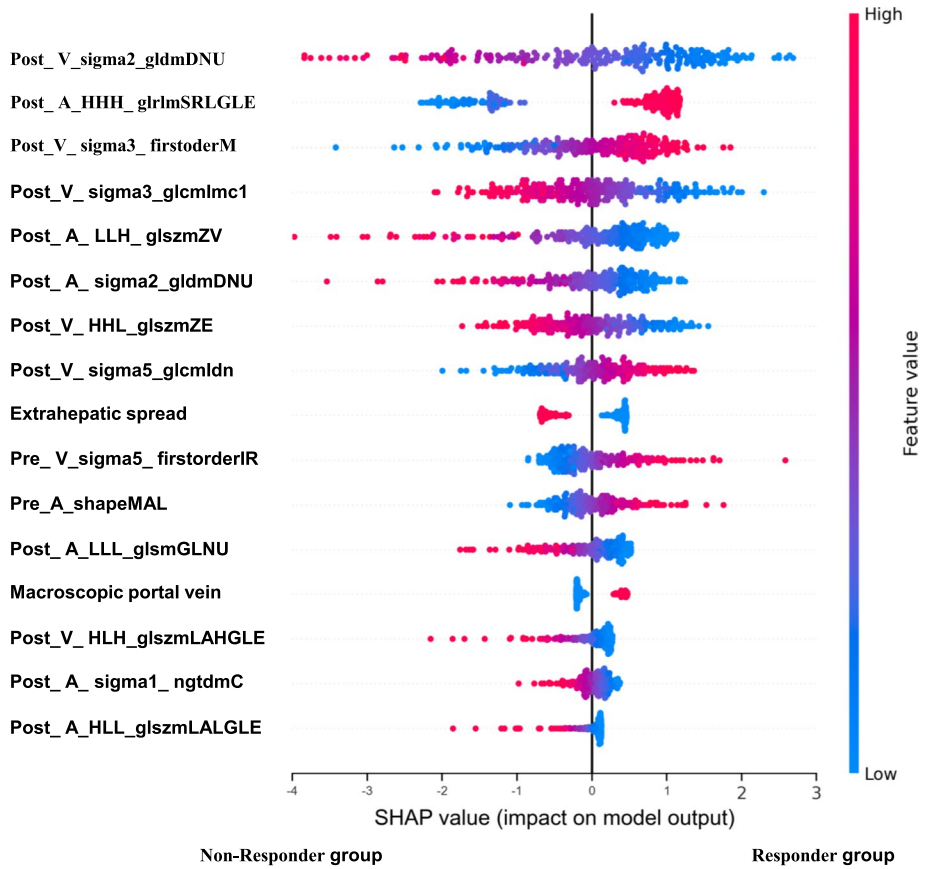

Fig. 6 SHAP summary plots for the LWCTR model illustrate feature relevance and combined feature attributions impacting predictive performance

Therefore, OS and CSS trends are usually highly consistent. A prospective study of 1080 HCC patients showed that after a median follow-up of 3.6 years, 601 patients died, of whom 552 (92%) were attributed to HCC progression [44]. Similarly, another study involving 967 HCC patients, with a median follow-up of 1017 days, reported 393 deaths, of which 360 (92%) were clearly due to HCC progression. The proportion of advanced HCC in both studies was approximately 50% [45]. In our study, a total of 99 deaths were recorded in the internal validation and external test cohorts, of which 93 (93.9%) were attributed to HCC, consistent with the above findings. To assess the applicability of the model at the CSS endpoint, we constructed a CSS-adjusted cohort. The results showed that the AUC of the LWCTR model in the CSS-adjusted cohort was 0.883, which was highly consistent with the AUC of the OS-based overall validation cohort (0.882), indicating that the model's predictive ability remained robust when using CSS as the endpoint. It is important to note that the implementation of surgical intervention can significantly affect the distribution of HCC-related and non-related mortality [46]. Therefore, in our study design, we excluded patients who underwent radical surgery before and after treatment, which effectively reduced the potential difference between OS and CSS assessment results.

To enhance the clinical applicability of our model, we utilized the SHAP algorithm to interpret the model. In our target model, the most critical feature contributing to the model's prediction output is Post_ V_sigma2_gldmDNU. This radiomics feature quantifies the degree of irregularity or heterogeneity in the spatial dependence of voxel intensities within the tumor and liver. A higher value of dependence nonuniformity indicates a more anomaly in the spatial dependence of voxel intensities, suggesting a more heterogeneous texture within the cancer. Specifically, the recruitment of peripherally activated T cells into the tumor microenvironment (TME) to eliminate tumor cells is considered the primary mechanism of immunotherapy [47]. The pathological features of HCC, such as the abundance of immature supplying arteries and draining veins, result in characteristic "arterial hyperenhancement" and "washout" features observed on CT scans. These characteristics may hinder T-cell infiltration into TME. Contrast-enhanced CT images displayed attenuation contrasts between the

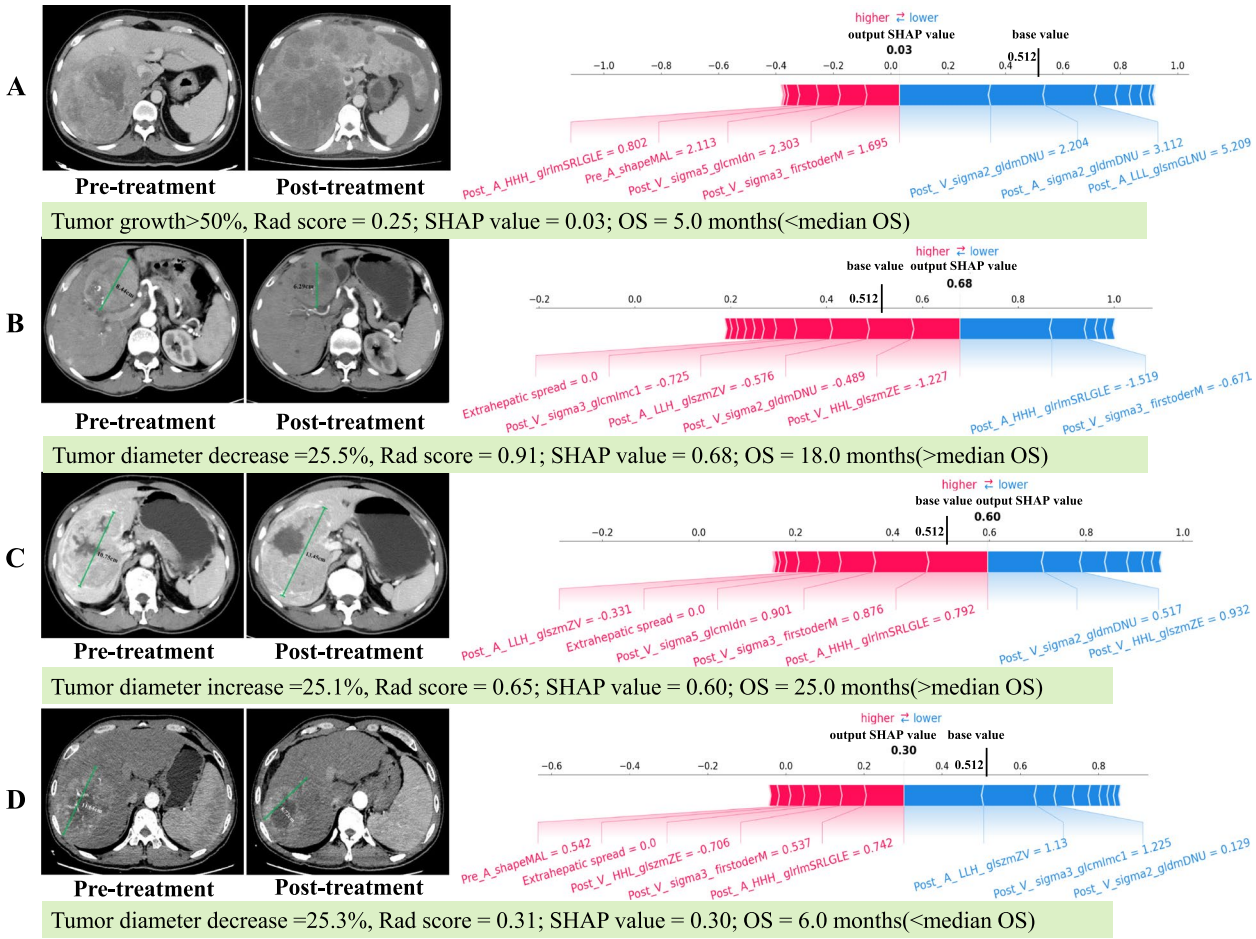

Fig. 7 The SHAP force plots revealed how the LWCTR model distinguished the treatment responses among four patients. Patient **A** had a negative response (non-responder) and experienced rapid tumor growth (hyperprogression). Patient **B** responded positively (responder) with a successful treatment outcome. Patient **C** also showed a positive response (responder) but exhibited signs of treatment-related tumor growth (pseudoprogression). Patient **D**, on the other hand, revealed a negative response (non-responder) and displayed an atypical response pattern. Although there was an initial reduction in tumor burden during the first assessment, the patient had a notably short survival time. Portal vein invasion: 0 = Presence of portal vein tumor thrombus; 1 = Absence of portal vein tumor thrombus. Extrahepatic spread: 0 = Absence of extrahepatic spread; 1 = Presence of extrahepatic spread

tumor and liver tissues, signifying substantial tumor heterogeneity and higher values of dependence nonuniformity [48]. Our findings suggest that patients with lower Post_ V_sigma2_gldmDNU values on post-treatment CT scans after receiving ICIs-based combination therapy showed improved responses to immunotherapy and more prolonged survival. That result may be attributed to the normalization of the tumor vasculature induced by combination therapy, leading to an enhanced portal vein blood supply to the tumor. Increased blood retention within the tumor microenvironment provides more opportunities for T cell migration into the tumor stroma, thereby improving the antitumor immune response. This multi-center study pioneers using a longitudinal CT-based radiomics model to predict the treatment response and prognosis of ICIs-based combination therapy in advanced HCC.

Our study had several limitations. First, the LWCTR model was developed using OS as the primary endpoint. Although OS and CSS are highly concordant in advanced HCC patients, OS may still be affected by non-HCC-related mortality. Future research will expand the sample size, integrate CSS data, and investigate OS versus CSS differences across HCC subgroups. Additionally, we plan to incorporate multi-omics data (e.g., genomics and radiomics) and apply deep learning techniques to construct a more precise and biologically interpretable multidimensional predictive model. Second, unlike the conventional tumor-based radiomic model that solely captures tumor information, the LWCTR model integrates tumor and

global liver information. Although the whole-liver analysis yielded favorable results, the most salient predictive value likely resides within features directly derived from the tumor itself. Combining whole-liver contextual features with tumor-specific radiomic approaches may yield optimal performance, warranting further investigation. Third, its retrospective nature may introduce inherent biases, as most enrolled patients had HCC related to viral etiology. Further validation is crucial to extend the applicability of this model to NAFLD patients. Finally, using CT scans from different centers may result in variations in radiomic features. To mitigate scanner variability, we standardized feature extraction through resampling and normalizing slice thickness, enhancing model reproducibility and robustness, which could facilitate broader utilization.

Conclusions

In conclusion, our study has shown that noninvasive tracking based on whole-liver analysis with longitudinal CT images using the LWCTR model holds promise for predicting treatment response and prognosis in advanced HCC. Although further research is required to establish the model's generalizability, noninvasive whole-liver tracking shows promise in guiding adaptive and personalized immunotherapy strategies for HCC, providing valuable insights for clinicians.

Abbreviations

HCC	Hepatocellular carcinoma
ICIs	Immune checkpoint inhibitors
OS	Overall survival
mRECIST	Modified Response Evaluation Criteria in Solid Tumors
AUC	Area under the receiver operating characteristic curve
CI	Confidence interval
ROC	Receiver operating characteristic
OS	Overall survival
PFS	Progression-free survival
CSS	Cancer-specific survival
LASSO	Least absolute shrinkage and selection operator
CT	Computed tomography
DCR	Disease control rate
ORR	Objective response rate
CR	Complete response
PD	Progressive disease
PR	Partial response
SD	Stable disease
SHAP	Shapley Additive Explanations
SVM	Support vector machine

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13978-4.

Additional file 1.

Acknowledgements

The authors acknowledge the First Affiliated Hospital of the University of Science and Technology of China, the First Affiliated Hospital of Anhui Medical University, and the First Affiliated Hospital of Wannan Medical College (Yijishan Hospital of Wannan Medical College) for their help and support in this project.

Authors' contributions

Conception and design: H.L., T.W., L-Z.Y., X.L., H.W., T.C.W., and J.X. Collection and curation of the clinical datasets: J.X., J.L., Z.Zhu., Yo.W., and R.S. Image processing and programming: J.X., T.W., and Z.Zhang. Data analysis and interpretation: J.X., J.L., T.W., Yi.W., and Z.Zhang. Data and algorithms quality control: H.L., L-Z.Y, X.L., R.S., and T.C.W. Drafting of the manuscript: J.X., J.L., and T.W. Critical analysis and manuscript revision: all authors. All authors read and approved the final manuscript.

Funding

This work was supported by the National Key R&D Program of China (2017YFB1300204), Anhui Provincial Key Research and Development Project (2023s07020001), the National Natural Science Foundation of China (82371931), John S Dunn Research Foundation and T. T. and W. F. Chao Foundation.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This study has obtained approval from the First Affiliated Hospital of the University of Science and Technology of China(Approval No.: 2023-ky197); the First Affiliated Hospital of Wannan Medical College (Approval No.: 2024–35); the First Affiliated Hospital of Anhui Medical University (Approval No.: 2023–510). The requirement to obtain informed consent was waived by the Ethics Committee of the First Affiliated Hospital of the University of Science and Technology of China, the First Affiliated Hospital of Wannan Medical College; the First Affiliated Hospital of Anhui Medical University because of the retrospective nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Anhui Province Key Laboratory of Medical Physics and Technology, Institute of Health and Medical Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, People's Republic of China. [2]University of Science and Technology of China, Hefei 230026, People's Republic of China. [3]Department of Intervention, The First Affiliated Hospital of University of Science and Technology of China, Hefei 230001, People's Republic of China. [4]Department of Oncology, Hefei Cancer Hospital, Chinese Academy of Sciences, Hefei 230031, People's Republic of China. [5]Department of Radiology, The First Affiliated Hospital of University of Science and Technology of China, Hefei 230001, People's Republic of China. [6]Yangtze Delta Region Institute (Huzhou) & School of Resources and Environment, University of Electronic Science and Technology of China, Huzhou, Chengdu 313099, 611731, China. [7]Department of Radiology, The First Affiliated Hospital of Anhui Medical University, Hefei 230022, People's Republic of China. [8]Department of Radiology, The First Affiliated Hospital of Wannan Medical College (Yijishan Hospital of Wannan Medical College), Wuhu 241001, People's Republic of China. [9]Department of Hepatobiliary Surgerydivision of Life Sciences and Medicineanhui Province Key Laboratory of Hepatopancreatobiliary Surgery, Anhui Provincial Clinical Research Center for Hepatobiliary Diseases, The First Affiliated Hospital of USTC, the University of Science and Technology of China, Hefei 230001, People's Republic of China. [10]Department of Systems Medicine and Bioengineering, Houston Methodist Cancer Center, Houston Methodist Hospital, Houston, TX 77030, USA. [11]Department of Radiology, Weill Cornell Medical College, New York, NY 10065, United States.

Received: 7 February 2024 Accepted: 19 March 2025
Published online: 03 April 2025

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Finn RS, Ikeda M, Zhu AX, Sung MW, Baron AD, Kudo M, et al. Phase Ib study of lenvatinib plus pembrolizumab in patients with unresectable hepatocellular carcinoma. J Clin Oncol. 2020;38(26):2960–70.
3. Ren Z, Xu J, Bai Y, Xu A, Cang S, Du C, et al. Sintilimab plus a bevacizumab biosimilar (IBI305) versus sorafenib in unresectable hepatocellular carcinoma (ORIENT-32): a randomised, open-label, phase 2–3 study. Lancet Oncol. 2021;22(7):977–90.
4. Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. N Engl J Med. 2020;382(20):1894–905.
5. Llovet JM, Lencioni R. mRECIST for HCC: performance and novel refinements. J Hepatol. 2020;72(2):288–306.
6. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30(1):52–60.
7. Litiere S, Isaac G, De Vries EGE, Bogaerts J, Chen A, Dancey J, et al. RECIST 1.1 for response evaluation apply not only to chemotherapy-treated patients but also to targeted cancer agents: a pooled database analysis. J Clin Oncol. 2019;37(13):1102–10.
8. Dercle L, Sun S, Seban RD, Mekki A, Sun R, Tselikas L, et al. Emerging and evolving concepts in cancer immunotherapy imaging. Radiology. 2023;306(1):32–46.
9. Zhou L, Shao L, Gao S, Cui C, Chi Z, Sheng X, et al. Impact of response patterns for patients with advanced acral melanoma treated with anti-programmed death-1 monotherapy. Br J Dermatol. 2023;188(1):112–21.
10. Yu H, Bai Y, Xie X, Feng Y, Yang Y, Zhu Q. RECIST 1.1 versus mRECIST for assessment of tumour response to molecular targeted therapies and disease outcomes in patients with hepatocellular carcinoma: a systematic review and meta-analysis. BMJ Open. 2022;12(6): e052294.
11. Lim M, Muquith M, Miramontes B, Lee CJ, Espinoza M, Huang YH, Hsiehchen D. Surrogate and modified endpoints for immunotherapy in advanced hepatocellular carcinoma. Hepatology. 2023;78(6):1755–62.
12. Xia TY, Zhou ZH, Meng XP, Zha JH, Yu Q, Wang WL, et al. Predicting microvascular invasion in hepatocellular carcinoma using CT-based radiomics model. Radiology. 2023;307(4): e222729.
13. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500–10.
14. Gao R, Zhao S, Aishanjiang K, Cai H, Wei T, Zhang Y, et al. Deep learning for differential diagnosis of malignant hepatic tumors based on multi-phase contrast-enhanced CT and clinical data. J Hematol Oncol. 2021;14(1):154.
15. Wu M, Tan H, Gao F, Hai J, Ning P, Chen J, et al. Predicting the grade of hepatocellular carcinoma based on non-contrast-enhanced MRI radiomics signature. Eur Radiol. 2019;29(6):2802–11.
16. Yu Y, Fan Y, Wang X, Zhu M, Hu M, Shi C, Hu C. Gd-EOB-DTPA-enhanced MRI radiomics to predict vessels encapsulating tumor clusters (VETC) and patient prognosis in hepatocellular carcinoma. Eur Radiol. 2022;32(2):959–70.
17. Aujay G, Etchegaray C, Blanc JF, Lapuyade B, Papadopoulos P, Pey MA, et al. Comparison of MRI-based response criteria and radiomics for the prediction of early response to transarterial radioembolization in patients with hepatocellular carcinoma. Diagn Interv Imaging. 2022;103(7–8):360–6.
18. Cui H, Zeng L, Li R, Li Q, Hong C, Zhu H, et al. Radiomics signature based on CECT for non-invasive prediction of response to anti-PD-1 therapy in patients with hepatocellular carcinoma. Clin Radiol. 2023;78(2):e37–44.
19. Yuan G, Song Y, Li Q, Hu X, Zang M, Dai W, et al. Development and validation of a contrast-enhanced CT-based radiomics nomogram for prediction of therapeutic efficacy of anti-PD-1 antibodies in advanced HCC patients. Front Immunol. 2020;11: 613946.
20. Kim H, Park CM, Keam B, Park SJ, Kim M, Kim TM, et al. The prognostic value of CT radiomic features for patients with pulmonary adenocarcinoma treated with EGFR tyrosine kinase inhibitors. PLoS ONE. 2017;12(11): e0187500.
21. Liu Y, Xun Z, Ma K, Liang S, Li X, Zhou S, et al. Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. J Hepatol. 2023;78(4):770–82.
22. Carroll HK, Duffy AG, O'Farrelly C. Liver immunology, immunotherapy, and liver cancers: time for a rethink? Semin Liver Dis. 2022;42(2):212–24.
23. Lui TKL, Cheung KS, Leung WK. Machine learning models in the prediction of 1-year mortality in patients with advanced hepatocellular cancer on immunotherapy: a proof-of-concept study. Hepatol Int. 2022;16(4):879–91.
24. Vogel A, Meyer T, Sapisochin G, Salem R, Saborowski A. Hepatocellular carcinoma. Lancet. 2022;400(10360):1345–62.
25. Zhou J, Sun H, Wang Z, Cong W, Wang J, Zeng M, et al. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (2019 edition). Liver Cancer. 2020;9(6):682–720.
26. Reig M, Forner A, Rimola J, Ferrer-Fabrega J, Burrel M, Garcia-Criado A, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. J Hepatol. 2022;76(3):681–93.
27. Zhu HD, Li HL, Huang MS, Yang WZ, Yin GW, Zhong BY, et al. Transarterial chemoembolization with PD-(L)1 inhibitors plus molecular targeted therapies for hepatocellular carcinoma (CHANCE001). Signal Transduct Target Ther. 2023;8(1):58.
28. Papadimitrakopoulou V, Lee JJ, Wistuba II, Tsao AS, Fossella FV, Kalhor N, et al. The BATTLE-2 study: a biomarker-integrated targeted therapy study in previously treated patients with advanced non-small-cell lung cancer. J Clin Oncol. 2016;34(30):3638–47.
29. Menis J, Hasan B, Besse B. New clinical research strategies in thoracic oncology: clinical trial design, adaptive, basket and umbrella trials, new end-points and new evaluations of response. Eur Respir Rev. 2014;23(133):367–78.
30. Lu L, Dercle L, Zhao B, Schwartz LH. Deep learning for the prediction of early on-treatment response in metastatic colorectal cancer from serial medical imaging. Nat Commun. 2021;12(1):6654.
31. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18(2):203–11.
32. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intelligent Systems and their Applications. 1998;13(4):18–28.
33. Chen H, Lundberg SM, Lee SI. Explaining a series of models by propagating Shapley values. Nat Commun. 2022;13(1):4512.
34. Zhang R, Hong M, Cai H, Liang Y, Chen X, Liu Z, et al. Predicting the pathological invasiveness in patients with a solitary pulmonary nodule via Shapley additive explanations interpretation of a tree-based machine learning radiomics model: a multicenter study. Quant Imaging Med Surg. 2023;13(12):7828–41.
35. Wang Y, Lang J, Zuo JZ, Dong Y, Hu Z, Xu X, et al. The radiomic-clinical model using the SHAP method for assessing the treatment response of whole-brain radiotherapy: a multicentric study. Eur Radiol. 2022;32(12):8737–47.
36. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565–74.
37. Wj Y. Index for rating diagnostic tests. Cancer. 1950;3:32–5.
38. Chen D, Liu J, Zang L, Xiao T, Zhang X, Li Z, et al. Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. Int J Biol Sci. 2022;18(1):360–73.
39. Hua Y, Sun Z, Xiao Y, Li H, Ma X, Luo X, et al. Pretreatment CT-based machine learning radiomics model predicts response in unresectable hepatocellular carcinoma treated with lenvatinib plus PD-1 inhibitors and interventional therapy. J Immunother Cancer. 2024;12(7):e008953.
40. Xin H, Lai Q, Liu Y, Liao N, Wang Y, Liao B, et al. Integrative radiomics analyses identify universal signature for predicting prognosis and therapeutic vulnerabilities across primary and secondary liver cancers: a multi-cohort study. Pharmacol Res. 2024;210: 107535.
41. Llovet JM, Kudo M, Merle P, Meyer T, Qin S, Ikeda M, et al. Lenvatinib plus pembrolizumab versus lenvatinib plus placebo for advanced hepatocellular carcinoma (LEAP-002): a randomised, double-blind, phase 3 trial. Lancet Oncol. 2023;24(12):1399–410.
42. Abou-Alfa GK, Lau G, Kudo M, Chan SL, Kelley RK, Furuse J, et al. Tremelimumab plus durvalumab in unresectable hepatocellular carcinoma. NEJM Evidence. 2022;1(8):EVIDoa2100070.

43. Qin S, Chan SL, Gu S, Bai Y, Ren Z, Lin X, 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(10408):1133–46.
44. Wusiman M, Huang SY, Liu ZY, He TT, Fang AP, Li MC, et al. Serum S-adenosylhomocysteine, rather than homocysteine, is associated with hepatocellular carcinoma survival: a prospective cohort study. Am J Clin Nutr. 2024;120(3):481–90.
45. Shu J, Zhang M, Dong X, Long J, Li Y, Tan P, et al. Vitamin D receptor gene polymorphisms, bioavailable 25-hydroxyvitamin D, and hepatocellular carcinoma survival. J Natl Cancer Inst. 2024;116(10):1687–96.
46. Yen YH, Kee KM, Li WF, Liu YW, Wang CC, Hu TH, et al. Causes of death among patients with hepatocellular carcinoma according to chronic liver disease etiology. Cancers (Basel). 2023;15(6):1687.
47. Yost KECH, Satpathy AT. Recruiting T cells in cancer immunotherapy. Science. 2021;372(6538):130–1.
48. Feng Y, Zhang H, Ren Q, Li C, Liu S, Zheng C, Xia X. Contrast-enhanced CT parameters predict short-term tumor response in patients with hepatocellular carcinoma who received sequential combined anti-angiogenesis and immune checkpoint inhibitor treatment. Eur J Radiol. 2023;162:110784.

Publisher's Note