BMC Cancer

RESEARCH Open Access

Prediction of prognosis of immune checkpoint inhibitors combined with anti-angiogenic agents for unresectable hepatocellular carcinoma by machine learning-based radiomics

Xuni Xu[1,3†], Xue Jiang[2,3†], Haoran Jiang[3†], Xiaoye Yuan[3], Mengjing Zhao[4], Yuqi Wang[3], Gang Chen[5*], Gang Li[3*] and Yuxia Duan[4*]

Abstract

Objectives This study aims to develop and validate a novel radiomics model utilizing magnetic resonance imaging (MRI) to predict progression-free survival (PFS) in patients with unresectable hepatocellular carcinoma (uHCC) who are receiving a combination of immune checkpoint inhibitors (ICIs) and antiangiogenic agents. This is an area that has not been previously explored using MRI-based radiomics.

Methods 111 patients with uHCC were enrolled in this study. After performing univariate cox regression and the least absolute shrinkage and selection operator (LASSO) algorithms to extract radiological features, the Rad-score was calculated through a Cox proportional hazards regression model and a random survival forest (RSF) model. The optimal calculation method was selected by comparing the Harrell's concordance index (C-index) values. The Rad-score was then combined with independent clinical risk factors to create a nomogram. C-index, time-dependent receiver operating characteristics (ROC) curves, calibration curves, and decision curve analysis were employed to assess the forecast ability of the risk models.

Results The combined nomogram incorporated independent clinical factors and Rad-score calculated by RSF demonstrated better prognosis prediction for PFS, with C-index of 0.846, 0.845, separately in the training and the validation cohorts. This indicates that our model performs well and has the potential to enable more precise patient

†Xuni Xu, Xue Jiang and Haoran Jiang contributed equally to this work.

*Correspondence:
Gang Chen
chen.gang@wmu.edu.cn
Gang Li
andrewlee0923@wzhospital.cn
Yuxia Duan
doughter1@163.com

Full list of author information is available at the end of the article

stratification and personalized treatment strategies. Based on the risk level, the participants were classified into two distinct groups: the high-risk signature (HRS) group and the low-risk signature (LRS) group, with a significant difference between the groups ($P < 0.01$).

Conclusion The effective clinical-radiomics nomogram based on MRI imaging is a promising tool in predicting the prognosis in uHCC patients receiving ICIs combined with anti-angiogenic agents, potentially leading to more effective clinical outcomes.

Keywords Hepatocellular carcinoma, Radiomics, Progression-free survival, Immune checkpoint inhibitors, Anti-angiogenic agents

Introduction

Hepatocellular carcinoma (HCC) was the sixth most commonly observed cancer and the third most common cause of cancer-related deaths worldwide in 2020 [1]. Numerous patients suffer from late-stage disease as a result of delayed identification. The median overall survival is typically ranges from only 6 to 8 months [2].

Recently, systematic treatments for patients with advanced HCC include tyrosine kinase inhibitors (TKIs) and immune checkpoint inhibitors (ICIs). TKIs including sorafenib, regorafenib, lenvatinib, bosutinib, and ramucirumab have been used as first-line and second-line treatments for HCC patients [3–8]. Additionally, ICIs, which encompass the inhibitors of programmed death 1 (PD-1) and programmed death ligand 1 (PD-L1), have exhibited remarkable therapeutic advantages for patients with advanced HCC [9]. PD-1 monoclonal antibodies, such as nivolumab, pembrolizumab, camrelizumab, and sintilimab, have received approval for their use in advanced HCC following successful clinical trials [10–13]. Studies have demonstrated this combination therapy has a synergistic effect to improve overall survival (OS) and PFS [14–16]. However, some patients may not experience benefits from the combination of TKIs and ICI therapy, which is also associated with significant adverse events, including gastrointestinal hemorrhage, pneumonia, sepsis, and cardiac arrest. This highlights the urgent need for novel biomarkers to inform clinical decision-making [15, 17].

Radiomics is an emerging field of image analysis that transforms images into quantifiable data, which can be analyzed to construct models that support decision-making [18]. This approach can serve as a guideline for the diagnosis, staging, treatment planning, prediction of treatment response, and determination of patient prognosis in HCC [19, 20]. Dong et al. conducted a study where they used a machine learning (ML) algorithm on noninvasive CT imaging to determine the effectiveness of TKI plus anti-PD-1 antibodies treatment in advanced HCC patients [21]. Using radiomics based on MRI is a valuable method for forecasting the reaction to transarterial chemoembolization (TACE) in HCC [22, 23]. It has also been shown to be effective in assessing the effect of combined targeted therapy [24]. PFS has historically been the most commonly used surrogate endpoint in late-stage clinical trials, and early assessment of PFS is a robust surrogate endpoint for OS in immunotherapy trials for HCC [25, 26]. However, there is currently no published research focusing on the use of MRI-based radiomics to prognosticating progression-free survival (PFS) among patients diagnosed with uHCC.

As an emerging technology, artificial intelligence (AI) radiomics has profoundly changed the field of imaging analysis. It has important application in diagnosis, individualized treatment and prognosis of HCC. ML is a classical technique of AI used to generate predictive models that are widely used in HCC research [27]. In this study, the objective is to develop an MRI-based radiomics model using machine learning ML. We also hypothesize that a comprehensive model, which combines radiomic features with clinical features, will enhance the accuracy of predicting PFS in patients with uHCC.

Methods and materials

Patient selection

From September 2019 to August 2022, 111 patients diagnosed with uHCC according to the national comprehensive cancer network guidelines [28] were enrolled in the study. These patients were treated with ICIs and anti-angiogenic agents at the First Affiliated Hospital of Wenzhou Medical University. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the First Affiliated Hospital of Wenzhou Medical University. Given that this investigation was retrospective, and involved regular examinations utilizing MRI and the analysis of clinical data, obtaining individual informed consent was not considered mandatory. The eligibility criteria for inclusion in this study were as follows: (1) individuals who had been diagnosed with HCC either clinically or pathologically; (2) patients who received complete treatment with PD-1 inhibitors such as toripalimab (*Coherus Biosciences*, *China*), sintilimab (*Innovent Biologics*, *China*), or camrelizumab (*Hengrui Medicine*, *China*), along with anti-angiogenic agents therapy using sorafenib (*Bayer Pharma AG*, *Germany*), lenvatinib (*Eisai Co.*, *Ltd.*,

Japan), or regorafenib (*Bayer Pharma AG*, *Germany*); (3) patients who underwent enhanced MRI before starting the treatment to obtain imaging data; (4) patients who had comprehensive clinical and treatment information accessible. The exclusion criteria for this study are as follows: (1) patients who underwent alternative anticancer therapies, like interventional therapy, were not included; (2) patients diagnosed with other forms of cancer, like cholangiocarcinoma, were not included; (3) patients with liver function classified as Child-Pugh class C were also not included. (4) patients with MRI data of low quality that was challenging to visualize or quantify were not included. The patients were allocated randomly into two cohorts (at a ratio of 2:1): training cohort ($n = 74$) and validation cohort ($n = 37$). The flowchart of patient enrollment was illustrated in Fig. 1.

Treatment

The treatment protocol was determined by an experienced multidisciplinary team. The therapeutic regiments included the administration of ICIs and TKIs. Toripalimab, at a dose of 3 mg/kg body weight or a fixed amount of 240 mg, was intravenously given every two weeks. Camrelizumab and Sintilimab, at a fixed dose of 200 mg, were intravenously administered every three weeks. Sorafenib, given orally, was provided at a fixed dosage of 400–800 mg per day. For patients with a body weight less than 60 kg, Lenvatinib was orally administered at a fixed dose of 8 mg per day, while patients with a body weight of 60 kg or more received a dosage of 12 mg per day. Regorafenib, taken orally, was administered at a fixed dose of 160 mg per day, every 3 weeks within a 4-week cycle.

Follow-up and clinical endpoint

The subjects underwent regular monthly follow-up appointments throughout the initial year via the outpatient service. Subsequently, they were reassessed every 3–6 months. The endpoint of the study was to determine the PFS of patients. PFS refers to the duration between the introduction of ICIs and TKIs and the occurrence of disease progression or death. Disease progression is determined by RECIST guideline (version 1.1) [29]. The deadline for observation was established until March 30, 2023. The mean follow-up time for both the training and validation cohorts was 6 months.

MRI data acquisition

The images, compliant with DICOM standards, were obtained from the Picture Archiving and Communication System (PACS) records. Acquiring the images involved the utilization of an 8-channel phased-array body coil-equipped Siemens 3.0 T Prisma MR scanner from Siemens Medical Solutions in Erlangen, Germany. The imaging protocol included T1-weighted (T1w), T2-weighted (T2w), diffusion weighted imaging (DWI),

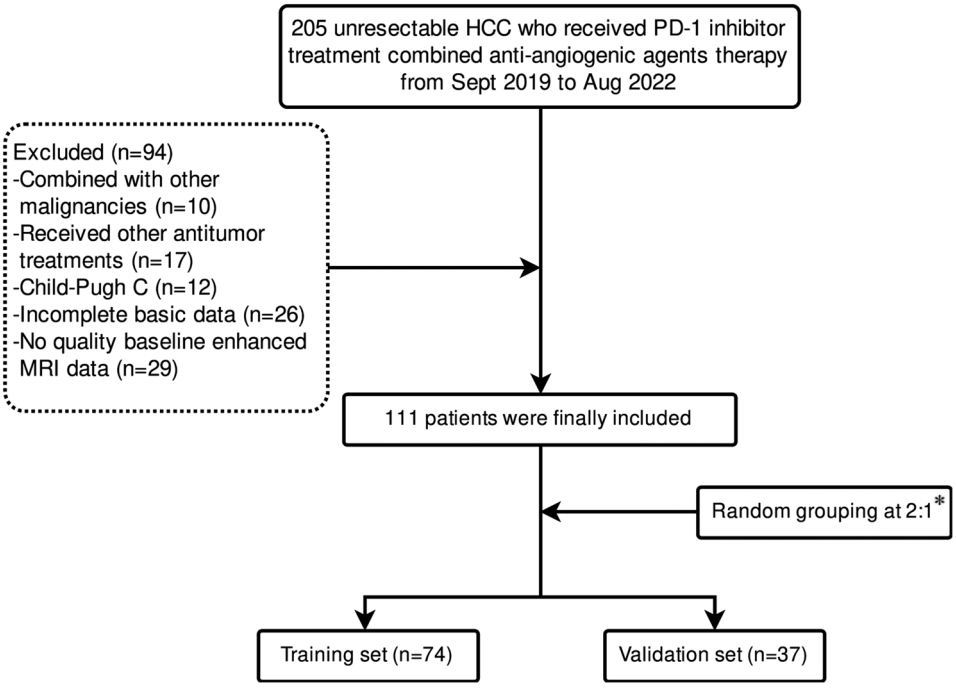

Fig. 1 Flowchart of patient enrollment. HCC, hepatocellular carcinoma; PD-1 Programmed death 1. *For training and validation of the model, use about 2/3 to 4/5 of the sample for training, and the rest for validation [49]. The 2:1 ratio is a commonly used approach in machine learning and radiomics studies to balance the need for a comprehensive training dataset with the requirement for a substantial validation dataset

and multiphase dynamic contrast-enhanced (DCE) imaging. For the enhanced sequences, the following parameters were used: repetition time (TR) = 2 or 3 ms, echo time (TE) = 1 ms, field of view (FOV) = 420 × 420 mm, matrix (Mat) = 320 × 208, flip angle (FA) = 11°, and slice thickness = 3.0 mm. The MRIs of all patients were conducted using 3.0-T machines within a period of 3 weeks prior to treatment. For the purpose of this study, only the enhanced sequences from the baseline MRI data were selected. Gadodiamide Injection (Omniscan®,0.2 mL/kg body weight) was intravenously administered with a flow rate of 1 mL/s. MR scans were performed in the arterial phase, portal venous phase, and delayed phase at 20–30 s, 60–70 s, and 180 s after injection, respectively.

Tumor segmentation

The feature data in this study were obtained from specific regions of interest (ROIs). Each patient underwent multiparameter MRIs using a 3.0-T MR scanner. The images were then imported into a 3D Slicer software (version 4.11; https://www.slicer.org/) for manual segmentation. Tumor segmentation was conducted by a radiologic oncologist and a proficient radiologist (5 years of radiological experience) on each transverse slice of 3 enhancement phases. The two reviewers were blinded to clinical data and then independently assessed the images. In cases where patients had multiple liver lesions, the largest lesion was assessed to extract radiomics features. Additionally, a senior radiologist (20 years of experience in abdominal tumor radiology) validated and evaluated each segmentation. Any discrepancies or issues that arose during the segmentation process were resolved by seeking the opinion of another senior radiologist.

Feature extraction

This study used the PyRadiomics platform (version 3.0.1, https://pyradiomics.readthedocs.io/en/latest/index.html) to employ various feature algorithms to quantify information on first order features, shape-based features, texture features (Fig. 2). MR images were resampled into isotropic resolution of $2.0 \times 2.0 \times 2.0$ mm^3 to eliminate the inconsistency of spatial resolution. This study utilized this platform to extract 971 features from each phase of the medical images. To ensure consistency and comparability, all radiomics features were normalized using z-scores.

Development and validation of PFS-Predictive models

All radiomics features were analyzed using a univariable Cox proportional hazard regression approach. Further investigation only considered variables that demonstrated a significant level ($P < 0.05$) in the univariate analysis. The initial feature set consisted of 2736 features extracted from the MRI images. While this high-dimensional data provides a comprehensive representation of the imaging information, it also increases the risk of overfitting. Therefore, the optimal radiomics features were selected using the LASSO algorithms. The Rad-score for three phase of enhanced MRI was calculated using a linear combination of selected features, which are weighted by their respective LASSO coefficients. Based on Cox regression, a prediction model including three phases of rad-score was established.

Random survival forest (RSF) is an ensemble tree method used to analyze right-censored data, which is a comprehensive method of random forest and survival analysis [30]. RSF can rank the importance of variables, and is known for its high predictive ability and interpretability of the relationship between variables. The RSF method was also used to incorporate the selected features to obtain the final Rad-score, and use Rad-score to build a prediction model.

Subsequently, the prediction ability of the two models was compared by C-index. The more predictive Rad-score enters later studies. Using the Rad-score obtained through the computational model, the participants were categorized into two groups based on their risk level: high-risk signature (HRS) group and low-risk signature (LRS) group. The threshold for categorization was determined by taking the median value. To analyze the relationship between the Rad-score and PFS, Kaplan-Meier survival analysis was conducted in both the training cohort and the validation cohort. In addition to the radiomics model, clinical characteristics were also considered in predicting PFS. Univariate and multivariate logistic regression analyses were conducted to select independent risk factors. Independent clinical factors were utilized to build a Clinical nomogram, while both the Rad-score and clinical risk factors were employed to create a Combined nomogram. These predictive models were then employed to forecast PFS in two cohorts.

Statistical analysis

This study utilized the t test or Mann–Whitney U test for continuous data analysis for between-group comparisons. The chi-square test or Fisher exact test were used to analyze categorical data. The Kaplan–Meier method was employed for survival curve analysis, which was later compared using the Log rank test. Additionally, Harrell's concordance index (C-index), time-dependent receiver operating characteristics (ROC) curves were employed to assess the forecast ability of the risk models. Moreover, calibration curves, decision curve analysis, subgroup analysis and N-fold cross-validation were used to measure the agreement between the predicted probabilities generated by our model and the actual observed outcomes and to evaluate the net benefit of using our predictive model across a range of probability thresholds.

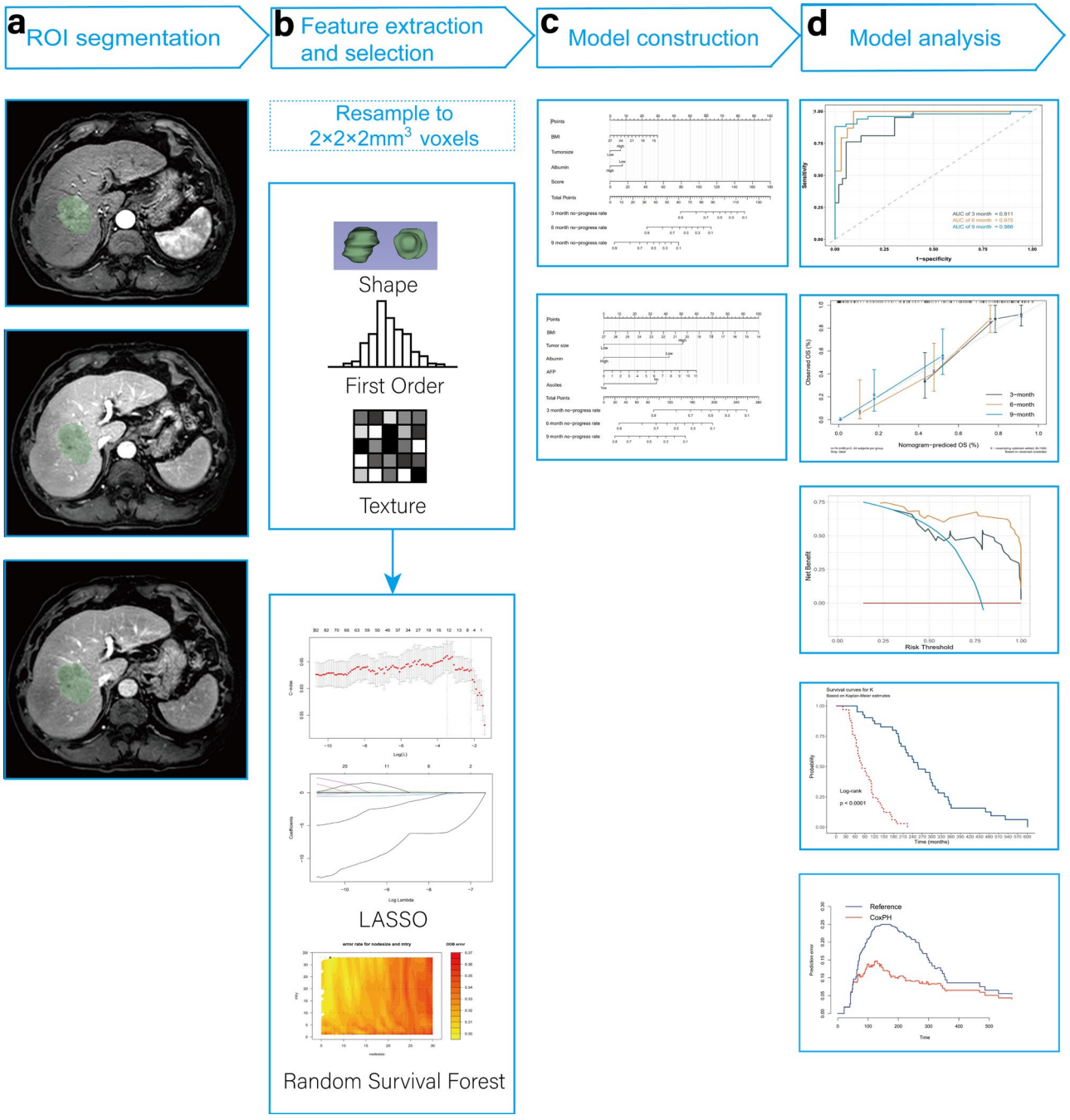

Fig. 2 Radiomics workflow in this study. **a** Image segmentation was performed in arterial phase, portal venous phase, and delayed phase of Contrast-enhanced MR images. **b** Features extracted from tumors were classified into shape, first-order statistics, texture, and the features were extracted by least absolute shrinkage and selection operator (LASSO) regression and random survival forest. **c** clinical nomogram and combined nomogram were established. **d** Two nomograms were compared by employing time-dependent receiver operating characteristics (ROC) curves, calibration curves, and decision curve analysis. Shape features capture the geometric properties of the tumor, which can be associated with tumor growth and invasiveness. First-order statistics describe basic image characteristics like mean intensity and variance. Texture features reflect the spatial distribution of pixel intensities, indicating tumor heterogeneity and complexity

R software (version 4.0.3, http://www.rproject.org/) was utilized for conducting statistical analyses. The packages employed in the analyses include 'survival', 'rms', 'ggplot2', 'RMS', 'survminer', 'randomForestSRC', 'dplyr', 'timeROC', and 'interp'. The code for the modeling described in this paper is available on GitHub at https://github.com/NI-design56/effective-funicular.

Results

Baseline characteristics

In this study, 111 patients who were treated with ICIs and anti-angiogenic agents were included. The mean age was 56.1 years, with 94.6% being male. Tumor size was significant, with 66.7% having tumors larger than 5 cm, and 49.5% having multiple tumors. Portal vein invasion was present in 47.7% of patients. The detailed flowchart of the study was shown in Fig. 1. Subsequently, 74 patients were designated for the training cohort, while 37 patients were allocated to the validation cohort. No significant variations in the baseline characteristics of the patients were observed between the two cohorts (Table 1).

Feature extraction

Each phase consisted of 912 features, with a total of 2736 features initially selected. Subsequently, univariable Cox proportional hazard regression analysis and the LASSO algorithm were applied in each phase, resulting in the retention of 32 features (12, 10, and 10 in the arterial, portal venous, and delayed phase images).

Rad-score calculation

Rad-score calculated by lasso-cox regression

The computational formulae for determining the Rad-score for each patient across the three distinct imaging phases is detailed in the supplementary materials provided with this study. The Rad-score of three phases was used to predict PFS by cox regression. The C-index of the training cohort was 0.760 (95% confidence interval (CI) 0.723–0.789), and the C-index of the Validation cohort was 0.748 (95%CI 0.692–0.763).

Rad-score calculated by RSF

The Grid Search (GS) method was used to determine the combination of mtry and nodesize values when the out-of-bag (OOB) error rate was lowest. As shown in Fig. 3a, under the parameter combination of mtry = 20 and nodesize = 9, the OOB error rate of the RSF model was the lowest (29.2%). Through parameter debugging, the error rate of the model tends to be stable when ntree was 1000 (Fig. 3b). Using the parameters obtained by the above method, the RSF model was constructed, and the importance of the independent variables was ranked using the VIMP method, as shown in the Table 2. Ultimately, 32 features were identified that were considered relatively valuable in all three phases. Considering the potential interrelationships of texture features, the established RSF model algorithm was used to predict PFS by calculating the Rad-score. The C-index of the training cohort was 0.837, and the C-index of the validation cohort was 0.830. Since the C-index of the RSF model was slightly higher than that of the Lasso-Cox regression model, the RSF model was finally used in this study to calculate the Rad-score.

Prediction models Establishing

In the training cohort, a total of 11 clinical variables were found to be predictive by univariate analysis (Table 3). Among these variables, 5 clinical factors consisting of BMI, tumor size, ALB, AFP, and ascites were independently selected as predictors for recurrence through the application of multivariate analysis. The clinical nomogram was constructed using these variables in the training cohort through Cox regression (Fig. 4a). However, when the combined nomogram was established, it was observed that the value of ALB and AFP for predicting survival was relatively insignificant. Therefore, the combined nomogram was modified to include BMI, tumor size, albumin, and Rad-score (Fig. 4b).

This study further assessed the predictive effectiveness of the aforementioned two models by comparing the C-index value and ROC curve. Results indicated that the composite clinical-and-radiomics models, which combined Rad-score with clinical risk factors, exhibited superior ROC curves compared to models using only clinical variables (Fig. 5). In terms of the C-index, the combined nomogram (0.846, 95% CI 0.804–0.879) demonstrated better prognosis prediction for PFS compared to the clinical nomogram (0.752, 95%CI 0.692–0.789). A parallel performance was observed in the validation cohort (C-index, 0.709, 95%CI 0.586–0.772 vs. 0.845, 95%CI0.767-0.893) (Table 4).

Verification of the radiomics nomogram

This study plotted ROC curves for both the clinical nomograms and the combined nomogram separately in the two cohorts. ROC curves evaluate the model's ability to distinguish patients with different outcomes. The higher the AUC, the stronger the discriminative power. In our study, the AUC of the combined model (0.966 for 9-month PFS in the training cohort) was higher than that of the clinical nomogram (0.850), indicating better risk stratification. A higher AUC helps identify patients at high risk of disease progression, thereby enabling more targeted treatment plans. The calibration curves (Fig. 6) showed a strong agreement between the predicted and observed PFS in both cohorts. Moreover, Calibration curves assess the consistency between predicted and observed outcomes. Well-calibrated predictions ensure

Table 1 Patient baseline characteristics in training and validation cohort

Characteristic	Entire cohort	Training cohort	Validation cohort	P-value*
	(*n* = 111)	(*n* = 74)	(*n* = 37)	
Age (years)				0.667
Mean (SD)	56.1(12.7)	56.5(12.4)	55.4(13.4)	
Median [Min, Max]	56[28,84]	57[31,84]	55[28,84]	
Gender				0.398
Male	105(5.4%)	71(95.9%)	34(91.9%)	
Female	6(94.6%)	3(4.1%)	3(8.1%)	
Smoking				0.783
No	43(38.7%)	28(37.8%)	15(40.5%)	
Yes	68(61.3%)	46(62.2%)	22(59.5%)	
Drinking				0.219
No	45(40.5)	27(36.5)	18(48.6%)	
Yes	66(59.5)	47(63.5)	19(51.4%)	
Ascites				0.13
No	89(80.2)	56(75.7%)	33(89.2%)	
Yes	22(19.8%)	18(24.3%)	4(10.8%)	
BMI(kg/m^2)				0.785
Mean (SD)	22.3(2.66)	22.3(2.66)	22.2(2.70)	
Median [Min, Max]	22.4[14.9,27.3]	22.4[14.9,26.9]	22.4[15.6,27.3]	
Chronic hepatitis				0.684
No	7(6.3)	4(5.4)	3(8.1)	
Yes	104(93.7)	70(94.6)	34(91.9)	
Liver cirrhosis				0.174
No	30(27)	17(23)	13(35.1)	
Yes	81(73)	57(77)	24(64.9)	
AFP(ng/mL)[a]				0.425
Mean (SD)	5.25(3.24)	5.42(3.34)	4.90(3.06)	
Median [Min, Max]	5.10[0.293,10.9]	5.40[0.507,10.9]	4.61[0.293,10.9]	
ALT(U/L)				0.574
Mean (SD)	49.2(54.8)	51.4(51.2)	44.8(62.1)	
Median [Min, Max]	33[10,383]	40[11,319]	29[10,383]	
AST(U/L)				0.1
Mean (SD)	66.9(60.1)	73.2(63.1)	54.4(52.3)	
Median [Min, Max]	49[18,406]	50[21,406]	40[18,327]	
ALB(g/L)				0.077
> 37.5	47(57.7)	27(36.5)	20(45.9)	
≤ 37.5	64(42.3)	47(63.5)	17(54.1)	
GGT(U/L)				0.121
Mean (SD)	188(238)	213(270)	138(146)	
Median [Min, Max]	118[19,1881]	132[19,1881]	88[21,678]	
TBIL(μmol/L)				0.841
Mean (SD)	18.3(10.6)	18.7(11.7)	17.3(8.04)	
Median [Min, Max]	15[5,69]	16[5,69]	14[6,48]	
DBIL(μmol/L)				0.142
Mean (SD)	7.05(5.45)	7.59(6.09)	5.97(3.69)	
Median [Min, Max]	6[2,29]	6.50[2,29]	5[2,17]	
PLT(×10^9/L)				0.652
Mean (SD)	172(156)	158(82.7)	198(243)	
Median [Min, Max]	147[15,1577]	148[15,459]	136[68,1577]	
LYM(×10^9/L)				0.691
Mean (SD)	1.34(0.604)	1.31(0.556)	1.40(0.70)	
Median [Min, Max]	1.20[0.40,3.7]	1.20[0.4,3.1]	1.20[0.4,3.7]	

Table 1 (continued)

Characteristic	Entire cohort	Training cohort	Validation cohort	P-value*
	(*n* = 111)	(*n* = 74)	(*n* = 37)	
PT(seconds)				0.892
Mean (SD)	14.3(1.87)	14.2(1.38)	14.3(2.6)	
Median [Min, Max]	14.1[11.6,28.4]	14.3[11.6,21.8]	13.9[12,28.4]	
INR				0.452
Mean (SD)	1.12(0.20)	1.12(0.14)	1.14(0.30)	
Median [Min, Max]	1.1[0.9,2.8]	1.1[0.9,1.9]	1.1[0.9,2.8]	
Child-Pugh class				0.476
A	92(82.9%)	60(81.1%)	32(86.5%)	
B	19(17.1%)	14(18.9%)	5(13.5%)	
BCLC stage				0.37
A	12(10.8%)	6(8.1%)	6(16.2%)	
B	26(23.4%)	19(25.7%)	7(18.9%)	
C	73(65.8%)	49(66.2%)	24(64.9%)	
ALBI grade				0.071
1	36(32.4%)	19(25.7%)	17(45.9%)	
2	73(65.8%)	53(71.6%)	20(54.1%)	
3	2(1.8%)	2(2.7%)	0	
TNM Classification				0.115
II	16(14.4%)	7(9.5%)	9(24.3%)	
III	19(17.1%)	14(18.9%)	5(13.5%)	
IV	76(68.5%)	53(71,6%)	23(62.2%)	
Tumor size(cm)				0.117
≤ 5.0	37(33.3%)	21(28.4%)	16(43.2%)	
> 5.0	74(66.7%)	53(71.6%)	21(56.8%)	
Tumor number				0.591
Solitary	56(50.5%)	36(48.6%)	20(54.1%)	
Multiple	55(49.5%)	38(51.4%)	17(45.9%)	
Portal vein invasion				0.282
No	58(52.3%)	36(48.6%)	22(59.5%)	
Yes	53(47.7%)	38(51.4%)	15(40.5%)	
Progressionfree survival (days)				0.789
Mean (SD)	184(131)	187(133)	179(131)	
Median [Min, Max]	144[21,601]	147[21,601]	144[21,600]	
Progress				0.21
No	12(10.8%)	6(8.1%)	6(16.2%)	
Yes	99(89.2%)	68(91.9%)	31(83.8%)	
Immune checkpoint inhibitors				0.598
Sintilimab	50(45.1%)	31(41.9%)	19(51.4%)	
Camrelizumab	35(31.5%)	24(32.4%)	11(29.7%)	
Toripalimab	26(23.4%)	19(25.7%)	7(18.9%)	
Anti-angiogenic agents				0.17
Lenvatinib	57(51.4%)	37(50%)	20(54.1%)	
Sorafenib	26(23.4%)	21(28.4%)	5(13.5%)	
Regorafenib	28(25.2%)	16(21.6%)	12(32.4%)	

[a] The value is the natural logarithm of clinical alpha-fetoprotein

Abbreviations: SD, standard deviation; BMI, Body Mass Index; AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; GGT, glutamyl transpeptidase; TBIL, total bilirubin; DBIL, Direct Bilirubin; PLT, platelet count; LYM, *lymphocyte;* PT, prothrombin time; INR, international normalized ratio; BCLC, Barcelona Clinic Liver Cancer; ALBI, albumin-bilirubin; TNM, Tumor Node Metastasis

P-value* < 0.05 indicates a significant difference between the training and validation cohorts

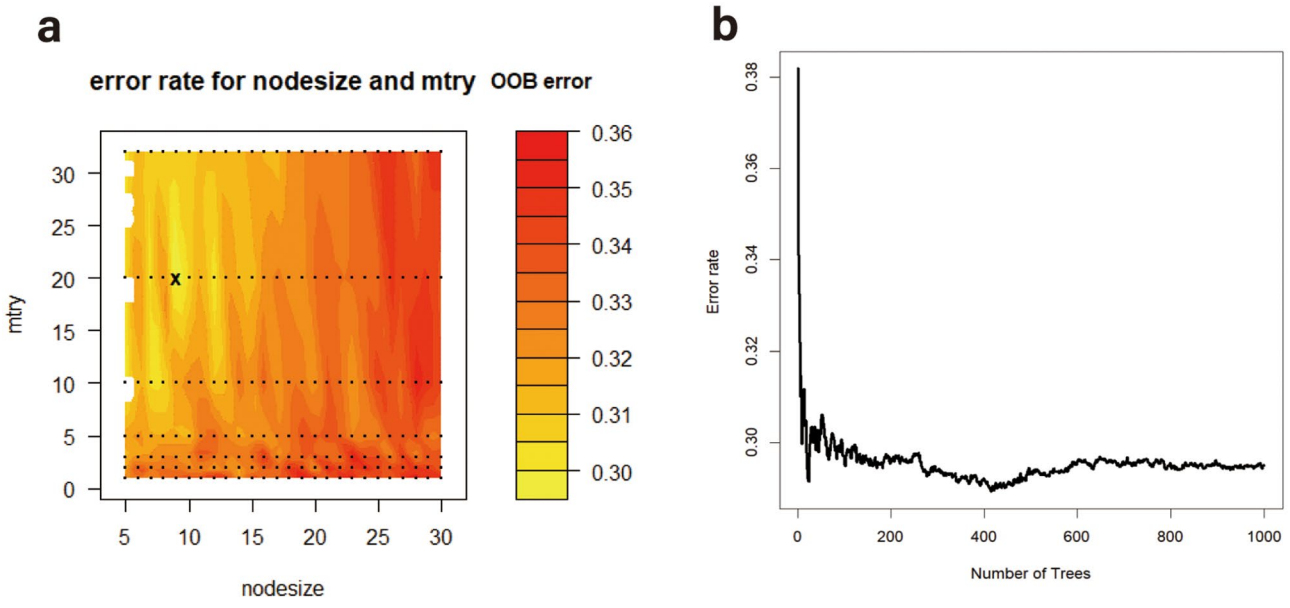

Fig. 3 **a** Tuning parameter of the RSF model. The OOB error rate is a measure of model accuracy, with lower rates indicating better predictive performance. The black sign in the figure is the parameter combination of mtry = 20 and nodesize = 9 and the corresponding OOB error rate is 29.2%. **b** Curve of the OOB error rate for RSF model

that the model's outputs can be confidently used for treatment planning and resource allocation. Our combined model showed good calibration in both cohorts (Fig. 7).

Furthermore, to validate the robustness of our model, we conducted subgroup analyses based on body mass index (BMI). Patients were categorized into high BMI (BMI ≥ 25 kg/m²) and low BMI (BMI < 25 kg/m²) subgroups. The combined nomogram maintained high predictive accuracy within both BMI subgroups, with C-index values of 0.809 (95% CI 0.756–0.848) in the high BMI subgroup and 0.874 (95% CI 0.804–0.931) in the low BMI subgroup, confirming the model's generalizability across different BMI categories. This subgroup validation using BMI subgroups further reinforces the clinical utility of our model in predicting PFS in patients with uHCC.

Additionally, we used cross-validation to verify the survival prediction model. The cross-validation results showed that the combined nomogram had a significantly lower prediction error compared to a reference model, with a prediction error of 0.071 for the combined nomogram versus 0.129 for the reference model (Fig. 8). This indicates that the combined nomogram not only performs well in terms of discrimination and calibration but also has a robust predictive performance, further validating its reliability and utility in clinical settings.

Prognostic validation of rad-score

The Rad-score has been validated as a standalone prognostic determinant for individuals diagnosed with uHCC. The classification of Rad-score was determined by establishing a threshold using the median value. This approach effectively segregated patients into either HRS or LRS groups. Survival analysis was conducted in this study using the Kaplan-Meier method, with the main focus being on monitoring disease progression as the primary endpoint (Fig. 9).

Discussion

HCC is one of the most common malignant tumors worldwide, with a particularly poor prognosis for patients. The prognosis of HCC is characterized by significant heterogeneity, and the disease is challenging to treat. Therefore, there is an urgent need to characterize the disease status of individual patients to enable patient stratification and personalized treatment, ultimately maximizing prognosis. In this study, we employed machine learning-based radiomics to predict the prognosis of patients receiving immune checkpoint inhibitors in combination with anti-angiogenic drugs for uHCC. This approach could assist in identifying high-risk patients who would benefit from appropriate treatment.

At present, there are multiple treatment alternatives accessible for advanced and uHCC, involving local-regional therapy, immunotherapy, and molecular targeted medications [31]. The emergence of ICIs has significantly enhanced the prognosis of different solid tumors, including HCC. Moreover, it has displayed synergistic impacts when merged with anti-angiogenic agents [32]. A study was carried out by Shigeta et al. to examine the effects of a combination therapy consisting of anti-PD-1/VEGFR-2 on survival rates in HCC using orthotopic

Table 2 Imaging features selected for the construction of combined model

Phase	Features number	Features selected	Importance
Arterial phase	12	wavelet.LLH_glszm_LargeAreaLowGrayLevelEmphasis	0.027912
		wavelet.HLL_glcm_Autocorrelation	0.016868
		wavelet.LLL_glszm_GrayLevelNonUniformity	0.018511
		wavelet.HHL_firstorder_Maximum	0.01982
		wavelet.LLH_glcm_MaximumProbability	0.020024
		wavelet.LHH_glcm_ClusterProminence	0.008468
		wavelet.HLH_gldm_LargeDependenceHighGrayLevelEmphasis	0.011627
		wavelet.LLL_glcm_Imc2	0.010737
		wavelet.HHH_firstorder_Median	0.005421
		wavelet.HLH_glszm_LargeAreaLowGrayLevelEmphasis	0.00187
		wavelet.LLL_glrlm_GrayLevelNonUniformity	0.017364
		wavelet.LHH_glcm_Imc2	0.002767
Portal venous phase	10	wavelet.LLL_glcm_Imc2	0.177879
		wavelet.HHH_firstorder_Mean	0.007811
		log.sigma.3.0.mm.3D_glcm_ClusterShade	0.013591
		wavelet.HLH_glszm_LargeAreaLowGrayLevelEmphasis	0.002256
		wavelet.LLH_firstorder_Kurtosis	0.006625
		wavelet.LLL_glszm_LargeAreaHighGrayLevelEmphasis	0.000742
		original_glszm_LargeAreaHighGrayLevelEmphasis	0.003997
		log.sigma.5.0.mm.3D_glcm_Imc2	0.002265
		wavelet.LLL_gldm_GrayLevelNonUniformity	0.003511
		wavelet.LHH_glrlm_ShortRunEmphasis	0.003647
Delayed phase	10	wavelet.LLL_glcm_Imc2	0.030602
		wavelet.LLH_firstorder_Kurtosis	0.006942
		log.sigma.3.0.mm.3D_glcm_ClusterShade	0.014806
		log.sigma.5.0.mm.3D_firstorder_90Percentile	0.00555
		wavelet.HHH_glcm_Imc2	0.003582
		wavelet.HLL_glcm_Correlation	0.010021
		wavelet.HHH_glszm_LargeAreaLowGrayLevelEmphasis	0.006388
		wavelet.LLL_gldm_LargeDependenceHighGrayLevelEmphasis	0.007395
		log.sigma.3.0.mm.3D_gldm_SmallDependenceHighGrayLevelEmphasis	0.001644
		wavelet.LLL_glszm_GrayLevelNonUniformity	0.01647

Wavelet Transform Features: These features are extracted using wavelet transform, a mathematical tool that allows for the analysis of data at different scales and orientations. "Wavelet. LLH_glszm_LargeAreaLowGrayLevelEmphasis" specifically measures the emphasis of large areas with low gray-level intensities within the tumor, which may indicate regions of necrosis or hypoxia and are associated with tumor aggressiveness and treatment resistance

GLSZM (Gray Level Size Zone Matrix) Features: Derived from the gray level size zone matrix, these features describe the distribution of pixel intensities within zones of uniform intensity. They provide information about the spatial arrangement of tissue characteristics within the tumor and can reflect tumor heterogeneity

GLCM (Gray Level Co-occurrence Matrix) Features: Calculated from the gray level co-occurrence matrix, these features quantify the spatial relationship between pixel intensities. Features such as "Wavelet. HLL_glcm_Autocorrelation" reflect the similarity of pixel values across the image, indicating the uniformity of the tumor texture

transplantation or induced mouse models [33]. By concurrently targeting VEGFR-2 and PD-1, this dual blockade can promote the normalization of blood vessels, alter the immune microenvironment, and augment the anti-tumor immune response in HCC. Since selecting the most suitable treatment approach for advanced HCC patients from the available options can be challenging, it is crucial to identify potential HCC patients who will respond positively to ICI and anti-angiogenic agents. Optimal treatment strategies and the attainment of therapeutic success heavily rely on the accurate identification of respective conditions. Regrettably, there is a dearth of studies that have prioritized the implementation of ML radiomics techniques in prognosticating the responsiveness of HCC to combination therapy consisting of ICIs and anti-angiogenic agents. In this investigation, researchers adeptly devised radiomics model utilizing ML to forecast the reaction to ICIs in concurrence with anti-angiogenic agents.

Radiomics analysis involves the use of high-throughput feature extraction algorithms to quantitatively measure macroscopic disease features within and between tumors [18, 34], extending beyond the field of HCC to explore its clinical value in different types of tumors. ML is a well-established AI technique employed to develop predictive models, and it is extensively utilized in hepatology

Table 3 The results of univariate and multivariate Cox proportional hazard analysis in this study

Variable	Univariate cox regression		Multivariate cox regression	
	P-value*	Hazard ratio (95% confidence interval)	*P*-value*	Hazard ratio (95% confidence interval)
Age	0.633	0.99(0.98, 1.02)		
Gender				
Male	0.637	0.75 (0.23, 2.43)		
Female				
Smoking				
No				
Yes	0.775	0.93 (0.57,1.53)		
Drinking				
No				
Yes	0.691	1.11(0.67, 1.83)		
Ascites				
No				
Yes	0.0148	1.99(1.14, 3.46)	0.0446	1.86 (1.02, 3.41)
BMI	0.001	0.84 (0.76,0.93)	0.0173	0.87(0.77, 0.98)
Chronic hepatitis				
No				
Yes	0.144	0.46(0.16, 1.30)		
Liver cirrhosis				
No				
Yes	0.89	1.04 (0.58, 1.87)		
AFP [a]	0.0004	1.16 (1.07, 1.26)	0.021	1.10 (1.02, 1.20)
ALT	0.731	1.00 (0.99, 1.005)		
AST	0.0428	1.00 (1, 1.007)	0.68	0.999(0.994, 1.00)
ALB				
≤ 37.5	0.003	2.18 (1.30, 3.66)	0.0103	2.16 (1.20, 3.90)
> 37.5				
GGT	0.191	1.00 (0.99, 1.001)		
TBIL	0.125	1.02(1.00, 1.04)		
DBIL	0.025	1.05(1.01, 1.09)	0.52	1.02(0.96, 1.08)
PLT	0.112	1.00 (0.99, 1.01)		
LYM	0.919	1.02 (0.66, 1.60)		
PT	0.0793	1.17(0.98, 1.39)		
INR	0.0263	6.01(1.24, 29.21)	0.061	6.11(0.92, 40.62)
Child-Pugh class				
A				
B	0.0002	3.65(1.85, 7.19)	0.46	0.66(0.21, 2.01)
BCLC stage				
A				
B	0.542	1.36(0.50, 3.69)		
C	0.334	1.58(0.62, 4.03)		
ALBI grade				
1				
2	0.072	1.69(0.95, 2.98)		
3	0.094	3.59(0.80, 16.01)		
TNM Classification				
II				
III	0.37	1.56(0.59, 4.14)		
IV	0.252	1.65(0.70, 3.89)		
Tumor size				
≤ 5.0	0.017	0.53 (0.30, 0.89)	0.004	0.40 (0.21, 0.74)
>5				
Tumor number				

Table 3 (continued)

Variable	Univariate cox regression		Multivariate cox regression	
	P-value*	Hazard ratio (95% confidence interval)	*P*-value*	Hazard ratio (95% confidence interval)
Solitary				
Multiple	0.015	1.863(1.13, 3.08)	0.081	1.614(0.94, 2.77)
Portal vein invasion				
No				
Yes	0.013	1.865(1.14, 3.06)	0.465	1.228(0.71, 2.13)

[a] The value is the natural logarithm of clinical alpha-fetoprotein

P-value* < 0.05 indicates a significant difference between the training and validation cohorts

Abbreviations: BMI, Body Mass Index; AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; GGT, glutamyl transpeptidase; TBIL, total bilirubin; DBIL, Direct Bilirubin; PLT, platelet count; LYM, *lymphocyte;* PT, prothrombin time; INR, international normalized ratio; BCLC, Barcelona Clinic Liver Cancer; ALBI, albumin-bilirubin; TNM, Tumor Node Metastasis

research [35]. Previous studies have demonstrated that AI radiomics holds significant promise in forecasting the prognosis of patients with HCC. Akai et al. included 127 hepatocellular carcinoma (HCC) patients who underwent initial hepatectomy and developed a radiomic model utilizing the Random Survival Forest (RSF) algorithm to effectively distinguish between high-risk and low-risk groups for survival risk, demonstrating high consistency [36]. Wang et al. included 201 HCC patients who underwent curative hepatectomy in a retrospective, multicenter study. The 30 radiomics features most associated with survival were selected from preoperative MRI, and combined with preoperative AFP and AST as independent clinical risk factors, the model was constructed by RSF. The model demonstrated a good ability to predict 5-year survival, with a mean AUC of 0.9804 and 0.7578 in the training and validation cohorts, respectively [37]. Bo Z et al. introduced a valuable machine learning radiomics model that showed excellent performance in predicting the response to lenvatinib monotherapy in uHCC, achieving the highest AUC of 0.97 [38].

In the current investigation, researchers constructed a combined nomogram that integrates the Rad-score and three clinical characteristics (body mass index, tumor dimension, and albumin levels) to anticipate the efficacy of ICIs coupled with anti-angiogenic agents in managing uHCC. In this study, the combined model achieved a C-index of 0.846, which is significantly higher than predictive models that rely solely on clinical factors [39]. The models formulated in this study can provide valuable guidance in the judicious administration of ICIs and anti-angiogenic agents in real-world medical practice. The models formulated in this study can provide valuable guidance in the judicious administration of ICIs and anti-angiogenic agents in real-world medical practice. The combined nomogram can identify patients at higher risk of disease progression, allowing for the customization of treatment strategies. For instance, a patient classified as high risk might be considered for more aggressive or alternative therapies, potentially leading to improved patient outcomes. Furthermore, the combined nomogram provides a quantitative risk assessment, which can be used to inform patients about their prognosis and potential treatment outcomes. This information is crucial for shared decision-making, allowing patients to understand their disease trajectory and participate in choosing the most suitable treatment path.

In some radiomics studies, a large number of radiological features have been employed without appropriate feature selection methods, leading to overfitting and consequently poor performance on new, unseen data [40]. This study adopted a rigorous feature selection process using univariate Cox regression and the LASSO algorithm. This approach helps to reduce overfitting by selecting only the most predictive features. Additionally, this study utilized RSF to construct our model, which is less prone to overfitting compared to other machine learning algorithms [41]. Moreover, some radiomics studies have focused solely on imaging features without integrating them with clinical data, which can limit the model's predictive power as clinical factors often provide important additional information [42]. This study integrates radiological features with clinical variables such as BMI, tumor size, and albumin levels. This integrated approach provides a more comprehensive view of the patient's condition and enhances the model's predictive capability.

The use of MRI in this study is a significant differentiator from other imaging modalities such as CT. MRI provides superior soft tissue contrast compared to CT, which is crucial for imaging the liver and detecting small tumors or subtle changes in tumor morphology. This enhanced contrast allows for more accurate delineation of tumor boundaries and the identification of intra-tumoral heterogeneity, which is particularly important in the management of HCC. Furthermore, MRI offers multiparametric imaging capabilities, including T1-weighted, T2-weighted, diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE) sequences. Each sequence provides unique information about the tumor's characteristics, such as cellularity, vascularity, and diffusion properties. This multi-parametric approach provides

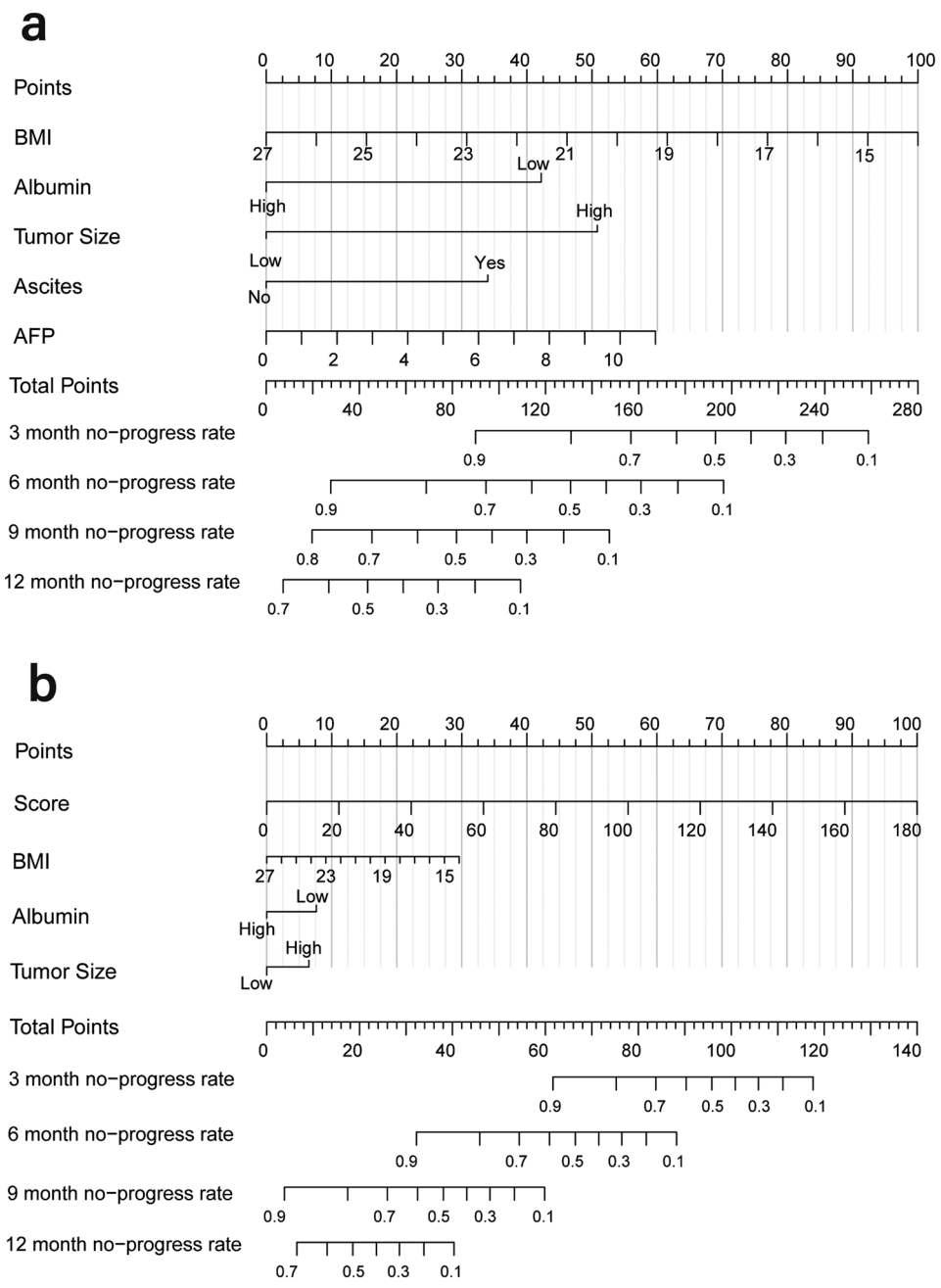

Fig. 4 Established clinical nomogram for combination regimen of unresectable hepatocellular carcinoma. (**a**) The Clinic nomogram for progression-free survival (PFS) including selected clinical risk factors of Table 3. For instance, consider a patient with a BMI of 27 (25 points), high albumin levels (23 points), large tumor size (25 points), presence of ascites (10 points), and an AFP score of 6 (6 points), totaling 89 points. Referring to the nomogram, this patient would fall into the high-risk stratum, with a 3-month PFS rate of approximately 0.5, indicating a moderate to high risk of disease progression. (**b**) The Combined nomogram for RFS including clinical risk factors and rad-score

a more comprehensive view of the tumor's biology and behavior, enhancing the predictive power of our model [43–45]. Variations in MRI imaging protocols and scanner settings can significantly impact the reproducibility of radiomic features. To mitigate this issue, the study used a standardized imaging protocol across all patients in our study. This included the use of a consistent MRI scanner model and settings, ensuring that the images were acquired under the same conditions.

Nomograms are commonly used predictive models in evaluating tumor prognosis. Since a majority of hepatocellular carcinoma (HCC) instances are linked to chronic hepatitis and cirrhosis of the liver, the prognostic outcome for liver cancer relies not solely on the tumor, but also on various factors encompassing the patient's overall

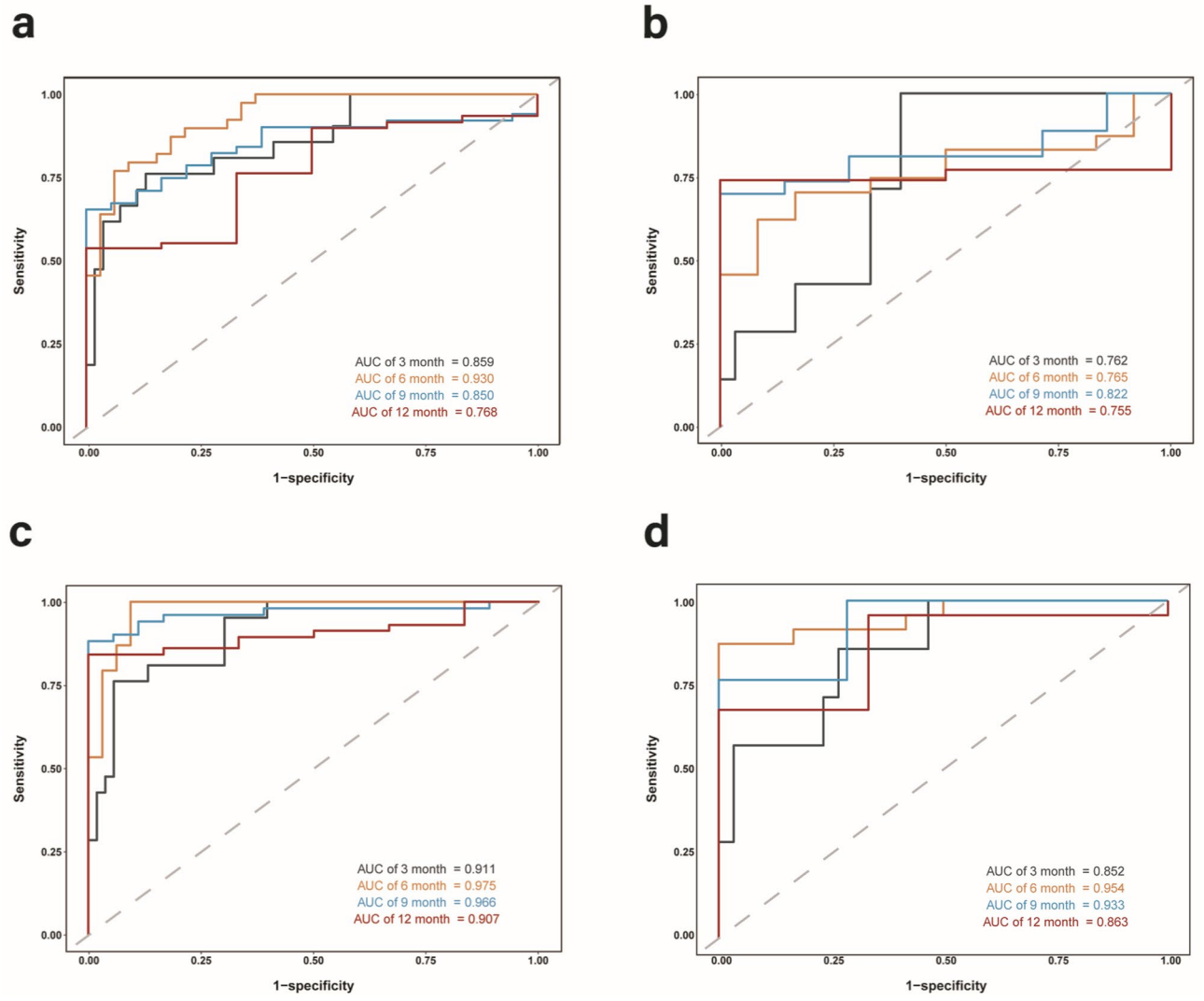

Fig. 5 **a** The time-dependent ROC curves of clinical nomogram in the training cohort. **b** The time-dependent ROC curves of clinical nomogram in validation cohort. **c** The time-dependent ROC curves of clinical and radiomics nomogram in the training cohort. **d** The time-dependent ROC curves of clinical and radiomics nomogram in the validation cohort. AUC Area under the curve

Table 4 The performances of the models in the training and validation cohort

Cohort	Rad-Score C-Index (95% CI)	Clinical Nomogram C-Index (95% CI)	Combined Nomogram C-Index (95% CI)	% Improvement Over Clinical Nomogram
Training cohort	0.837 (0.796, 0.878)	0.752 (0.692, 0.789)	0.846 (0.804, 0.879)	**12.8%**
Validation cohort	0.830 (0.744, 0.912)	0.709 (0.586, 0.772)	0.845 (0.767, 0.893)	**19.1%**

nutritional condition, the localized inflammatory microenvironment, and the functioning of the liver reserve. This study found that BMI, tumor size, and albumin levels are associated with the prognosis of uHCC patients receiving combination therapy. BMI serves as a straightforward and intuitive indicator to measure individual obesity levels. Previous research findings support the observation that patients with a higher BMI exhibit an extended period of PFS subsequent to treatment for HCC [46]. This could be attributed to the fact that advanced HCC is a highly consumptive disease, and patients with higher BMI may have better nutritional, physical, and mental statuses, along with greater energy reserves to combat tumor consumption. In addition, tumor size and albumin are important variables in predicting the prognosis of HCC. Many staging systems use a tumor size of 5 cm as a cutoff point for treatment and prognosis of HCC [47, 48]. However, this study did not yield any statistically significant findings concerning ALBI grading. This outcome might be attributed to the inclusion

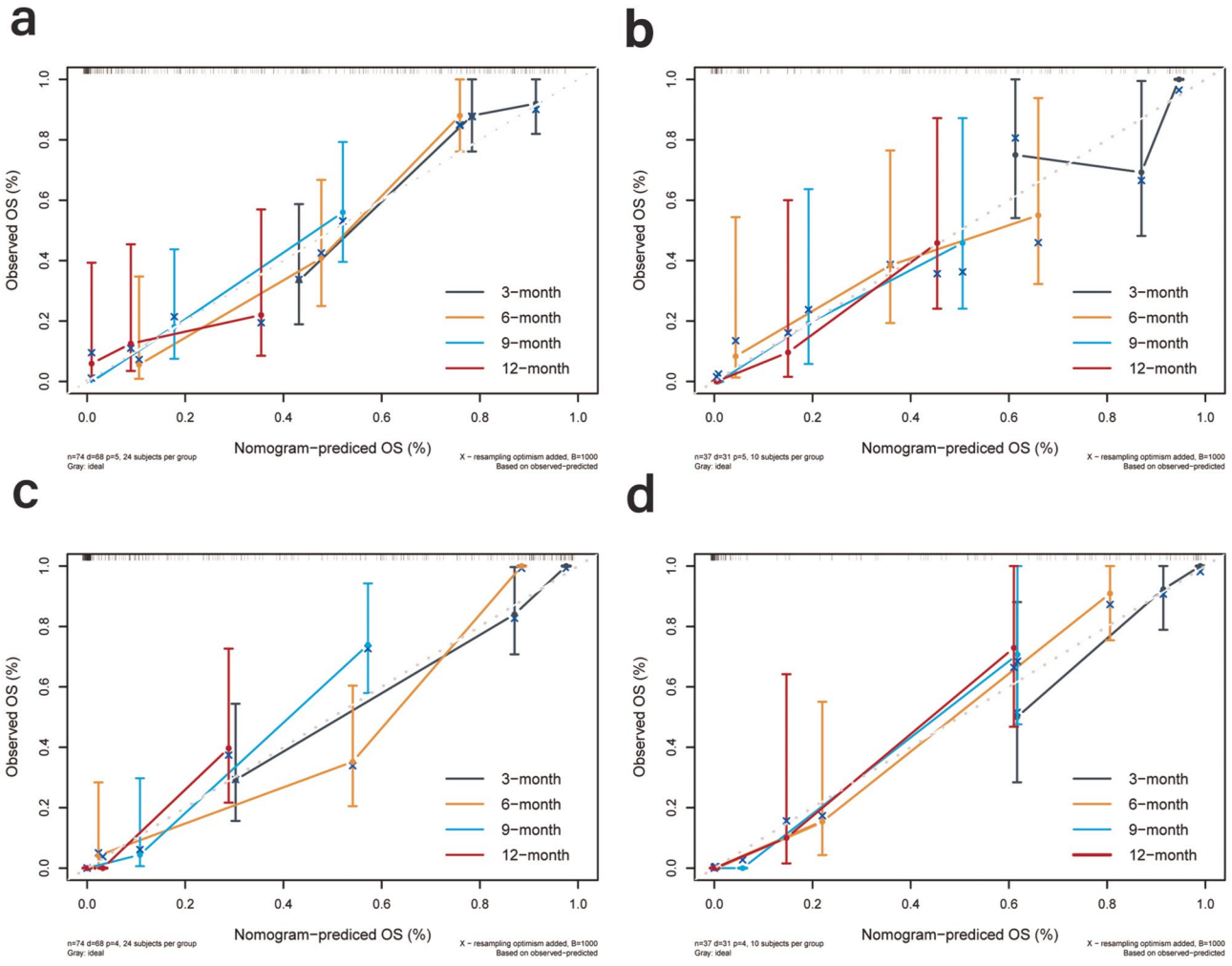

Fig. 6 Calibration curves of clinical nomogram and clinicalandradiomics nomogram in the training (**a**, **c**) and validation cohort (**b**, **d**)

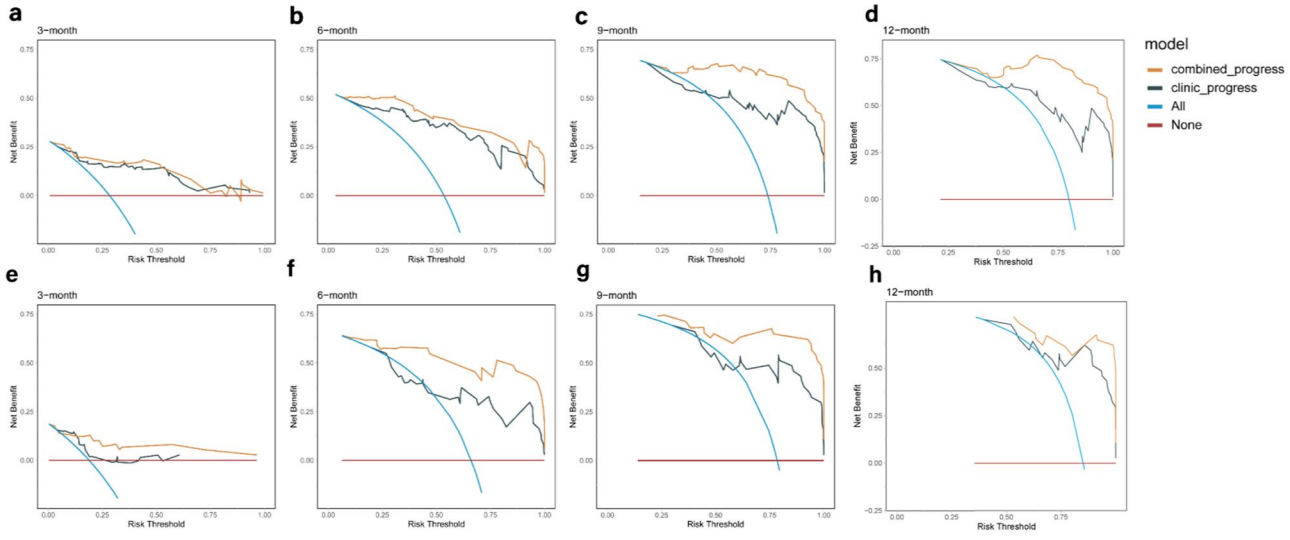

Fig. 7 Decision curve analysis for clinical and clinicalandradiomics nomogram models in the training (**a**, **b**, **c**, **d**) and validation cohorts (**e**, **f**, **g**, **h**)

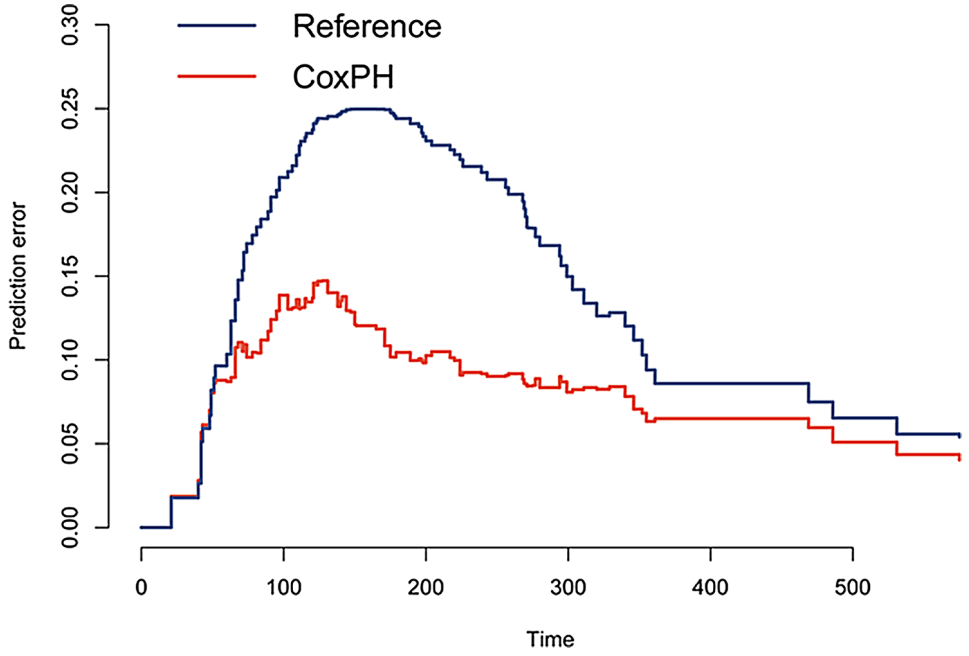

Fig. 8 Combined model validation: prediction error curves

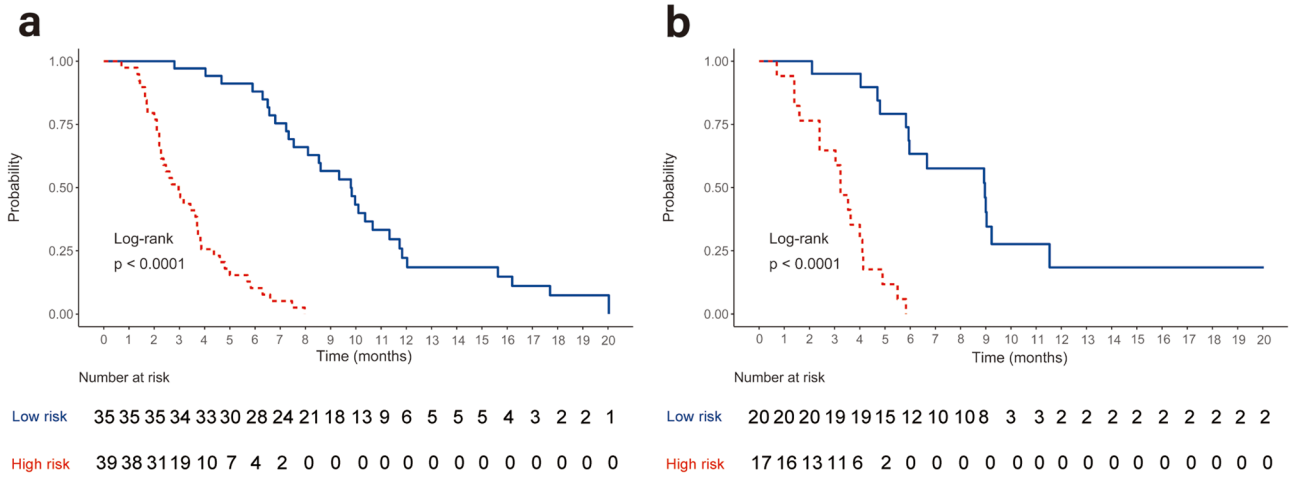

Fig. 9 Kaplan–Meier diagrams of HCC patients stratified according to rad-score **a** PFS survival curve of patients in the training cohort. **b** PFS survival curve of patients in the validation cohort

of patients with adequate liver function and the lack of substantial variations in bilirubin levels among them. Furthermore, AFP was omitted from the final integrated model, likely because the HCC patients encompassed in our study were predominantly in the intermediate and advanced stages, with the majority exhibiting high AFP values. When other clinical and radiological features were incorporated into the model, the hazard ratio (HR) of AFP failed to reach statistical significance, and its exclusion did not notably impact the overall performance of the nomogram.

This study has several limitations that warrant acknowledgment. Firstly, the retrospective nature of the study design introduces potential selection biases in the collection and retrieval of MRI and clinical data. This inherent limitation highlights the necessity for additional prospective studies to validate the findings of this investigation. Secondly, there is variability in the use of ICIs and TKIs among patients. However, no studies have yet conducted comparative analyses to establish differences in effectiveness between various PD-1 inhibitors. Thirdly, the short-term follow-up, averaging six months, may restrict our ability to obtain long-term outcomes and fully assess the durability of PFS predictions. To address this limitation, we performed internal validation using N-fold cross-validation to evaluate the model's performance on a subset of

the data, thereby providing a more conservative estimate of its predictive power. Lastly, the study was conducted at a single institution, which may limit the generalizability of the findings to patients in different regions and institutions. In light of these limitations, we plan to conduct subgroup analyses by drug class in future studies to assess the impact of different ICIs and TKIs on treatment outcomes. We also intend to extend the follow-up period in our ongoing research to obtain long-term results, which will enable us to more accurately assess the durability of PFS predictions and validate the model's performance over a longer time horizon. Furthermore, we plan to explore the use of automatic segmentation techniques, such as deep learning-based methods, to improve the accuracy and repeatability of tumor segmentation. Given that the combination of TKIs and ICIs is a relatively new approach, large-scale, multicenter subgroup analysis studies should be prioritized for future research.

Conclusion

This study successfully developed an MRI-based radiomics model that accurately predicts progression-free survival in patients with unresectable hepatocellular carcinoma. The combined nomogram, which integrates clinical factors and radiomic features, demonstrated superior predictive accuracy compared to clinical nomogram alone. The model's effectiveness across BMI subgroups highlights its broad applicability. Future research will focus on external validation and the potential inclusion of additional biomarkers to further enhance model performance.

Abbreviations

HCC	Hepatocellular carcinoma
TKI	Tyrosine kinase inhibitor
ICI	Immune checkpoint inhibitor
PD-1	Programmed death 1
PFS	Progression-free survival
LASSO	Least absolute shrinkage and selection operator
RSF	Random survival forests
C-index	Harrell's concordance index
ROC	Receiver operating characteristics
AUC	Area under the curve
CI	Confidence interval
AFP	Alpha-fetoprotein
BMI	Body Mass Index
ALBI	Albumin-bilirubin

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-14247-0.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

G.L., Y.D. and G.C. conceived and designed the study. X.X. and X.J. acquired the data. H.J. collected the follow-up data. X.X., M.Z., Y.D., took part in performing image segmentation and data extraction. X.X., X.J., H.J., X.Y. and Y.W. analyzed the data and developed the model. X.X. and Y.D. were major contributors in writing the paper. All authors read and approved the final manuscript.

Funding

This work was supported by Beijing CSCO Clinical Oncology Research Foundation [Y-2020Sciclone/zb/ms-0006]; Life Oasis Public Service Center [BJHA-CRP-070]; Wenzhou Bureau of Science and Technology [Y20190159 and Y2023217]; the Zhejiang Provincial Natural Science Foundation of China [ZCLMS25H1606].

Data availability

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

Declarations

Ethics approval and consent to participate

The study was carried out according to the ethical guidelines of the Helsinki Declaration and was approved by the Ethics Review Committee of the First Affiliated Hospital of Wenzhou Medical University (Approval Number KY2024-R022). The Ethics Committee of the First Affiliated Hospital of Wenzhou Medical University has waived the requirement for informed consent due to the retrospective nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Radiology, Shaoxing Central Hospital, The Central Affiliated Hospital, Shaoxing University, Shaoxing 312000, China
[2]Department of Pathology, Jinhua Municipal Central Hospital, Affiliated Jinhua Hospital, Zhejiang University School of Medicine, Jinhua 321000, China
[3]Department of Radiation and Chemotherapy Oncology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China
[4]Department of Radiology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang 325000, China
[5]Department of Hepatobiliary Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang 325000, China

Received: 1 December 2024 / Accepted: 29 April 2025
Published online: 19 May 2025

References

1. Sung H, Ferlay J, Siegel RL, et al. 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. Park JW, Chen M, Colombo M, et al. Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE study. Liver Int. 2015;35(9):2155–66.
3. Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008;359(4):378–90.
4. Cheng AL, Kang YK, Chen Z, et al. Efficacy and safety of Sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol. 2009;10(1):25–34.
5. Bruix J, Qin S, Merle P, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on Sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2017;389(10064):56–66.

6. Abou-Alfa GK, Meyer T, Cheng AL, et al. Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. N Engl J Med. 2018;379(1):54–63.
7. Kudo M, Finn RS, Qin S, et al. Lenvatinib versus Sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet. 2018;391(10126):1163–73.
8. Zhu AX, Kang YK, Yen CJ, et al. Ramucirumab after Sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2019;20(2):282–96.
9. Cheng AL, Hsu C, Chan SL, Choo SP, Kudo M. Challenges of combination therapy with immune checkpoint inhibitors for hepatocellular carcinoma. J Hepatol. 2020;72(2):307–19.
10. Zhu AX, Finn RS, Edeline J, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with Sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet Oncol. 2018;19(7):940–52.
11. Yau T, Kang YK, Kim TY, et al. Efficacy and safety of nivolumab plus ipilimumab in patients with advanced hepatocellular carcinoma previously treated with Sorafenib: the checkmate 040 randomized clinical trial. JAMA Oncol. 2020;6(11):e204564.
12. Xu J, Shen J, Gu S, et al. Camrelizumab in combination with apatinib in patients with advanced hepatocellular carcinoma (RESCUE): A nonrandomized, Open-label, phase II trial. Clin Cancer Res. 2021;27(4):1003–11.
13. Zhang W, Gong C, Peng X, et al. Serum concentration of CD137 and tumor infiltration by M1 macrophages predict the response to sintilimab plus bevacizumab biosimilar in advanced hepatocellular carcinoma patients. Clin Cancer Res. 2022;28(16):3499–508.
14. Finn RS, Ikeda M, Zhu AX, et al. Phase Ib study of lenvatinib plus pembrolizumab in patients with unresectable hepatocellular carcinoma. J Clin Oncol. 2020;38(26):2960–70.
15. Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. N Engl J Med. 2020;382(20):1894–905.
16. Ren Z, Xu J, Bai Y, 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.
17. Pinato DJ, Mauri FA, Spina P, et al. Clinical implications of heterogeneity in PD-L1 immunohistochemical detection in hepatocellular carcinoma: the Blueprint-HCC study. Br J Cancer. 2019;120(11):1033–6.
18. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77.
19. Lewis S, Hectors S, Taouli B. Radiomics of hepatocellular carcinoma. Abdom Radiol (NY). 2021;46(1):111–23.
20. Xia T, Zhao B, Li B, et al. MRI-Based radiomics and deep learning in biological characteristics and prognosis of hepatocellular carcinoma: opportunities and challenges. J Magn Reson Imaging. 2024;59(3):767–83.
21. Dong W, Ji Y, Pi S, Chen QF. Noninvasive imaging-based machine learning algorithm to identify progressive disease in advanced hepatocellular carcinoma receiving second-line systemic therapy. Sci Rep. 2023;13(1):10690.
22. Kuang Y, Li R, Jia P, et al. MRI-Based radiomics: nomograms predicting the short-term response after transcatheter arterial chemoembolization (TACE) in hepatocellular carcinoma patients with diameter less than 5 cm. Abdom Radiol (NY). 2021;46(8):3772–89.
23. Zhao Y, Wang N, Wu J, et al. Radiomics analysis based on Contrast-Enhanced MRI for prediction of therapeutic response to transarterial chemoembolization in hepatocellular carcinoma. Front Oncol. 2021;11:582788.
24. Zhang L, Xia W, Yan ZP, et al. Deep learning predicts overall survival of patients with unresectable hepatocellular carcinoma treated by transarterial chemoembolization plus Sorafenib. Front Oncol. 2020;10:593292.
25. Anagnostou V, Yarchoan M, Hansen AR, et al. Immuno-oncology trial endpoints: capturing clinically meaningful activity. Clin Cancer Res. 2017;23(17):4959–69.
26. Cabibbo G, Celsa C, Enea M et al. Progression-Free survival early assessment is a robust surrogate endpoint of overall survival in immunotherapy trials of hepatocellular carcinoma. Cancers (Basel). 2020;13(1).
27. Bo Z, Song J, He Q, et al. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. Comput Biol Med. 2024;173:108337.
28. Benson AB, D'Angelica MI, Abbott DE, et al. Hepatobiliary cancers, version 2.2021, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw. 2021;19(5):541–65.
29. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228–47.
30. Ishwaran HM, Lu. Random survival forests. Wiley StatsRef: Statistics Reference Online; 2019. pp. 1–13.
31. Bruix J, Chan SL, Galle PR, Rimassa L, Sangro B. Systemic treatment of hepatocellular carcinoma: an EASL position paper. J Hepatol. 2021;75(4):960–74.
32. Leone P, Solimando AG, Fasano R et al. The evolving role of immune checkpoint inhibitors in hepatocellular carcinoma treatment. Vaccines (Basel). 2021;9(5).
33. Shigeta K, Datta M, Hato T, et al. Dual programmed death Receptor-1 and vascular endothelial growth factor Receptor-2 Blockade promotes vascular normalization and enhances antitumor immune responses in hepatocellular carcinoma. Hepatology. 2020;71(4):1247–61.
34. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
35. Pellat A, Barat M, Coriat R, Soyer P, Dohan A. Artificial intelligence: A review of current applications in hepatocellular carcinoma imaging. Diagn Interv Imaging. 2023;104(1):24–36.
36. Akai H, Yasaka K, Kunimatsu A, et al. Predicting prognosis of resected hepatocellular carcinoma by radiomics analysis with random survival forest. Diagn Interv Imaging. 2018;99(10):643–51.
37. Wang XH, Long LH, Cui Y, et al. MRI-based radiomics model for preoperative prediction of 5-year survival in patients with hepatocellular carcinoma. Br J Cancer. 2020;122(7):978–85.
38. Bo Z, Chen B, Zhao Z, et al. Prediction of response to lenvatinib monotherapy for unresectable hepatocellular carcinoma by machine learning radiomics: A multicenter cohort study. Clin Cancer Res. 2023;29(9):1730–40.
39. Li X, Sun W, Ding X, Li W, Chen J. Prognostic model of immune checkpoint inhibitors combined with anti-angiogenic agents in unresectable hepatocellular carcinoma. Front Immunol. 2022;13:1060051.
40. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging-how-to guide and critical reflection. Insights Imaging. 2020;11(1):91.
41. Liu QP, Hu. Radiogenomic association of deep MR imaging features with genomic profiles and clinical characteristics in breast cancer. Biomark Res. 2023;11(1):9.
42. Fusco R, Granata V, Simonetti I, et al. An informative review of radiomics studies on Cancer imaging: the main findings, challenges and limitations of the methodologies. Curr Oncol. 2024;31(1):403–24.
43. Ichikawa T, Saito K, Yoshioka N, et al. Detection and characterization of focal liver lesions: a Japanese phase III, multicenter comparison between Gadoxetic acid disodium-enhanced magnetic resonance imaging and contrast-enhanced computed tomography predominantly in patients with hepatocellular carcinoma and chronic liver disease. Invest Radiol. 2010;45(3):133–41.
44. Tsuboyama T, Onishi H, Kim T, et al. Hepatocellular carcinoma: hepatocyte-selective enhancement at Gadoxetic acid-enhanced MR imaging–correlation with expression of sinusoidal and canalicular transporters and bile accumulation. Radiology. 2010;255(3):824–33.
45. Piana G, Trinquart L, Meskine N, Barrau V, Beers BV, Vilgrain V. New MR imaging criteria with a diffusion-weighted sequence for the diagnosis of hepatocellular carcinoma in chronic liver diseases. J Hepatol. 2011;55(1):126–32.
46. Cha B, Yu JH, Jin YJ, Suh YJ, Lee JW. Survival outcomes according to body mass index in hepatocellular carcinoma patient: analysis of nationwide Cancer registry database. Sci Rep. 2020;10(1):8347.
47. Vauthey JN, Lauwers GY, Esnaola NF, et al. Simplified staging for hepatocellular carcinoma. J Clin Oncol. 2002;20(6):1527–36.
48. Yau T, Tang VY, Yao TJ, Fan ST, Lo CM, Poon RT. Development of Hong Kong liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma. Gastroenterology. 2014;146(7):1691–e17001693.
49. Zhou Z-H. Model Selection and Evaluation. Machine Learning, Z.-H. Zhou, Editor. Singapore, Springer Singapore. 2021:25–55.

Publisher's note