

CT Radiomics Nomogram for the Preoperative Prediction of Severe Post-Hepatectomy Liver Failure in Patients with Huge (≥ 10 cm) Hepatocellular Carcinoma

Fei Xiang

Zhejiang University School of Medicine Second Affiliated Hospital

Xiaoyuan Liang

Zhejiang University School of Medicine Second Affiliated Hospital

Lili Yang

Zhejiang University School of Medicine First Affiliated Hospital

Xingyu Liu

Zhejiang University School of Medicine Second Affiliated Hospital

Sheng Yan (✉ shengyan@zju.edu.cn)

<https://orcid.org/0000-0002-4153-3546>

Research

Keywords: Hepatocellular carcinoma, Liver failure, Radiomics, Nomogram, Decision tree

Posted Date: July 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-732725/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at World Journal of Surgical Oncology on December 1st, 2021. See the published version at <https://doi.org/10.1186/s12957-021-02459-0>.

Abstract

Background To establish a radiomics-based nomogram for predicting severe (grade B or C) post-hepatectomy liver failure (PHLF) in patients with huge (≥ 10 cm) hepatocellular carcinoma (HCC).

Methods 186 patients with huge HCC ($n = 131$ for training dataset and $n = 55$ for test dataset) who underwent curative hepatic resection were included. The least absolute shrinkage and selection operator approach was applied to develop the radiomics signature for grade B or C PHLF prediction in the training dataset. A multivariable logistic regression model was used by incorporating radiomics signature and other clinical predictors to establish a radiomics nomogram. A decision tree was created to stratify the risk for severe PHLF.

Results The radiomics signature consisting of nine features predicted severe PHLF with an AUC of 0.766 and 0.745 in the training and test datasets, respectively. The radiomics nomogram was generated by integrating the radiomics signature, the extent of resection and model for end-stage liver disease (MELD) score. The nomogram exhibited satisfactory discrimination and calibration, with an AUC of 0.842 and 0.863 in the training and test datasets, respectively. Decision tree split patients into 3 risk classes: low-risk patients with radiomics score < -0.247 and MELD score < 10 or radiomics score ≥ -0.247 and underwent partial resections; intermediate-risk patients with radiomics score < -0.247 but MELD score ≥ 10 ; high-risk patients with radiomics score ≥ -0.247 and underwent extended resections.

Conclusions The radiomics nomogram was able to predict severe PHLF in huge HCC patients. Decision tree may be useful in surgical decision-making for huge HCC hepatectomy.

Introduction

Due to the high prevalence of hepatitis B virus (HBV) infection, China has experienced high rates of hepatocellular carcinoma (HCC) incidence, accounting for approximately half of cases worldwide [1, 2]. In these cases, huge HCC (≥ 10 cm) is common due to a lack of awareness for early surveillance and detection. Studies have shown a relatively satisfactory overall survival in selected patients with huge HCC hepatectomy [3–5]. However, patients with huge HCC often require to undergo major or extended hepatic resection, which puts them at high risk of post-hepatectomy liver failure (PHLF).

PHLF is not only a predominant cause of postoperative mortality, with reported mortality as high as 50% [6], but is also associated with a prolonged hospital stay, compromised long-term overall survival, and increased costs in patients undergoing this surgical procedure. To prevent PHLF, the appropriate selection of patients to undergo hepatectomy through a detailed assessment of liver function is paramount. Numerous methods have been used to predict PHLF, including clinical parameters and scoring systems [7–9], dynamic quantitative liver function tests [10, 11], and volume of the liver remnant [12, 13]. However, the predictive outcomes are variable, and no single method alone can accurately predict PHLF. Therefore, it may be possible to establish comprehensive models by combining multiple methods.

An emerging methodology named Radiomics through the high-throughput extraction of imaging features, including features based on intensity, shape, texture, and higher-order features, which has the potential to characterize diseases and guide clinical decision making. It has been initially applied in oncology studies, but gradually in other non-oncologic diseases [14]. Recent studies demonstrate that radiomics has improved the accuracy in the diagnosis of liver fibrosis and cirrhosis and could be potentially used to assess liver function [15, 16].

Therefore, we aimed to establish a CT-based radiomics signature and a nomogram by combining radiomics features and independent clinical factors for predicting severe (grade B or C) PHLF in patients with huge HCC.

Materials And Methods

Patients

From January 2012 to December 2020, a total of 1267 patients with HCC underwent hepatic resection in our hospital. Of these, 254 patients with huge HCC who underwent curative surgical resection were recruited. However, 68 patients were excluded, and 186 patients who met the following inclusion criteria, were enrolled into this study: (1) patients who did not receive any treatment before surgery; (2) Liver function was classified as Child-Pugh grade A or B; (3) Eastern Cooperative Oncology Group (ECOG) performance score 0-2; (4) Undergone enhanced CT scanning within 7 days before surgery; (5) histologically confirmed HCC. The detailed recruitment process is presented in Fig. 1. Then, patients were divided into training and test datasets at a ratio of 7:3. The training dataset was used to construct prediction model, and the test dataset was used to validate the performance of model. The Ethics Review Board of the Second Affiliated Hospital of Zhejiang University School of Medicine approved this study (No. 2021-0376).

Clinical Characteristics

Baseline demographic, clinical, and laboratory characteristics (including liver and kidney function tests, platelet count, blood coagulation index, serum alpha-fetoprotein level), clinical grading scores were collected. The operative variables (including surgical methods, intraoperative blood loss, intraoperative blood transfusion, and intraoperative vascular occlusion methods) that correlate to PHLF were also recorded.

Diagnosis and definitions

PHLF was diagnosed according to the International Study Group of Liver Surgery (ISGLS) criteria [17]. The INR was set at 1.5 and the bilirubin level of more than 20 micromol/L (1.2mg/dL). The severity of PHLF was divided into 3-classes according to the clinical management: grade A, no further clinical management necessary; grade B, requires an active therapeutic intervention without invasive approach;

grade C, invasive approach. We defined grade B and C PHLF as severe PHLF and put as primary outcome since grade A PHLF does not require any additional management.

CT scan acquisition

Multi-detector CT systems (16-slice SOMATOM Perspective, SIEMENS; 16-slice SOMATOM Sensation, SIEMENS, Germany) were scanning devices in our department. Dynamic contrast-enhanced CT imaging were obtained following the administration of iodinated contrast material (Iohexol, GE Healthcare, USA) at 3.0 mL/sec. Scanning parameters were 120 KV, 160 mAs; rotation time 0.5 s; 350 mm×350 mm field of view; matrix of 388×388; slice thickness, 3 mm. Arterial phase and portal phase images were obtained at 40s and 72s after injection of contrast medium.

Image segmentation and radiomics features extraction

The region of interest (ROI) was drawn manually using the freely available application ITK-SNAP (version 3.6.0). ROI was delineated in the liver along the border of the whole liver parenchyma by avoiding major blood vessels, focal lesions, and artifacts on the portal phase images. Features were extracted from each segmented ROI, divided into textual and non-textual features using PyRadiomics [18], an open-source python package for medical imaging.

To get reproducible radiomics features, standardized computation of radiomics features is necessary [19]. In our study, the sitkBSpline interpolation was applied to resample the images with a pixel size of 1×1mm. Voxel intensities were discretized

using a bin-width of 25 HU. 788 radiomics features were extracted from the liver ROI, including 18 original first-order histogram features, 14 original shape features, 68 original textural features and 688 high-order wavelet features. The list of radiomics features is shown in the supplemental table 1.

Inter-observer and intra-observer agreement

To ensure reproducibility, CT images of 20 patients were randomly selected and re-segmented by reader 1 (X.F. with 7 years' experience in liver imaging) at an interval of 2 week and reader 2 (Y.L.L. with 8 years' experience in liver imaging) independently performed the segmentation. The intra-observer reproducibility and inter-observer reliability of features extraction were assessed by using intra- and inter-class correlation coefficients (ICCs). Features with ICC > 0.75 represent a good agreement and were retained.

Feature selection and radiomics signature construction

The extracted radiomics features were normalized by Z-scores method. Radiomics features with ICCs lower than 0.75 were excluded. Univariate analyses were conducted using univariate logistic regression analysis. Features were considered to be associated with severe PHLF when p values < 0.1. Based on the selected features, the least absolute shrinkage and selection operator (LASSO) algorithm was conducted to identify significant features with nonzero coefficients. The penalty parameter (λ) was optimized

through the tenfold cross-validation method. A radiomics signature was constructed by summing the selected features multiplied by their coefficients. The area under the receiver operating characteristic curve (AUC area under the ROC curve) was calculated for assessing the predictive ability of the established radiomics signature.

Development of the clinical-radiomic nomogram

To develop a comprehensive clinical-radiomic nomogram, the clinical characteristics and radiomics signature were analyzed by univariate logistic regression. Significant factors ($p < 0.05$) were used to build the multivariate logistic model. Finally, the clinical-radiomic nomogram model integrated the clinical predictors and radiomics signature was built in the training dataset.

Assessing the accuracy of Nomogram model and comparison with conventional methods

We determined the discriminatory ability of the nomogram model, by comparing the radiomics signature, albumin-bilirubin score (ALBI) score, the model for end-stage liver disease (MELD) score, Child-Pugh score with the areas under the receiver operating characteristic curve (AUC). DeLong's test was used to compare the nomogram model with conventional methods based on the AUC values in both datasets. To evaluate the consistency of the nomogram, we plotted a calibration curve with the Hosmer-Lemeshow goodness-of-fit test.

Clinical Use

To assist in surgical decision-making, a decision tree for safe huge HCC hepatectomy was built based on the associated risk factors. In addition, to evaluate the clinical usefulness of the nomogram model, radiomics signature, MELD, ALBI, and Child-Pugh scores, decision curve analysis (DCA) was conducted to evaluate the net benefits across a variety of threshold risks.

Statistical Analysis

The radiomics analysis workflow is shown in Fig. 2. Continuous variables and categorical variables were compared by Mann-Whitney U test and chi-square test, respectively. We considered two-tailed values of $p < 0.05$ as statistically significant for all analyses. All analyses were conducted with R software (version 3.6.1).

Results

Patient Demographic

A total of 186 patients (71 men, 66 women) were included in present study. The patients were assigned into training ($n = 131$) and test datasets ($n = 55$) at a ratio of 7:3. The clinical variables did not differ significantly between the 2 datasets, except for HBsAg positive ($P = 0.044$) and intraoperative blood

transfusion ($P = 0.002$). The percentage of severe PHLF was 31.3% ($n = 41$) and 23.64% ($n = 13$) in the training and test dataset, respectively. The baseline characteristics are presented in Table 1.

Table 1
Comparison of patient demographics and clinicopathological features of the two datasets

Variable	Training dataset (n = 131)	Test dataset (n = 55)	p value
Sex, n (%)			0.363
male	118 (90.08)	47 (85.46)	
female	13 (9.92)	8 (14.54)	
Age, years	58.69 ± 11.13	53.25 ± 17.19	0.387
BMI, kg/m ²	23.97 ± 3.79	23.89 ± 4.01	0.963
HBsAg positive, n (%)	104 (79.38)	36 (65.45)	0.044
Child-Pugh class, n (%)			0.923
A	115 (87.79)	48 (87.27)	
B	16(12.21)	7(12.72)	
AFP, median (IQR), ng/ml	38.7 (4-3849)	171.5 (9-9312)	0.717
Liver function tests			
ALB (g/L)	38.07 ± 4.84	37.55 ± 6.57	0.837
TBIL (µmol/L)	16.07 ± 8.75	18.94 ± 12.43	0.541
ALT (U/L)	48.00 ± 53.05	58.88 ± 47.64	0.641
AST (U/L)	81.23 ± 63.99	75.75 ± 66.02	0.853
GGT(U/L)	130.62 ± 109.54	167.13 ± 166.12	0.549
Cr (µmol/L)	50.46 ± 11.81	55.63 ± 12.83	0.358
PLT(10 ⁹ /L)	202.92 ± 83.81	198.50 ± 63.50	0.900
PT (s)	13.45 ± 1.32	12.59 ± 1.26	0.155
INR	1.04 ± 0.12	1.02 ± 0.96	0.631
Tumor size, mm	125.54 ± 25.78	127.13 ± 52.94	0.930
Cirrhosis, n (%)	76 (58.02)	29 (52.72)	0.507

BMI, body mass index; HBsAg, hepatitis B surface antigen; AFP, alpha fetoprotein; SD, standard deviation; ALB, albumin; TBIL, total bilirubin; ALT, alanine aminotransferase; AST, aspartate transaminase; GGT, γ-glutamyl transpeptidase; PLT, platelets; PT, prothrombin time; INR, international normalized ratio; MELD, model for end-stage liver disease; ALBI, albumin to bilirubin ration index; PHLF, posthepatectomy liver failure. † Median (range); ‡ Median (IQR).

Variable	Training dataset (n = 131)	Test dataset (n = 55)	p value
Extent of resection			0.515
Extended (≥ 4 segments)	84 (64.12)	38 (69.09)	
Partial (< 4 segments)	47 (35.88)	17 (30.91)	
Conventional predictive models			
Child-Pugh score†	5 (5–8)	5 (5–8)	0.463
MELD score†	7 (6–15)	7 (6–13)	0.568
ALBI score‡	-2.47 (-0.51~ -3.42)	-2.55 (-1.09~ -3.32)	1.000
Intraoperative blood loss, ml ‡	400 (200–800)	500 (300–800)	0.199
Intraoperative blood transfusion, n (%)			
Yes	87 (66.4%)	23 (41.8%)	0.002
No	44 (33.6%)	32 (58.2%)	
Pringle maneuver, n (%)			0.872
Yes	65 (49.6%)	28 (50.9%)	
No	66 (50.4%)	27 (49.1%)	
PHLF (B/C), n (%)			0.293
Yes	41 (31.30)	13 (23.64)	
No	90 (68.70)	42 (76.36)	
Postoperative mortality, n (%)	5 (3.8%)	3 (5.4%)	0.696
BMI, body mass index; HBsAg, hepatitis B surface antigen; AFP, alpha fetoprotein; SD, standard deviation; ALB, albumin; TBIL, total bilirubin; ALT, alanine aminotransferase; AST, aspartate transaminase; GGT, γ -glutamyl transpeptidase; PLT, platelets; PT, prothrombin time; INR, international normalized ratio; MELD, model for end-stage liver disease; ALBI, albumin to bilirubin ration index; PHLF, posthepatectomy liver failure. † Median (range); ‡ Median (IQR).			

Radiomics Signature Construction

Of 788 extracted radiomics features, 165 features were eliminated due to ICC lower than 0.75. Further, univariate logistic regression was used to select PHLF-associated features. Then, 30 features remained and were subjected to Lasso regression algorithm to screen the critical features and construct radiomics signature. Finally, 9 features with non-zero coefficients were screened by LASSO approach in the training dataset (Fig. 3A and 3B). Among the 9 features, 2 features were original shape features, and the

remaining were wavelet features. The radiomics signature was constructed using the 9 features, and the radiomics score computed as:

$$\begin{aligned} \text{Radscore} = & -0.93044761 + 0.20910827 * \text{original_shape_Maximum2DDiameterSlice} \\ & + 0.04625660 * \text{original_shape_SurfaceVolumeRatio} - 0.08693156 * \text{HHH_glszm_ZoneVariance} \\ & - 0.44200827 * \text{HHL_firstorder_Median} \\ & - 0.42800711 * \text{HHL_gldm_DependenceNonUniformityNormalized} - 0.04493315 \\ & * \text{HLH_firstorder_Maximum} - 0.35475442 * \text{HLH_glcm_ClusterProminence} + 0.01233872 \\ & * \text{LHH_glszm_LowGray} \\ & - 0.36996067 * \text{LLH_glszm_GrayLevelNonUniformity} \end{aligned}$$

In patients with PHLF, Radscore (median [range]) was significantly higher than those without PHLF in the training dataset (- 0.290 [-2.443~1.462] vs - 1.067 [-3.686~0.404], respectively, $P < 0.001$). The same trend was observed in the test dataset (- 0.536 [-1.375~1.461] vs - 0.930 [-3.875~1.138], respectively, $P = 0.007$). The distributions of Radscore for each patient in the training and test dataset are shown in supplemental Fig. 1.

Development The Clinical-radiomics Nomogram And Comparison With Conventional Models

Univariate and multivariate logistic regression analysis incorporating Radscore, MELD score, the extent of resection as significant predictive factors of severe PHLF (Table 2). An individualized nomogram model was developed using these significant independent risk factors (Fig. 4). The nomogram showed good discrimination ability, with a mean AUC of 0.842 (95% confidence interval (CI): 0.761–0.922) in the training dataset (Fig. 5A), and 0.863 (95% CI, 0.750–0.975) in the test dataset (Fig. 5B). In the training dataset, the nomogram model had a significantly higher AUC than that of Child-Pugh score ($P < 0.001$), MELD score ($P = 0.001$), and ALBI score ($P < 0.001$). The test dataset presented similar results (nomogram vs Child-Pugh score, MELD score, ALBI score; $P < 0.001$, $P = 0.002$, $P = 0.02$, respectively). The calibration curve demonstrated good agreement between the predicted and actual observation in the training and test datasets (Fig. 5C and 5D). The p value by Hosmer-Lemeshow test was 0.397 in the training dataset and 0.285 in the test dataset, suggesting a good fit of the nomogram.

Table 2

Univariable and multivariable logistic regression analyses of risk factors for severe PHLF in the training dataset

Variables	Univariate analysis			Multivariate analysis		
	OR	95%CI	P-value	OR	95% CI	P-value
Age	1.012	0.962–1.066	0.639			
Sex, male vs female	2.470	0.456–13.379	0.294			
BMI (≥ 25 vs < 25)	0.628	0.182–2.160	0.460			
HBV infection	1.703	0.419–6.922	0.457			
TBIL	0.949	0.871–1.034	0.232			
ALB	0.965	0.844–1.103	0.603			
PT	1.512	0.809–2.829	0.195			
INR (per 0.1 increase)	0.707	0.327–1.529	0.378			
PLT	0.999	0.993–1.006	0.820			
Tumor size	0.996	0.976–1.018	0.741			
Cirrhosis	0.660	0.177–2.465	0.537			
Extent of resection (Extended vs Partial)	4.903	1.381–17.405	0.014	4.483	1.591–12.633	0.005
blood loss (≥ 800 vs < 800 ml)	0.250	0.054–1.161	0.077			
blood transfusion	2.609	0.517–13.164	0.245			
Pringle maneuver	1.217	0.394–3.763	0.733			
Child-Pugh score	2.803	0.239–32.839	0.412			
MELD score	1.891	1.093–3.271	0.023	1.589	1.189–2.124	0.002
ALBI score	0.955	0.355–2.725	0.931			
Radscore (per 0.1 increase)	1.144	1.068–1.224	<0.001	1.139	1.066–1.216	<0.001

Clinical Use

On the basis of Radscore, MELD score and the extent of resection, decision tree stratified the risk for severe PHLF into the following three classes (Fig. 6A): low-risk patients with radiomics score < -0.247 and MELD score < 10 or, radiomics score ≥ -0.247 and underwent partial resections, the severe PHLF rate was about 18%; intermediate-risk patients with radiomics score < -0.247 but MELD score ≥ 10 , the severe PHLF rate was about 50%; high-risk patients with radiomics score ≥ -0.247 and underwent extended resections, the severe PHLF rate was about 82%. DCA (Fig. 6B) showed that the nomogram had high potential for clinical application with wider threshold probabilities compared with conventional models.

Discussion

The present study established a radiomics signature to preoperative predict severe PHLF in patients with huge HCC hepatectomy and further develop an individual clinical-radiomic nomogram comprising the radiomics signature and clinical predictors. The nomogram model integrated three predictive variables that reflect the preoperative clinical essentials, which yielded good predictive ability for severe PHLF. On the basis of radiomics score, MELD score, the extent of resection, a decision tree was built, and the whole series could be split into 3 risk classes.

In recent years, improved hepatic resection techniques and expanding indications have increased the use of extensive liver resection, leading to a higher risk of PHLF. Single-center studies report PHLF risk ranging between 25.8% and 35.3%, and the severe PHLF ranged between 11.3% and 28% [20–23]. Due to the large tumor diameters and major vascular invasion, approximately 62–80% of patients with huge HCC patients undergo major or extensive liver resection leading to morbidity and mortality rates ranging between 10.9–43.6% and 4.2–18.1% [24–27], respectively. Therefore, establishing an individualized PHLF prediction model in patients with a huge HCC is critical.

Radiomics is a high-throughput data mining method that involves extracting features from medical images and has been extensively used in oncology studies. Radiomics quantitatively assesses tumor heterogeneity by reflecting the distribution of gray level values and spatial arrangement of the pixels. Besides, it has gradually been applied to non-oncology diseases. In chronic liver diseases, studies have shown the potential benefits of radiomics in assessing the heterogeneity of the liver parenchyma, which can reflect architectural disturbance to predict liver function [28]. For example, radiomics of shear wave elastography, MRI, CT have been used to quantitatively assess liver fibrosis and have shown favorable diagnostic accuracy regardless of the image modalities [15, 16, 29]. Further, some studies use radiomic features to predict PHLF. In a study by Pak [30], the liver parenchyma in patients with PHLF demonstrate a more heterogeneous appearance, and pixel intensities vary widely. In contrast, normal patients have a more homogenous appearance. With the help of machine learning, significant features can be selected and established as radiomics signatures. In a study by Cai [31], the CT-based higher-order wavelet features were selected and radiomics score was calculated for the prediction of PHLF with an AUC of 0.82 and 0.76 in the training and test group, respectively. Zhu et al [32] reported an MRI-based radiomics model which combined first order and texture features associated with PHLF, resulting in an accuracy of 80.9% in the test dataset. Similarly, Chen et al [33] developed a liver failure model by incorporating PLT

count, tumor size and radiomics features from Gd-EOB-DTPA-enhanced MRI images, which shown a better performance than conventional clinical model. We reviewed these studies and compared the outcomes in the supplemental table 2. Unlike these studies, grade A PHLF was not included in our study due to grade A PHLF was asymptomatic and did not need to pay for specific treatments. We think that predicting symptomatic grade B or C PHLF is more helpful to guide surgeons for decision-making.

In this study, various prediction models were compared. The conventional scoring systems in combination with laboratory biochemical parameters have shown certain reference values. However, conventional scores are primarily used to roughly assess liver function. Single score often does not fully capture the liver function status. To accurately predict PHLF, integrated models that consider patient, liver, and surgery related risk factors are needed [34]. In view of this, a combined nomogram model which integrated radiomics score and other clinical factors was developed. In our nomogram model, three independent indicators including radiomics score, MELD score, and the extent of hepatectomy were incorporated in multivariate logistic regression. Radiomics score was calculated by wavelet and liver shape features. The wavelet features exhibited higher weights in the radiomics score and evidence shown that wavelet transformation can further reflect the spatial heterogeneity across multiple dimensions [35]. Even though the MELD score has been criticized due to several reasons, evidence shows that it presents good predictive accuracy in severe liver diseases [8]. Besides, numerous studies demonstrate that it is a significant factor in predicting PHLF and can be integrated with other factors to enhance the prediction accuracy [36, 37]. Extended hepatectomy was a well-established risk factor for PHLF [22]. Moreover, the incidence of PHLF is reported to increase with the number of segments resected [38].

To further assist in clinical decision-making, a decision tree was built by using these factors as determinants for risk stratification. As the root node of the decision tree, radiomics score is the most important factor associated with severe PHLF which in accordance with the results of multivariate regression analysis. The cutoff of radiomics score was -0.247 . Patients whose radiomics score greater than -0.247 and underwent extended resections which classified as high risk experienced with 82.1% probability risk for severe PHLF. For these patients, the surgical indication should be made cautiously, another option such as local treatments must be considered. For patients at intermediate risk whose radiomics score < -0.247 but MELD score ≥ 10 , additional clinical and diagnostic information will be required to determine whether hepatectomy will provide added benefit. For patients at low risk, clinical decision making is very straightforward if evidence shown that patients can benefit from surgery. We think the decision tree model is easy to understand and manipulated by generating a set of if-then rules along with the classification results which can simplify the decision-making process.

Limitations of this study include the retrospective nature that may leading to selection bias. Another limitation was lack of external validation using data from other hospitals. Therefore, further prospective multi-institutional studies should be conducted to assess the value of the radiomics nomogram in predicting severe PHLF.

Conclusion

The proposed clinical-radiomics nomogram, which integrates the radiomics signature and clinical predictors has satisfactory discrimination and calibration power in predicting severe PHLF. The radiomics nomogram in combination with the decision tree potentially provide alternative methods for clinical prediction and decision-making in patients with huge HCC hepatectomy. We hypothesize that this radiomics nomogram and decision tree play an important complementary role in predicting severe PHLF in patients with huge HCC after hepatectomy and improve the patient-selection criteria.

Abbreviations

PHLF post-hepatectomy liver failure

HCC hepatocellular carcinoma

AUC the areas under the receiver operating characteristic curve

MELD the extent of resection and model for end-stage liver disease

HBV hepatitis B virus

ECOG Eastern Cooperative Oncology Group

ISGLS the International Study Group of Liver Surgery

ROI the region of interest

ICC intra- and inter-class correlation coefficients

LASSO the least absolute shrinkage and selection operator

ALBI albumin-bilirubin

DCA decision curve analysis

Radscore radiomics score

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine (No. 2021-0376). Written informed consent was obtained from all participants.

Consent for publication

We have obtained the consent of all patients to use clinical data, test data, and graphical data for conference or journal presentation.

Availability of data and materials

The raw data of this paper are available upon reasonable request to the corresponding author.

Competing interests

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

Funding

This study was supported by grants from the National Natural Science Foundation of China (No.81572975) and Key research and development project of science and technology department of Zhejiang (No.2015C03053)

Authors' contributions

Fei Xiang and Sheng Yan were responsible for the conception of the work. Fei Xiang, Xiaoyuan Liang and Xingyu Liu obtained the data. Fei Xiang and Lili Yang segmented the liver images. Fei Xiang and Xiaoyuan Liang analyzed the data. Fei Xiang wrote the manuscript. Sheng Yan critically revised the manuscript. All authors are accountable for the contents of this work. The authors read and approved the final manuscript.

Acknowledgements

None.

References

1. Liu J, Liang W, Jing W, Liu M. Countdown to 2030: eliminating hepatitis B disease, China. *Bull World Health Organ.* 2019;**97**(3):230-238. doi: 10.2471/BLT.18.219469.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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):209-249. doi:10.3322/caac.21660.
3. Fang Q, Xie QS, Chen JM, Shan SL, Xie K, Geng XP, et al. Long-term outcomes after hepatectomy of huge hepatocellular carcinoma: A single-center experience in China. *Hepatobiliary Pancreat Dis Int.* 2019;**18**(6):532-537. doi:10.1016/j.hbpd.2019.09.001.
4. Pandey D, Lee KH, Wai CT, Waghlikar G, Tan KC. Long term outcome and prognostic factors for large hepatocellular carcinoma (10 cm or more) after surgical resection. *Ann Surg Oncol.* 2007;**14**(10):2817-23. doi:10.1245/s10434-007-9518-1.

5. Yamashita Y, Taketomi A, Shirabe K, Aishima S, Tsuijita E, Morita K, et al. Outcomes of hepatic resection for huge hepatocellular carcinoma (≥ 10 cm in diameter). *J Surg Oncol*. 2011;**104**(3):292-8. doi:10.1002/jso.21931
6. Allard MA, Adam R, Bucur PO, Termos S, Cunha AS, Bismuth H, et al. Posthepatectomy portal vein pressure predicts liver failure and mortality after major liver resection on noncirrhotic liver. *Ann Surg*. 2013;**258**(5):822-9; discussion 829-30. doi:10.1097/SLA.0b013e3182a64b38
7. Kok B, Abrales JG. Child-Pugh Classification: Time to Abandon? *Semin Liver Dis*. 2019;**39**(1):96103. doi:10.1055/s-0038-1676805.
8. Cucchetti A, Ercolani G, Vivarelli M, Cescon M, Ravaioli M, La Barba G, et al. Impact of model for end-stage liver disease (MELD) score on prognosis after hepatectomy for hepatocellular carcinoma on cirrhosis. *Liver Transpl*. 2016;**12**(6):966-0doi:10.1002/lt.20761.
9. Fagenson AM, Gleeson EM, Pitt HA, Lau KN. Albumin-Bilirubin Score vs Model for End-Stage Liver Disease in Predicting Post-Hepatectomy Outcomes. *J Am Coll Surg*. 2020;**230**(4):637-645. doi: 10.1016/j.jamcollsurg.2019.12.007.
10. Ohwada S, Kawate S, Hamada K, Yamada T, Sunose Y, Tsutsumi H, et al. Perioperative real-time monitoring of indocyanine green clearance by pulse spectrophotometry predicts remnant liver functional reserve in resection of hepatocellular carcinoma. *Br J Surg*. 2006;**93**(3):339-46. doi:10.1002/bjs.5258.
11. de Graaf W, van Lienden KP, Dinant S, Roelofs JJ, Busch OR, Gouma DJ, et al. Assessment of future remnant liver function using hepatobiliary scintigraphy in patients undergoing major liver resection. *J Gastrointest Surg*. 2010;**14**(2):369-78. doi:10.1007/s11605-009-1085-2.
12. Truant S, Oberlin O, Sergent G, Lebuffe G, Gambiez L, Ernst O, et al. Remnant liver volume to body weight ratio $> \text{ or } = 0.5\%$: A new cut-off to estimate postoperative risks after extended resection in noncirrhotic liver. *J Am Coll Surg*. 2007; **204**(1):22-33. doi: 10.1016/j.jamcollsurg.2006.09.007.
13. Kishi Y, Abdalla EK, Chun YS, Zorzi D, Madoff DC, Wallace MJ, et al. Three hundred and one consecutive extended right hepatectomies: evaluation of outcome based on systematic liver volumetry. *Ann Surg*. 2009;**250**(4):540-8. doi:10.1097/SLA.0b013e3181b674df.
14. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016; **278**(2):563-77. doi:10.1148/radiol.2015151169.
15. Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut*. 2019;**68**(4):729-741. doi:10.1136/gutjnl-2018-316204.
16. Lubner MG, Malecki K, Kloke J, Ganeshan B, Pickhardt PJ. Texture analysis of the liver at MDCT for assessing hepatic fibrosis. *Abdom Radiol (NY)*. 2017;**42**(8):2069-2078. doi:10.1007/s00261-017-1096-5.
17. Rahbari NN, Garden OJ, Padbury R, Brooke-Smith M, Crawford M, Adam R, et al. Posthepatectomy liver failure: a definition and grading by the International Study Group of Liver Surgery (ISGLS).

- Surgery. 2011;**149**(5):713-24. doi: 10.1016/j.surg.2010.10.001
18. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017;**77**(21): e104-e107. doi: 10.1158/0008-5472.CAN-17-0339.
 19. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology.* 2020; **295**(2):328-338. doi:10.1148/radiol.2020191145.
 20. Asenbaum U, Kaczirek K, Ba-Ssalamah A, Ringl H, Schwarz C, Waneck F, et al. Post-hepatectomy liver failure after major hepatic surgery: not only size matters. *Eur Radiol.* 2018;**28**(11):4748-4756. doi:10.1007/s00330-018-5487-y.
 21. Golriz M, Ghamarnejad O, Khajeh E, Sabagh M, Mieth M, Hoffmann K, et al. Preoperative Thrombocytopenia May Predict Poor Surgical Outcome after Extended Hepatectomy. *Can J Gastroenterol Hepatol.* 2018; **2018**:1275720. doi:10.1155/2018/1275720.
 22. Truant S, El Amrani M, Skrzypczyk C, Boleslawski E, Sergent G, Hebbar M, et al. Factors associated with fatal liver failure after extended hepatectomy. *HPB (Oxford).* 2017;**19**(8):682-687. doi: 10.1016/j.hpb.2017.04.006.
 23. Chen X, Zhai J, Cai X, Zhang Y, Wei L, Shi L, et al. Severity of portal hypertension and prediction of postoperative liver failure after liver resection in patients with Child-Pugh grade A cirrhosis. *Br J Surg.* 2012; **99**(12):1701-10. doi:10.1002/bjs.8951.
 24. Goh BK, Kam JH, Lee SY, Chan CY, Allen JC, Jeyaraj P, et al. Significance of neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio and prognostic nutrition index as preoperative predictors of early mortality after liver resection for huge (≥ 10 cm) hepatocellular carcinoma. *J Surg Oncol.* 2016;**113**(6):621-7. doi:10.1002/jso.24197.
 25. Shrager B, Jibara GA, Tabrizian P, Schwartz ME, Labow DM, Hiotis S. Resection of large hepatocellular carcinoma (≥ 10 cm): a unique western perspective. *J Surg Oncol.* 2013;**107**(2):111-7. doi:10.1002/jso.23246.
 26. Chen XP, Qiu FZ, Wu ZD, Zhang BX. Chinese experience with hepatectomy for huge hepatocellular carcinoma. *Br J Surg.* 2004;**91**(3):322-6. doi:10.1002/bjs.4413.
 27. Lim C, Compagnon P, Sebah M, Salloum C, Calderaro J, Luciani A, et al. Hepatectomy for hepatocellular carcinoma larger than 10 cm: preoperative risk stratification to prevent futile surgery. *HPB (Oxford).* 2015;**17**(7):611-23. doi:10.1111/hpb.12416.
 28. Wei J, Jiang H, Gu D, Niu M, Fu F, Han Y, et al. Radiomics in liver diseases: Current progress and future opportunities. *Liver Int.* 2020;**40**(9):2050-2063. doi:10.1111/liv.14555.
 29. Park HJ, Lee SS, Park B, Yun J, Sung YS, Shim WH, et al. Radiomics Analysis of Gadoteric Acid-enhanced MRI for Staging Liver Fibrosis. *Radiology.* 2019;**292**(1):269. doi:10.1148/radiol.2019194012.
 30. Pak LM, Chakraborty J, Gonen M, Chapman WC, Do RKG, Groot Koerkamp B, et al. Quantitative Imaging Features and Postoperative Hepatic Insufficiency: A Multi-Institutional Expanded Cohort. *J*

- Am Coll Surg. 2018;**226**(5):835-843. doi: 10.1016/j.jamcollsurg.2018.02.001.
31. Cai W, He B, Hu M, Zhang W, Xiao D, Yu H, et al. A radiomics-based nomogram for the preoperative prediction of posthepatectomy liver failure in patients with hepatocellular carcinoma. *Surg Oncol.* 2019;**28**:78-85. doi: 10.1016/j.suronc.2018.11.013.
 32. Zhu WS, Shi SY, Yang ZH, Song C, Shen J. Radiomics model based on preoperative gadoxetic acid-enhanced MRI for predicting liver failure. *World J Gastroenterol.* 2020;**26**(11):1208-1220. doi:10.3748/wjg.v26.i11.1208.
 33. Chen Y, Liu Z, Mo Y, Li B, Zhou Q, Peng S, et al. Prediction of Post-hepatectomy Liver Failure in Patients With Hepatocellular Carcinoma Based on Radiomics Using Gd-EOB-DTPA-Enhanced MRI: The Liver Failure Model. *Front Oncol.* 2021;**10**:11:60 doi:10.3389/fonc.2021.605296.
 34. Lafaro K, Buettner S, Maqsood H, Wagner D, Bagante F, Spolverato G, et al. Defining Post Hepatectomy Liver Insufficiency: Where do We stand? *J Gastrointest Surg.* 2015;**19**(11):2079-92. doi:10.1007/s11605-015-2872-6.
 35. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. *Transl Lung Cancer Res.* 2017;**6**(1):86-91. doi:10.21037/tlcr.2017.01.04.
 36. Schadde E, Raptis DA, Schnitzbauer AA, Ardiles V, Tschuor C, Lesurtel M, et al. Prediction of Mortality After ALPPS Stage-1: An Analysis of 320 Patients From the International ALPPS Registry. *Ann Surg.* 2015;**262**(5):780-5; discussion 785-6. doi:10.1097/SLA.0000000000001450.
 37. Cescon M, Cucchetti A, Grazi GL, Ferrero A, Viganò L, Ercolani G, et al. Indication of the extent of hepatectomy for hepatocellular carcinoma on cirrhosis by a simple algorithm based on preoperative variables. *Arch Surg.* 2009;**144**(1):57-63; discussion 63. doi:10.1001/archsurg.2008.522.
 38. Viganò L, Torzilli G, Aldrighetti L, Ferrero A, Troisi R, Figueras J, et al. Stratification of Major Hepatectomies According to Their Outcome: Analysis of 2212 Consecutive Open Resections in Patients Without Cirrhosis. *Ann Surg.* 2020;**272**(5):827-833. doi:10.1097/SLA.0000000000004338.

Figures

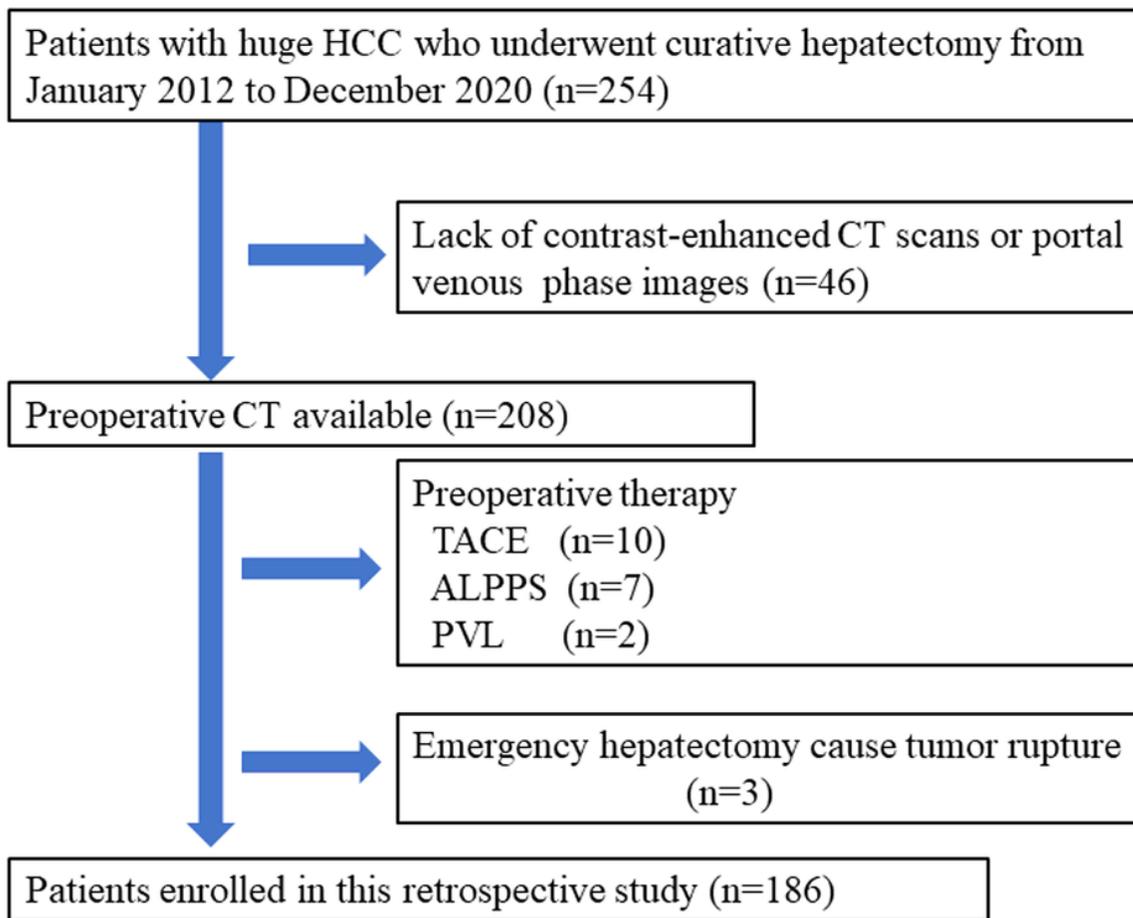


Figure 1

Flow chart of patients recruitment in this study. TACE, transarterial chemoembolization; ALPPS, associating liver partition and portal vein ligation for staged hepatectomy; PVL, portal vein ligation;

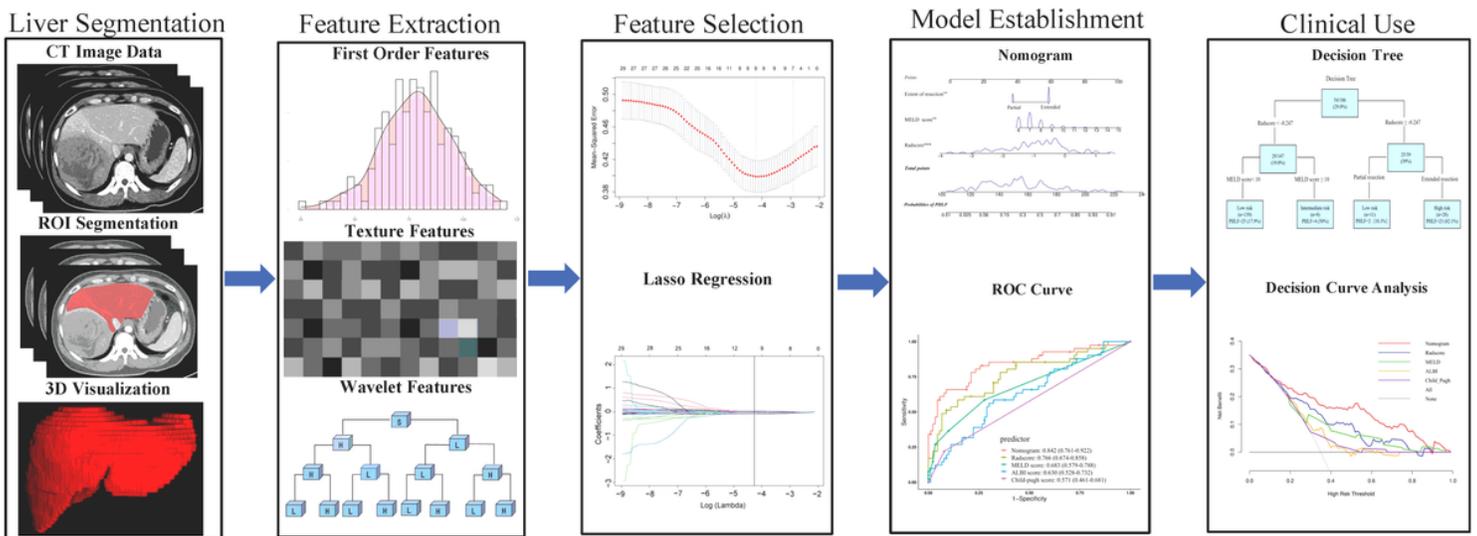


Figure 2

Workflow for the radiomics process. After CT images were acquired, segmentation of liver parenchyma was performed. The extracted radiomics features include intensity, shape, texture features, and wavelet features. Nine radiomics features were selected by LASSO algorithm. A nomogram was built that incorporates radiomics signature and independent clinical predictors for individualized predicting severe PHLF. The discrimination ability of nomogram and conventional models were compared with the ROC curve analysis and quantified by the AUC. A decision tree was built to stratify the risk for severe PHLF into three classes. Clinical benefits of nomogram and conventional models were compared by decision curve analysis.

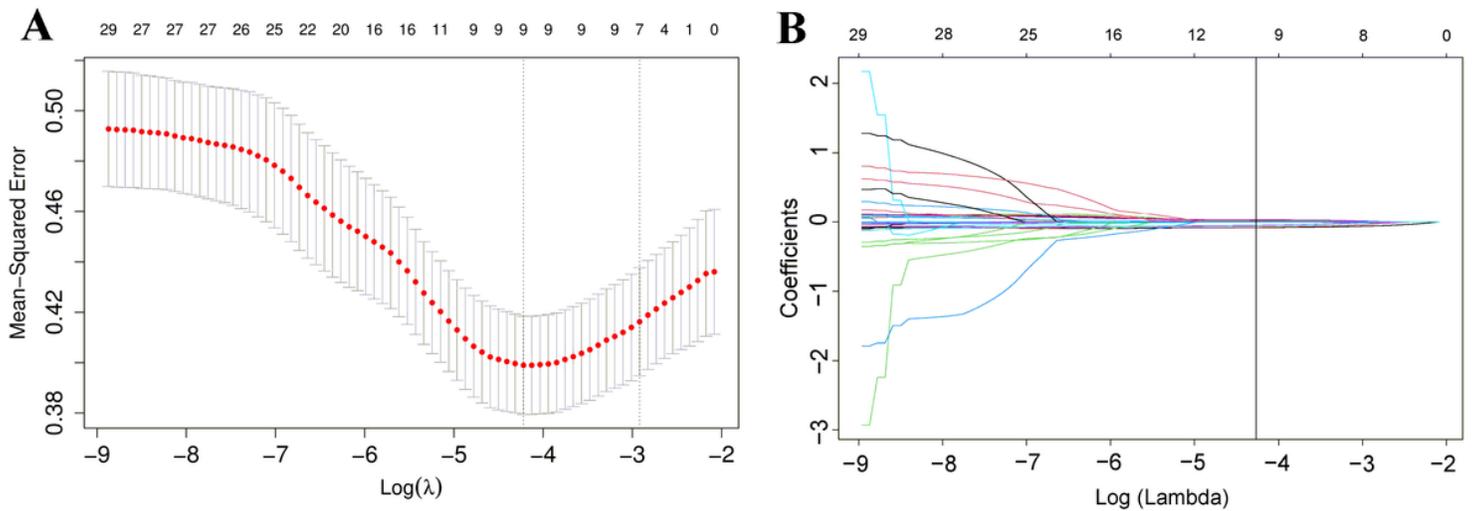


Figure 3

The LASSO algorithm was used to select predictive radiomics features. A. Tuning parameter (λ) in the LASSO model was selected by ten-fold cross-validation. The optimal λ value of 0.015 with $\log(\lambda)$ of -4.269 was chosen (at the minimum criteria). B. Coefficients of 30 features were shrunk with the penalty term increases. 9 features with non-zero coefficients were obtained with the optimal λ .

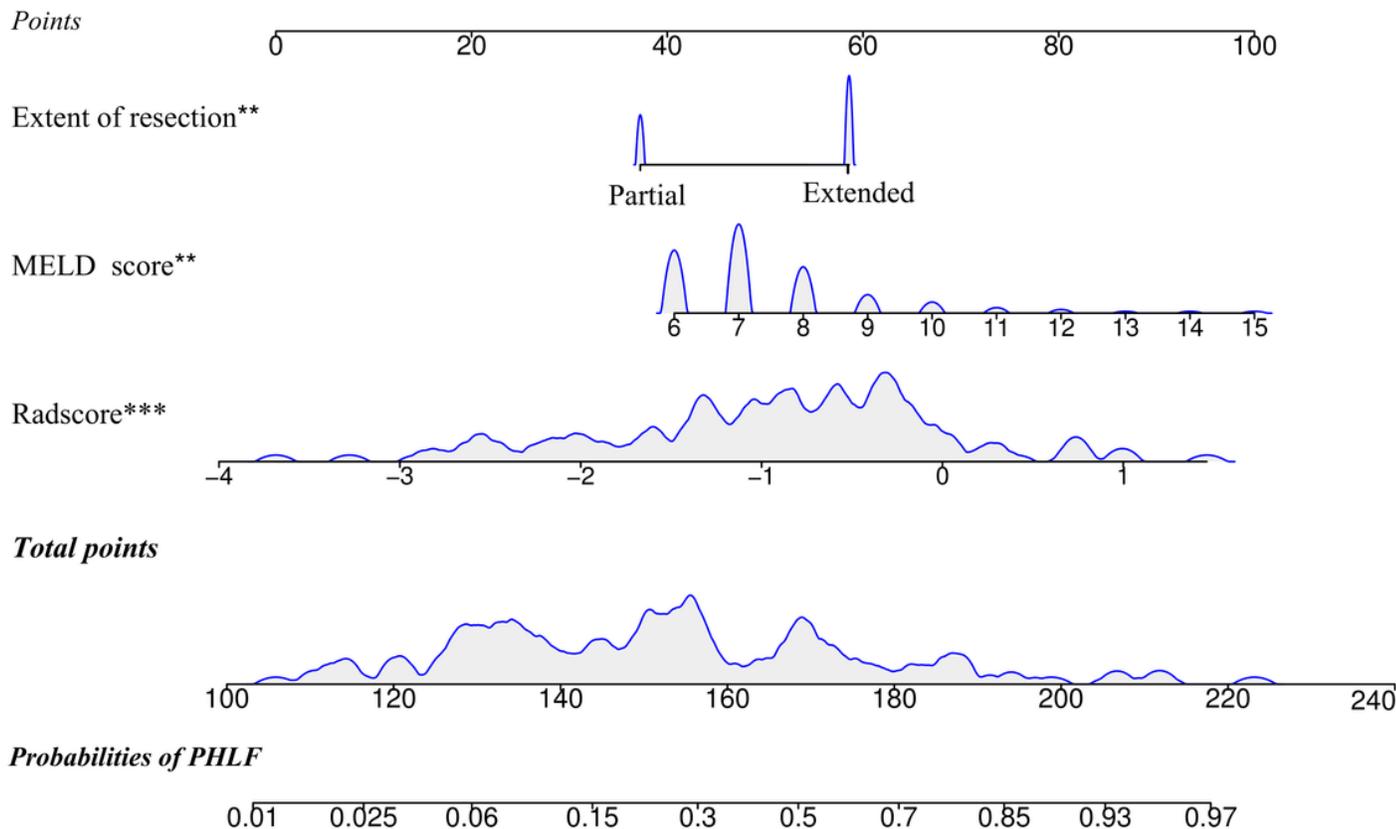


Figure 4

The radiomics nomogram developed by incorporating the Radscore, the MELD score and the extent of resection.

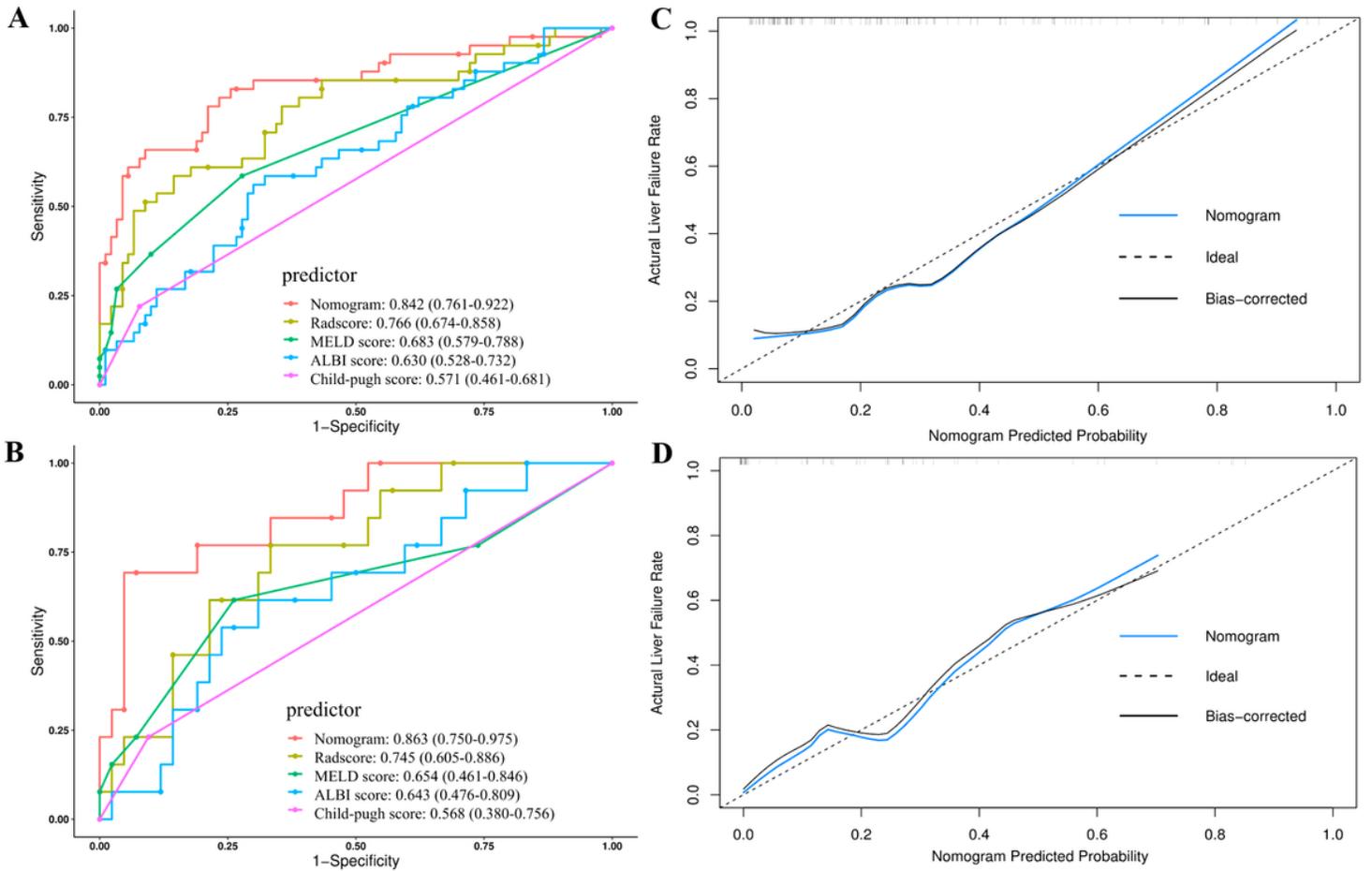


Figure 5

Assessing the accuracy of nomogram model and comparison with conventional methods. Nomogram shown a significant higher discrimination power than Radscore, MELD score, ALBI score and Child-Pugh score for the prediction of severe PHLF in the training (A) and test (B) datasets. The calibration curves demonstrated good agreement between the radiomics nomogram predicted and actual observation in the training (C) and test (D) datasets.

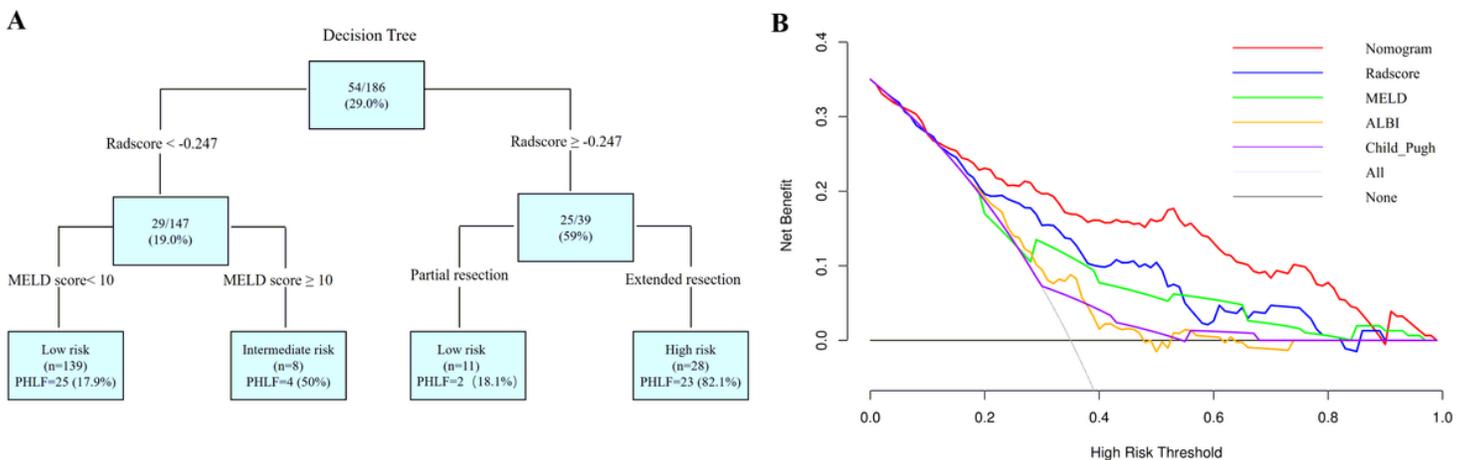


Figure 6

Clinical use. (A). Decision tree stratified the risk for severe PHLF into three classes. (B). DCA showed that the nomogram had wider threshold probabilities and earned more net benefit than conventional models.

Supplementary Files

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

- [Supplementarymaterial.docx](#)