Multiparametric radiomics nomogram may be used for predicting the severity of esophageal varices in cirrhotic patients

Shang Wan[1#], Yi Wei[1#], Xin Zhang[2], Xijiao Liu[1], Weiwei Zhang[1], Yuhao He[3], Fang Yuan[1], Shan Yao[1], Yufeng Yue[1], Bin Song[1]

[1]Department of Radiology, West China Hospital, Sichuan University, No.37, Guoxue Alley, Chengdu 610041, China; [2]Pharmaceutical Diagnostic team, GE Healthcare, Life Sciences, Beijing 100176, China; [3]Department of Neurosurgery, Third People's Hospital of Chengdu, Chengdu 610031, China

Contributions: (I) Conception and design: S Wan, B Song; (II) Administrative support: S Wan, B Song; (III) Provision of study materials or patients: X Zhang; (IV) Collection and assembly of data: S Wan; (V) Data analysis and interpretation: S Wan, X Zhang, S Yao, Y Yue, Y He; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Bin Song. Department of Radiology, The West China Hospital of Sichuan University, No. 37, Guoxue Alley, Chengdu 610041, China. Email: songlab_radiology@163.com.

Background: To explore whether a multiparametric radiomics nomogram on computed tomography (CT) images based on radiomics and relevant parameters of esophageal varices (EV) can be used for predicting the EV severity in patients with cirrhotic livers.

Methods: From January 2016 to August 2018, 136 consecutive patients with clinicopathologically confirmed liver cirrhosis were included for the development of a predictive model. The patients were then divided into two groups, including non-conspicuous EV group (mild-to-moderate EV, n=30) and conspicuous EV group (severe EV, n=106) by using the endoscopic validation as the reference standard. The radiomic scores (Rad scores) were constructed using the binary logistic regression model from the radiomics features of regions of interest (ROIs) in the left liver (LL) and right liver (RL), respectively. The multiparametric nomogram combined the best performance Rad-score and EV-relevant factors, and the calibration, discrimination, and clinical usefulness of developed nomogram were evaluated using calibration curves, decision curve analysis (DCA) and net reclassification index (NRI) analysis respectively.

Results: The LL Rad-score calculated from radiomics features was selected with a relatively higher area under the curve (AUC) (AUC; 0.88, training cohort; 0.87, the validation cohort) compared with RL Rad-score (AUC; 0.86, training cohort; 0.83, the validation cohort). In addition, cross-sectional surface area (CSA) was identified as the important predictor ($P<0.05$), the multiparametric nomogram containing LL Rad-score and CSA was shown to have a better predictive performance and good calibration in the training model (C-index, 0.953, 95% CI, 0.892 to 0.973) and the validation cohort (C-index, 0.938, 95% CI, 0.841 to 0.961), resulting in an improved NRI (categorical NRI of 25.9%, $P=0.0128$; continuous NRI of 120%, $P<0.001$) and integrated discriminatory improvement (IDI) (IDI =13.9%, $P<0.001$). DCA demonstrated that the multiparametric radiomics nomogram was clinically useful.

Conclusions: A multiparametric radiomics nomogram, which incorporates the liver radiomics signature and EV-relevant indices, is a useful tool for noninvasively predicting EV severity and may complement the standard endoscopy for evaluating EV severity in patients with cirrhosis.

Keywords: Esophageal varices (EV); endoscopy; computed tomographic; liver cirrhosis; radiomics; nomogram

Page 2 of 12

Wan et al. Radiomics nomogram of esophageal varices

Introduction

Esophageal varices (EV) is a common complication of patients with portal hypertension, resulted from liver cirrhosis, with a lifetime incidence as high as 80–90% (1), patients with EV have the potential to develop an episode of esophageal variceal bleeding (EVB), which is one of the leading causes of death in cirrhotic patients, with mortality as high as 30% (2). Many clinicians recommend the use of endoscopy for detection of EV severity and initial predicting the risk of variceal bleeding; in addition, some treatment of the varices can be performed by endoscopy, which is considered the gold standard and preliminary diagnostic modality for EV (3-5). However, patients have poor tolerance to endoscopy given the invasiveness nature of this procedure, and it is difficult for patients to conduct repeated examinations in a brief time. Furthermore, performing an endoscopy is unnecessary in some circumstances, such as low-risk patients (6).

Recently, a computed tomography (CT) can be used to effectively evaluate and diagnose EV. On the one hand, due to its noninvasive nature, CT has better compliance and can reduce the pain associated with endoscopy. On the other hand, a CT can reduce the risk of EVB and can be easily performed compared with endoscopy (7,8). Previous studies have reported that CT has been shown to have promising performance for detecting and diagnosing the EV (9,10); however, these studies mainly focused on varices and ignored some of the changes in liver parenchyma caused by cirrhosis, which is the major cause for developing portal hypertension and subsequent EV, thus, we must evaluate the original cause of EV.

Radiomics is the process of the conversion of medical images into high-dimensional, mineable data via high-throughput extraction of quantitative features, followed by subsequent data analysis for decision support (11-15), which contains much more valuable information and has been validated as an effective tool in imaging analysis. A prospective multicenter study used an enhanced CT-based radiomics model to detect of clinically significant portal hypertension (CSPH), they found that radiomics-based hepatic venous pressure gradient (rHVPG) showed a good performance in detection of CSPH with a C-index of 0.849 (95% CI: 0.786–0.911), which can be used as a noninvasive procedure for detecting CSPH (16). EV is known as the major collateral vessel resulted from portal hypertension to our best knowledge; there is no literature describing a radiomics signature that could help the noninvasive detection of EV in cirrhotic patients.

This study aims to analyze the original cause of EV by evaluating the value of liver cirrhosis in diagnosing EV severity. We also intend to use an effective radiomics method to explore the quality of multiparametric radiomics nomograms implemented in liver radiomics and EV-relevant indexes to predict the severity of EV in cirrhotic patients. We also asses weather the use of CT images can be used as a supplementary endoscopy monitoring tool.

Methods

Patients

This retrospective study had approval from the Institutional Review Board of West China Hospital, and the informed consent requirement was waived. Consecutive EV patients between January 2016 to August 2018 were enrolled according to the following inclusion criteria: (I) diagnosis of cirrhosis based on standard histopathological findings, or combination clinical features, laboratory examination and imaging findings (liver configuration, splenomegaly, ascites, and collateral vessels); (II) no prior treatment for including endoscopic variceal ligation (EVL) and β-blockers before admission to our hospital; (III) absence of portal vein emboli and hepatic carcinoma; (IV) no history of splenectomy; (V) available of thoracoabdominal triple-phase enhanced CT scans. CT was used to examine all within 4 weeks of upper gastrointestinal endoscopy. Patients were divided into a non-conspicuous EV group (mild-to-moderate EV) and a conspicuous EV group (severe EV) by endoscopic examination (17,18). All patients were divided into a training set and a validation set at a ratio of 7:3 (19,20) with a random sampling method, 95 patients constituted the training cohort, and the other 41 formed the validation cohort. The endoscopic study of our research was presented in supplementary materials (S-text-1).

CT image acquisition

Patients underwent a contrast-enhanced CT scan using one of the following systems: LightSpeed VCT (GE Healthcare), Sensation 64 CT (Siemens), or Sensation 16 CT (Siemens). Three phases, including non-enhanced, arterial, and portal vein phase, were obtained. Arterial phase scanning started about 20–30 s after the beginning of injection, and portal phase scanning was started 30–40 s after the beginning of the arterial phase. The reconstitution thickness was set at 1–2 mm, and reconstructions were done on a GE Advantage Windows 3D workstation
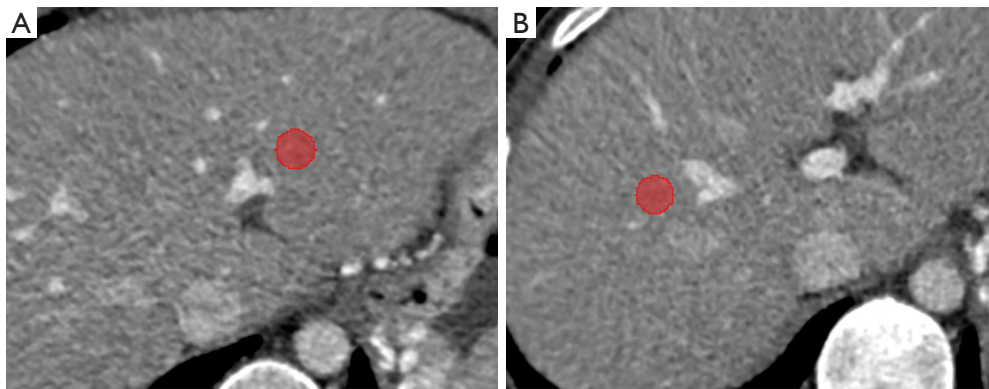
Figure 1 ROIs of the left liver (LL) and right liver (RL). (A) In LL, ROI was set 1–2 cm beside the stem of the left-portal vein (LPV); (B) in RL, ROI was set 1–2 cm beside the stem of the right-portal vein (RPV). The dyeing circles are ROIs. ROI, region of interest.

(GE Healthcare, Waukesha, WI, USA). The following parameters were used: tube voltage, 120 or 100 kVp; tube current, 150–600 mA; slice thickness, 1.25 mm; pitch, 1.375. All patients received an intravenous, nonionic contrast medium (iodine concentration, 370 mg/mL; volume, 1.5–2.0 mL/kg of body weight; contrast type, Iopromide Injection, Bayer Pharma AG) at a rate of 3–5 mL/s. A volume of 20 mL saline was injected after the injection of the contrast.

CT image analysis

A software package 3D Slicer version 4.10.1 (Boston, USA) was used as the measuring software, which is equipped with accurate 2D and 3D measurement tools, EV grade (EVG) and measurements of the EV diameter (EVD), cross-sectional surface area (CSA), EV volume (EVV), spleen volume (SV) and diameter of left gastric vein (DLGV) were made independently by two abdominal radiologists (X.J.L and W.W.Z) with over 10 years' experiences in analyzing CT images. Both observers were blinded to patient physical findings, laboratory data, earlier imaging findings, and endoscopy results. The portal venous phase was chosen as the observation period.

The EV grading system on CT images was according to criteria proposed by Kim (2), which was classified as I–IV mainly according to the EVD, measurement of EV-relevant indices (EVG, EVD, CSA, EVV, SV, DLGV) are reported in supplementary materials (S-text-2) (10,21).

Radiomics feature extraction

Portal venous phase CT images were used for radiomics

feature extraction (16,22), Regions of interest (ROIs) were drawn manually on 2 different areas of the left liver (LL) and right liver (RL) respectively according to branches of the portal vein (PV) with 3D-slicer software by the above two readers (*Figure 1*). In LL, ROI was set 1–2 cm beside the stem of the left-portal vein (LPV), which divided the LPV into the left lateral and left an inner branch of PV. In RL, ROI was set 1–2 cm beside the stem of the right-portal vein (RPV), which divided the RPV into the right posterior and right anterior branch of PV. The ROI was 1 to 2 cm^2 in size and excluded large vascular structures, artifacts, or liver lesions such as cirrhotic nodules on CT imaging (23), then, Artificial Intelligence Kit software (A.K. software; GE Healthcare, Life Sciences, Beijing, China) was used to extract feature parameters for each ROI, which was based on the Image biomarker standardization initiative (IBSI).

Data analysis

Statistical analysis

Statistical analyses were performed with R software version 3.4.3 (https://www.r-project.org). The LASSO and binary logistic regression were performed using the "glmnet" package (R language 3·0·2, R Core Team, 2013). The receiver operating characteristic (ROC) curves are created, and the area under the ROC curve (AUC) was calculated by using the "pROC" package. The nomogram and calibration plots were constructed using the "rms" package. C-index was calculated using the "Hmisc" package. The "generalhoslem" package was used to execute the Hosmer-Lemeshow (H-L) test. The "rmda" package was used in DCA. Chi-square test, *t*-test, and intraclass correlation coefficient (ICC) were calculated using MATLAB 2016b

Page 4 of 12

Wan et al. Radiomics nomogram of esophageal varices

(Mathworks, Natick, USA) with the "ICC" function supplemented. The Mann-Whitney U test compared continuous variables, and the Chi-square test compared categorical variables. ICC was considered to evaluate the inter-observer agreement, both in liver radiomics and EV-relevant indices; the specified criteria were reported in supplementary materials (S-text-3). Radiomics features with ICC values more than 0.87 (showing excellent stability) were selected for subsequent investigation (supplementary material, *Figure S1*). A P<0.05 (two-sided) was accepted as statistically significant.

EV-relevant indices choice

EV-relevant indices (EVG, EVD, CSA, EVV, SV, DLGV) with P<0.05 in univariable analysis were integrated into an adjusted odd ratio (OR) regression model. Multivariable analysis with a backward stepwise selection process by using the likelihood ratio test with Akaike's information criterion (24) was performed to select important clinical predictors.

Radiomics feature choice, multiparametric radiomics nomogram development, and evaluation

Radiomics features which had greater ICCs considering a threshold of 0.87 were robust and adopted for later analysis in LL and RL respectively, each radiomics feature was standardized. The least absolute shrinkage and selection operator (LASSO) method, which is suitable for the regression of high-dimensional data, was used to select the most useful predictive features. The LL and RL radiomic models were developed using a logistic regression model in the training cohort separately, and the Rad-score was computed for each patient by a linear combination of selected features weighted by their respective coefficients (25). A combined predictive model was built by combining EV-relevant indices and the better performance between LL and RL Rad-scores, and the multiparametric nomogram was established as a quantitative prediction tool.

The discrimination performance of the clinical, radiomics, and combined models was evaluated with the ROC curve analysis and AUC value in the training and validation cohort. Since the dataset used for model development was also used to assess model discrimination, the actual AUC estimates could be over-optimistic (26). In addition, considering the limited sample-set size of this study, an internal validation with 100 iterations of five-fold cross-validation was used for AUC evaluation (27,28). The Delong's test was used to compare the statistical significance between any two ROCs. The Hosmer-Lemeshow goodness of fit test was applied for the prognosis model (29). Furtherly, Harrell's C-index was calculated to quantify the discrimination performance of the multiparametric nomogram in both cohorts, the bootstrapping validation with 1,000 resampling being conducted to calculate a relatively corrected C-index. The integrated discrimination improvement (IDI) and the net reclassification index (NRI) calculation was also applied to quantify the improvement of usefulness added by the multiparametric radiomics nomogram (30). The decision curve analysis (DCA) was conducted to evaluate the clinical usefulness of the multiparametric radiomics nomogram by quantifying the net benefits at different threshold probabilities in the combined training and validation cohorts (31).

Results

Patients' demographics and endoscopic result

One hundred fifty-seven consecutive EV patients were enrolled in the initial population, according to the inclusion and exclusion criterion, 21 patients were excluded from the study, and 8 patients had a history of hepatic carcinoma, 8 patients had a history of EVL, and 5 patients had a history of splenectomy. Finally, 136 EV patients with liver cirrhosis were evaluated in this study, patients' characteristics are given in *Table 1*, there were no significant differences in the clinical characteristics between the training and the validation cohort. We analyzed the findings obtained by endoscopy, which is regarded as the reference standard in our study, the grading system and the standard endoscopic results of our study are given in supplementary materials (*Table S1*). The flowchart of our study is shown in *Figure 2*.

CT findings

Six EV-relevant features (EVG, EVD, CSA, EVV, SV, DLGV) are considered in the univariable analysis, and the results showed all indices (EVG, EVD, CSA, EVV, DLGV) except for SV were independent prognostic factors of EV, which were statistically significant (*Table 2*). ROC analysis of each EV-relevant index in both cohorts is depicted in the supplementary materials (*Figure S2*). Excellent inter-observer agreements were achieved in all 6 EV-relevant features (r_{ic} >0.9). A patient with severe EV (*Figure 3*) is depicted.

Table 1 Characteristics of 136 esophageal varices (EV) patients with cirrhosis

Demographics or characteristic	Training cohort (n=95)		Validation cohort (n=41)		Estimate risk*	P value**
	Non-conspicuous EV (n=21)	Conspicuous EV (n=74)	Non-conspicuous EV (n=9)	Conspicuous EV (n=32)		
Age, mean ± SD, years	56.3±14.7	60.4±12.4	54.7±14	52.5±12.5		
Gender, n (%)						0.144
Male	9 (42.9)	45 (60.8)	6 (66.7)	26 (81.2)	1	
Female	12 (57.1)	29 (39.2)	3 (33.3)	6 (18.8)	0.48 (0.18–1.29)	
Etiology, n (%)						0.998
Post-hepatic cirrhosis	12 (57.1)	48 (64.9)	4 (44.4)	18 (56.2)	–	
Alcoholic cirrhosis	6 (28.6)	14 (18.9)	4 (44.4)	10 (31.2)	–	
Combined cirrhosis	1 (4.8)	2 (2.7)	0 (0)	2 (6.3)	–	
Primary biliary cirrhosis	2 (9.5)	8 (10.8)	1 (11.1)	2 (6.3)	–	
Autoimmune hepatic cirrhosis	0 (0)	1 (1.4)	0 (0)	0 (0)		
Cryptogenic cirrhosis	0 (0)	1 (1.4)	0 (0)	0 (0)	–	
Child-Pugh class, n (%)						
Class A	8 (38.1)	18 (24.3)	3 (33.3)	9 (28.1)	1	0.195
Class B	8 (38.1)	38 (51.4)	2 (22.2)	15 (46.9)	2.11 (0.68–6.53)	
Class C	5 (23.8)	18 (24.3)	4 (44.4)	8 (25.0)	1.6 (0.44–5.84)	

*, the estimated risk of patients' characteristics with univariable analysis; **, t-test or chi-square test P values between the non-conspicuous EV group and the conspicuous EV group of both cohorts. EV, esophageal varices; SD, standard deviation.

Radiomics findings

Liver radiomics feature choice and the construction of the Rad-score system

In LL of the portal venous-CT images, 3 important radiomic features were selected from 1,033 features after two-step feature reduction. *Figure 4A,B* represent the features selected with non-zero coefficients in the LASSO logistic regression model in the training cohort. The AUCs of the LL radiomics model were 0.88 and 0.865 for the training cohort and validation cohort, respectively. The same method is applied to RL, 8 potential predictors with non-zero coefficients in the LASSO logistic regression model were selected from 1,033 radiomics (*Figure 4C,D*) features of RL parenchyma, the performance of the RL radiomics model for differentiation of mild-to-moderate and severe EV was 0.857, 0.828 of AUC for training cohort and validation cohort respectively, the AUC, accuracy (ACC), sensitivity, specificity of LL and RL Rad-score were given in *Table 3*. The LL and RL Rad-scores were

calculated according to the coefficients of each radiomics model, and the formulas were described in supplementary materials (S-text-4). Besides, our results represented that there was no significant difference in performance indices by ROC analysis of LL and RL Rad-score (P=0.7). Due to the better performance, the LL Rad score was selected as the radiomics signature.

Development, apparent performance, and validation of the multiparametric radiomics nomogram

In the multivariable regression model, the LL Rad-score and CSA were identified as the important predictors (all P<0.001). Then, a multiparametric nomogram having these 2 factors was established (*Figure 5A*).

The nomogram showed a satisfied diagnostic capability for the detection of severe EV in cirrhosis in both training and validation cohort, the C-indexes of the multiparametric radiomics nomogram were 0.953 (95% CI, 0.892 to 0.973) in the training cohort, and
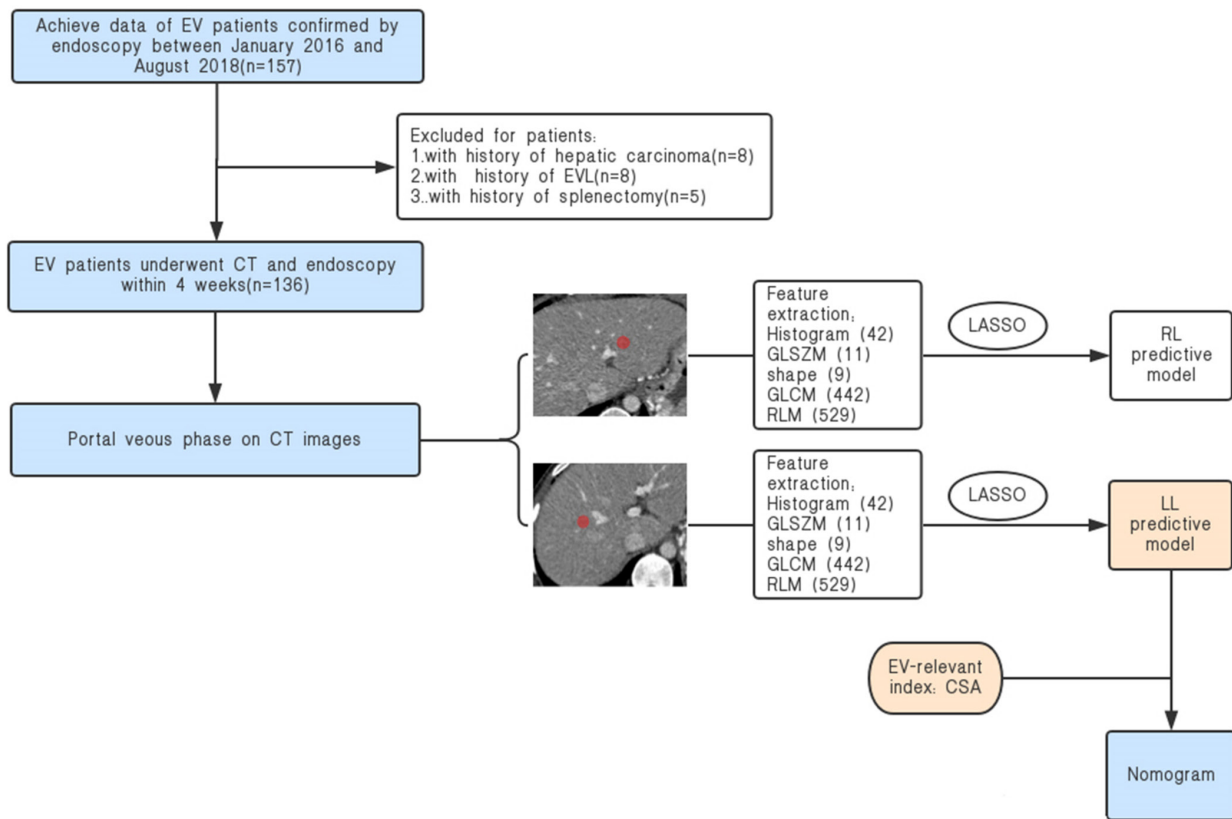
Page 6 of 12

Wan et al. Radiomics nomogram of esophageal varices

Figure 2 Flowchart of the study. EV, esophageal varices; LL, left liver; RL, right liver; CSA, cross-sectional surface area.

Table 2 Measurement of the indices in the individual group as referred to endoscopy, P values of CT for the detection of esophageal varices (EV) of each index

Variable	Training cohort (n=95)		Validation cohort (n=41)		Estimate risk*	P value**
	Non-conspicuous EV (n=21)	Conspicuous EV (n=74)	Non-conspicuous EV (n=9)	Conspicuous EV (n=32)		
EVG, n						
I	5	2	5	3	1	
II	10	5	17	10	1.70 (0.39–7.36)	0.478
III	6	2	52	19	8.67 (1.93–38.85)	0.005
EVD, cm	6.13 (4.82–7.81)	5.34 (4.16–7.31)	4.41 (3.44–5.12)	4.56 (3.98–4.83)	1.7 (1.23–2.36)	0.001
CSA, cm^2	69.07 (28.75–138.44)	61.21 (29.92–85.92)	24.77 (11.44–34.25)	23.27 (20.27–31.71)	1.03 (1.01–1.05)	0.003
EVV, cm^3	1.95×10^3 (1.03×10^3–5.14×10^3)	2.01×10^3 (8.36×10^2–3.26×10^3)	5.42×10^2 (4.12×10^2–1.46×10^3)	1.10×10^3 (6.8×10^2–1.37×10^3)	1.0008 (1.0002–1.0013)	0.006
SV, cm^3	7.33×10^5 (5.30×10^5–1.06×10^6)	8.62×10^5 (5.17×10^5–1.38×10^5)	6.79×10^5 (3.91×10^5–8.27×10^5)	4.47×10^5 (3.48×10^5–7.64×10^5)	1 (1,1)	0.107
DLGV, cm	6.00 (5–7.75)	6.00 (4.00–7.25)	4.00 (3.00–5.00)	4.00 (4.00–5.00)	1.41 (1.04,1.89)	0.026

*, the estimated risk of EV-relevant indices with univariable analysis; **, *t*-test or chi-square test P values between the non-conspicuous EV group and the conspicuous EV group of both cohorts. EV, esophageal varices; EVG, EV grade; EVD, EV diameter; EVV, EV volume; SV, spleen volume; DLGV, diameter of left gastric vein.
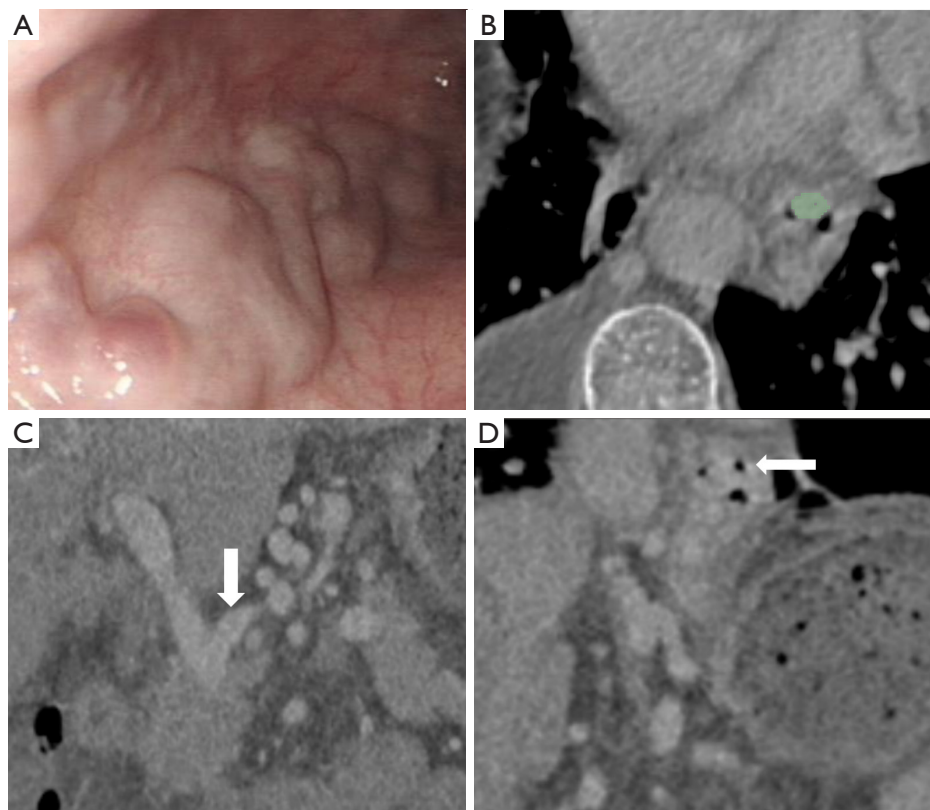
Figure 3 Severe EV according to endoscopy in a 68-year-old man (severe of endoscopy, maximum minor axis 7.48 mm). (A) The endoscopic image shows severe nodular varices; (B) the sum of the cross-sectional surface area (CSA) of EV (dyeing area) in the transverse section is depicted in 3D slicer; (C) the axial CT image shows the left gastric vein (LGV) originating from the portal vein (arrow); (D) 3D reconstruction has a satisfactory performance in visualizing severe EV of the lower esophagus (arrow). EV, esophageal varices.

the nomogram for the prediction of severe EV yielded a C-index of 0.938 (95% CI, 0.841 to 0.961) in the validation cohort. The ROC analysis and AUC of both cohorts were shown in *Figure 5B*, and the multiparametric nomogram showed better predictive performance and accuracy than a single radiomic predictive model (P=0.02, Delong test) in both cohorts. The AUC, sensitivity, specificity, cut-off values of both cohorts for the diagnosis of EV severity in cirrhosis were summarized in *Table 3*. However, in LL predictive model, the mean AUC over 100 iterations of five-fold cross-validation was lower (mean AUC =0.70, *Table 3*) than the model built in this study (training, AUC =0.88, validation, AUC =0.87), which can also be obtained from RL predictive model (*Table 3*). Thus, the models built in this study based on dividing cohorts were selected as the final models with satisfactory predictive performance.

The calibration curves of the nomograms for training and validation cohorts (*Figure 5C*) revealed good agreement between prediction and actual observation in both cohorts, H-L test revealed that the results for the calibration curves (training cohort, P=1; validation cohort, P=0.4) were reliable in both cohorts, which indicated good calibrations.

Clinical use

The DCA for the multiparametric radiomics nomogram and the Rad-score were depicted in *Figure 5D*, the DCA showed that the multiparametric radiomics nomogram had a higher overall net benefit than Rad-score when the threshold probability for a doctor or patient was within a range from 0.36 to 0.99, and this nomogram added more net benefit than the "treat all" or "treat none" scheme.

The predictive capability for EV severity of the multiparametric radiomics nomogram was significantly improved compared with the liver radiomics signature
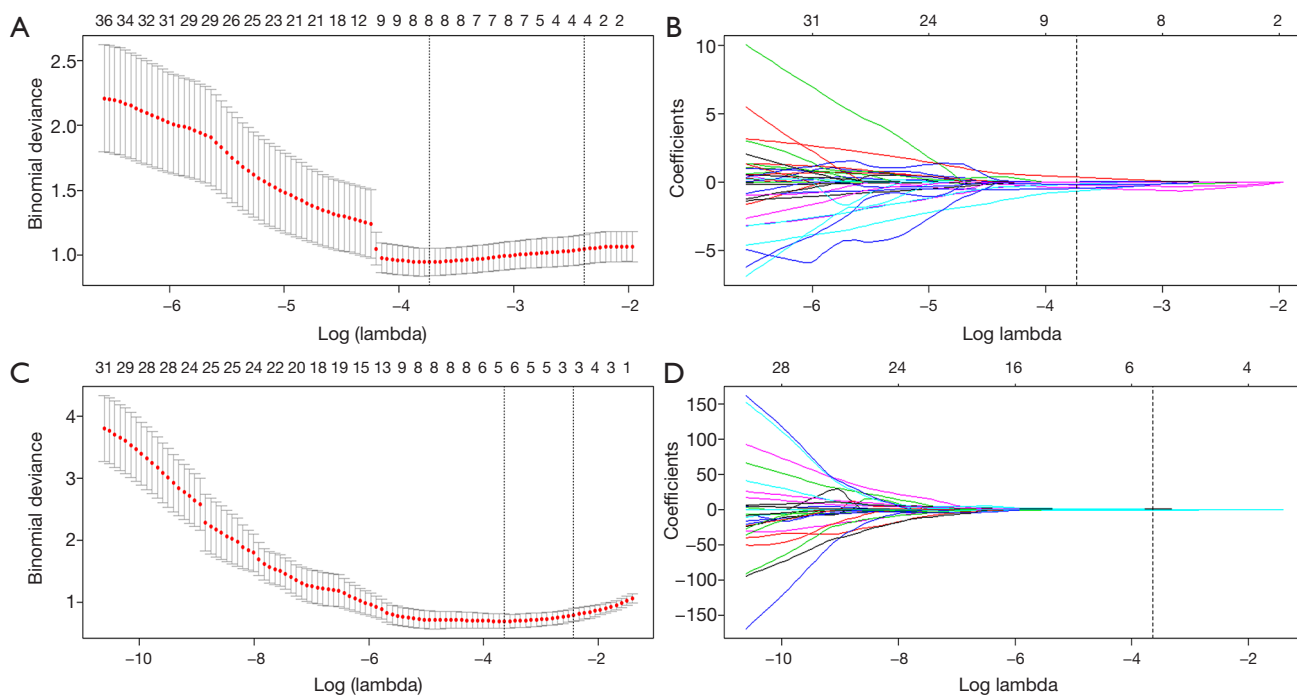
Page 8 of 12

Wan et al. Radiomics nomogram of esophageal varices

Figure 4 Radiomics feature selection of LL and RL using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. Ten-fold cross-validation via minimum criteria was used to select the tuning parameter (lambda) in the LASSO model of LL (A) and RL (C). The binomial deviance was plotted vs. log (lambda), and the numbers along the upper x-axis indicate the number of predictors. Dotted vertical lines were drawn at the optimal values by using the minimum criteria and one standard error of the minimum criteria. LASSO coefficient profiles of the 1,033 radiomics features of LL (B) and RL (D). The y-axis is the value of the coefficient, and the lower x-axis indicates a log (lambda). The upper x-axis is the number of non-zero coefficients. The vertical line was drawn at the value selected using 10-fold cross-validation, where optimal lambda resulted in seven coefficients. LL, left liver; RL, right liver.

Table 3 Performance of multiparametric radiomics nomogram and liver radiomics in diagnosing EV

Variable	Multiparametric radiomics nomogram		LL Rad-sore		RL Rad-sore	
	Train (n=95)	Validation (n=41)	Train (n=95)	Validation (n=41)	Train (n=95)	Validation (n=41)
AUC	0.95 (0.911–0.99)	0.92 (0.78–1)	0.88 (0.80–0.96)	0.87 (0.67–1)	0.86 (0.77–0.94)	0.83 (0.65–1)
Mean cross-validated AUC	0.87		0.70		0.76	
ACC	0.91	0.93	0.85	0.81	0.83	0.81
Sensitivity	0.86	0.89	0.71	0.78	0.70	0.78
Specificity	0.92	0.94	0.89	0.81	0.87	0.81
Cut off value	0.67		0.74		0.67	

Severity in cirrhosis in the training and validation cohorts. AUC, the area under the curve; ACC, accuracy; EV, esophageal varices.

(Rad-score), which was demonstrated with NRI and IDI, with a categorical NRI of 25.9% (P=0.0128), continuous NRI of 120% (P<0.001), and IDI of 13.9% (P=0.00013) for predicting severe EV in the validation cohort, the multiparametric radiomics nomogram incorporated the liver radiomics signature and EV-relevant indices showed marked discrimination accuracy compared to liver radiomics signature.
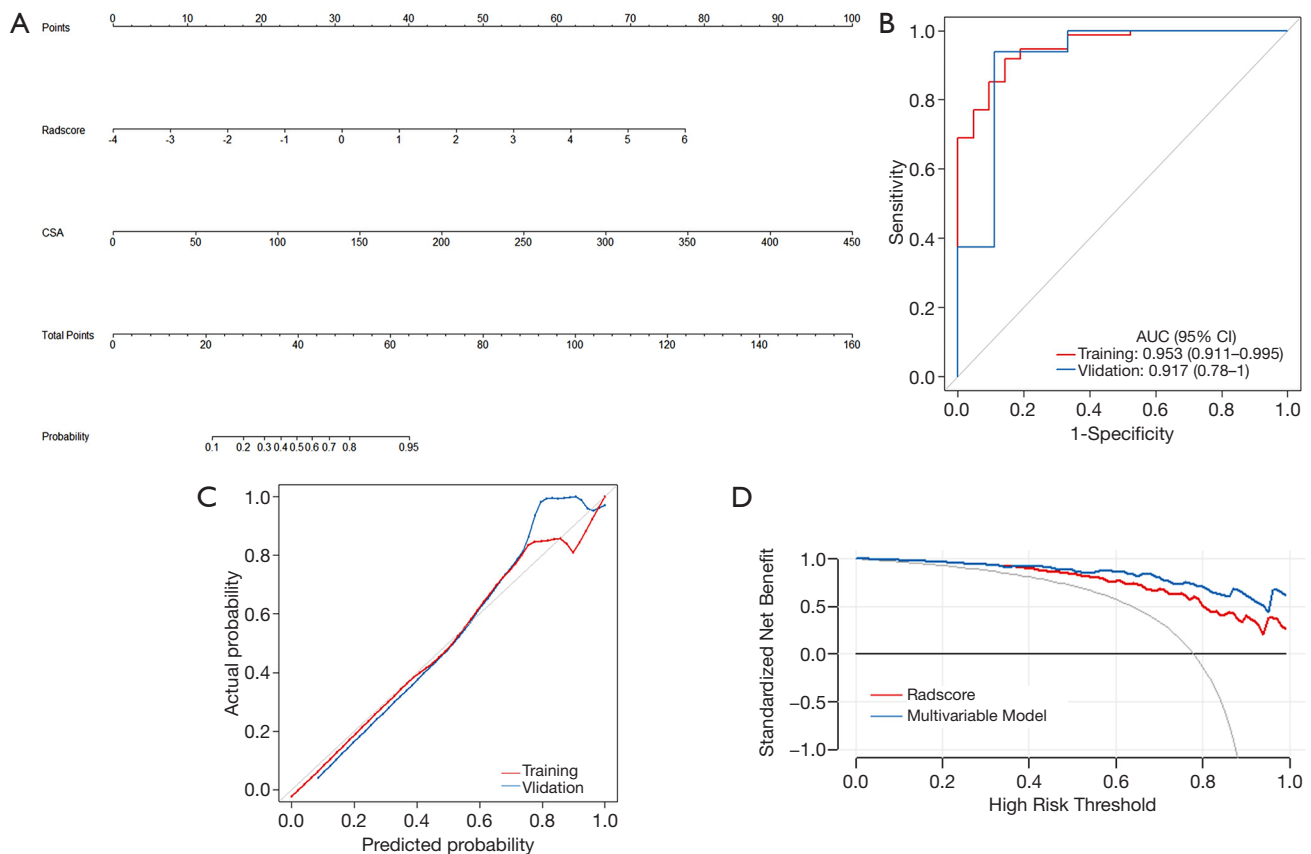
Figure 5 (A) Developed multivariable radiomics nomogram. The radiomics nomogram was developed in the training cohort and incorporated the LL Rad-score and CSA. (B) ROC curves analysis and a calibration curve of the nomogram, which shows the ROC curves for the detection of EV severity in cirrhosis in the training and validation cohorts, respectively, radiomics based on endoscopy. (C) The calibration curve of the multivariable radiomics nomogram in both cohorts, which indicates the calibration of the model in terms of an agreement between predicted risks and observed outcomes of EV severity. The solid red line and blue line represent the predictive performance of the training cohort and validation cohort of the nomogram, and the diagonal grey line represents an ideal model. Closer fit to the diagonal grey line indicates a better prediction. (D) Decision curve analysis for multivariable model and the LL Rad score. The x-axis represents the threshold probability, and the y-axis measures the net benefit; the blue line represents the multivariable model; the red line represents the LL Rad score. The threshold probability is where the expected benefit of treatment is equal to the expected benefit of avoiding treatment. The multivariable radiomics model had a higher net benefit compared to LL Rad-score when the threshold probability is within a range from 0.36 to 0.99; the more net benefit is added by using the nomogram for predicting severe EV than by treating either no or all patients. EV, esophageal varices; LL, left liver; CSA, cross-sectional surface area.

Discussion

In this study, we developed and validated a multiparametric radiomics nomogram for the noninvasive prediction of EV severity in patients with liver cirrhosis, which integrates radiomics and EV-relevant indices (EVG, EVD, CSA, EVV, SV, DLGV) based on CT images.

Our results demonstrated that the nomogram could successfully place stratified EV patients into mild-to-moderate EV and severe EV groups, and this multiparametric radiomics nomogram had better diagnostic performance compared with traditional radiomics, indicating it to have a potential role of this nomogram for evaluating EV severity, which may complement endoscopic surveillance in cirrhotic patients.

For the construction of the Rad score, the liver radiomic models were respectively set up by the texture features extracting from LL and RL ROIs. LL Rad-score was

Page 10 of 12

Wan et al. Radiomics nomogram of esophageal varices

considered into multivariable model due to relatively higher performance (AUC, 0.88 in training cohort; 0.865 in validation cohort) compared with RL Rad-score (AUC, 0.857 in training cohort; 0.828 in validation cohort), on the basis of our results and the principle for selecting the predictive model with better performance (20); whereas our results indicated that there was no significant difference in the diagnostic performance of LL and RL Rad-score. These results may correlate with the relative homogenization of liver changes due to cirrhosis.

It also should be noted of our study that all EV-relevant indices (EVG, EVD, CSA, EVV, DLGV) except SV can be helpful predictors for distinguishing conspicuous and non-conspicuous groups. Kim *et al.* concluded that a criterion EVD of 3 mm or larger on CT could be useful for identifying high-risk patients requiring prophylactic therapy (9), Miller *et al.* found a significant difference (P<0.018) in CSA between patients with EVB compared to those without bleeding using a cutoff of 0.45 cm^2 as measured by ultrasonic endoscopy, with 83% sensitivity and 75% specificity (10). Compared with the above previous studies (9,10), in this study, several new indices (EVG, EVV, SV, DLGV) were included, which have not been evaluated and verified, the results suggest that EVG, EVV, DLGV might be effective for identifying severe EV with accuracy comparable to EVD and CSA, which have been validated by the above reports. Further studies are needed to verify our results.

After building the multiparametric radiomics nomogram, incorporating the LL radiomics signature and EV-relevant indices, we found that this nomogram demonstrated promising discrimination with an AUC of 0.953 in the training cohort and 0.917 in the validation cohort for detection of severe EV in cirrhosis, the DCA showed that the multiparametric radiomics nomogram had a higher overall net benefit than LL Rad-score and the diagnostic performance for EV severity of the multiparametric radiomics nomogram was significantly improved compared with the LL radiomics signature (Rad-score), which was also demonstrated with a categorical NRI of 25.9% (P=0.0128), continuous NRI of 120% (P<0.001), and IDI of 13.9% (P=0.00013). A recent study revealed that a combination of multiple prognostic factors could show a better predictive performance than just one factor (32), which is consistent with our results.

Our results present several advantages when compared to previous reports (9,10). In the first place, most of the researchers mainly assess the varices themselves, to our best knowledge, esophageal and gastric fundic varices are the most common collateral vessels in portal hypertension, due to an increase in intrahepatic resistance combined with an increase in portal and hepatic arterial blood flow resulted from liver cirrhosis (33,34), thus, in our study, instead of only concentrating on the varices, we emphasized on liver cirrhosis which gives rise to EV and portal hypertension and simultaneously evaluated the varices. What's more, radiomics was originally focused on tumor tissues (13,35), recently, Liu *et al.* demonstrated that a noninvasive signature based on liver had an excellent diagnostic performance for CSPH in cirrhotic patients (16), suggesting a potential role of radiomics in diagnosing portal hypertension associated with liver cirrhosis. EV is known as a possible outcome of portal hypertension resulted from liver cirrhosis; on this basis, we evaluated the varices associated with portal hypertension with the approach of radiomics, which might can be used as a new method to predict EV severity with more valuable information. Finally, important benefits of the multiparametric radiomics nomogram is that the relative comprehensiveness of our predictive model, there is no previous literature integrated radiomics and EV-relevant indices for evaluating EV severity noninvasively in cirrhosis, and our results indicated a satisfactory predictive performance of the developed nomogram. Therefore, we suggested that the multiparametric radiomics nomogram might be served as a supplement surveillance method for diagnosing EV severity.

There are several limitations so this study. First, endoscopy is considered the reference standard in this study and there is a significant degree of inter-and interobserver disagreement due to the subjective nature of endoscopic grading systems (36). Since there is no other reference standard for the evaluation of EV, we recommend that further re-evaluation studies should be performed by trained endoscopists, which may eliminate any potential bias among observers.

In addition, although the consecutive patients of our study were included according to the inclusion and exclusion criteria strictly, the conspicuous EV patients composed the majority of our population due to the high prevalence of conspicuous EV of our hospital, which may bring some potential bias of our population composition, and polycentric large-scale samples with more non-conspicuous EV patients are warranted to obtain high-quality evidence for further validation. Finally, CT can evaluate EV severity noninvasively according to our results, while some therapeutic intervention cannot be performed

during CT, it is possible during endoscopy (37).

Conclusions

In summary, this study demonstrates that a multiparametric radiomics nomogram that incorporates both the liver radiomics signature and EV-relevant indices, it is validated as a useful predictive tool for noninvasive detection of EV severity and may be conveniently used to supplement the standard endoscopy for evaluating EV severity in patients with cirrhosis.

Acknowledgments

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study had approval from the Institutional Review Board of West China Hospital, and the informed consent requirement was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Kim SH, Kim YJ, Lee JM, et al. Esophageal varices in patients with cirrhosis: multidetector CT esophagography--comparison with endoscopy. Radiology 2007;242:759-68.
2. Solanki S, Haq KF, Chakinala RC, et al. Inpatient burden of esophageal varices in the United States: analysis of trends in demographics, cost of care, and outcomes. Ann Transl Med 2019;7:480.
3. de Franchis R. Updating consensus in portal hypertension: report of the Baveno III Consensus Workshop on definitions, methodology and therapeutic strategies in portal hypertension. J Hepatol 2000;33:846-52.
4. Schepke M, Kleber G, Nurnberg D, et al. Ligation versus propranolol for the primary prophylaxis of variceal bleeding in cirrhosis. Hepatology 2004;40:65-72.
5. D'Amico G, Pagliaro L, Bosch J. Pharmacological treatment of portal hypertension: an evidence-based approach. Semin Liver Dis 1999;19:475-505.
6. Spiegel BM, Targownik L, Dulai GS, et al. Endoscopic screening for esophageal varices in cirrhosis: Is it ever cost effective? Hepatology 2003;37:366-77.
7. Perri RE, Chiorean MV, Fidler JL, et al. A prospective evaluation of computerized tomographic (CT) scanning as a screening modality for esophageal varices. Hepatology 2008;47:1587-94.
8. Dessouky BAM, Aal ESMA. Multidetector CT oesophagography: An alternative screening method for endoscopic diagnosis of oesophageal varices and bleeding risk. Arab J Gastroenterol 2013;14:99-108.
9. Kim H, Choi D, Gwak GY, et al. Evaluation of esophageal varices on liver computed tomography: receiver operating characteristic analyses of the performance of radiologists and endoscopists. J Gastroenterol Hepatol 2009;24:1534-40.
10. Miller L, Banson F, Bazir K, et al. Risk of esophageal variceal bleeding based on endosscopic ultrasound evaluation of the sum of esophageal variceal cross sectional surface area (CSA). Am J Gastroenterol 2003;98:454-9.
11. Cai W, He B, Hu M, 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.
12. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.
14. Kolossváry M, Karady J, Szilveszter B, et al. Radiomic Features Are Superior to Conventional Quantitative Computed Tomographic Metrics to Identify Coronary Plaques With Napkin-Ring Sign. Circ Cardiovasc Imaging 2017. doi: 10.1161/CIRCIMAGING.117.006843.
15. Jiang Y, Chen C, Xie J, et al. Radiomics signature of computed tomography imaging for prediction of

Page 12 of 12

Wan et al. Radiomics nomogram of esophageal varices

survival and chemotherapeutic benefits in gastric cancer. EBioMedicine 2018;36:171-82.

16. Liu F, Ning Z, Liu Y, et al. Development and validation of a radiomics signature for clinically significant portal hypertension in cirrhosis (CHESS1701): a prospective multicenter study. EBioMedicine 2018;36:151-8.

17. Ding D. Endoscopic diagnosis and treatment of esophagus and gastric varices. Trial program of Chinese Society of Digestive Endoscopy. Chinese J Dig Endosc 2000;17.

18. Kawano Y, Sasaki A, Kai S, et al. Short- and Long-Term Outcomes after Hepatic Resection for Hepatocellular Carcinoma with Concomitant Esophageal Varices in Patients with Cirrhosis. Ann Surg Oncol 2008;15:1670-6.

19. Yin P, Mao N, Liu X, et al. Can clinical radiomics nomogram based on 3D multiparametric MRI features and clinical characteristics estimate early recurrence of pelvic chondrosarcoma? J Magn Reson Imaging 2020;51:435-45.

20. Tan Y, Zhang ST, Wei JW, et al. A radiomics nomogram may improve the prediction of IDH genotype for astrocytoma before surgery. Eur Radiol 2019;29:3325-37.

21. Zhou HY, Chen TW, Zhang XM, et al. Diameters of left gastric vein and its originating vein on magnetic resonance imaging in liver cirrhosis patients with hepatitis B: Association with endoscopic grades of esophageal varices. Hepatol Res 2014;44:E110-7.

22. Kim H, Choi D, Gwak GY, et al. High-risk esophageal varices in patients treated with locoregional therapies for hepatocellular carcinoma: evaluation with regular follow-up liver CT. Dig Dis Sci 2009;54:2247-52.

23. Chen BB, Hsu CY, Yu CW, et al. Clinical and histologic implications of delayed hepatobiliary enhancement on magnetic resonance imaging with gadolinium ethoxybenzyl diethylenetriaminepentaacetic Acid. Invest Radiol 2012;47:649-55.

24. Wu S, Zheng J, Li Y, et al. A Radiomics Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Bladder Cancer. Clin Cancer Res 2017;23:6904-11.

25. Farinati F, Vitale A, Spolverato G, et al. Development and Validation of a New Prognostic System for Patients with Hepatocellular Carcinoma. PLoS Med 2016;13:e1002006.

26. Giancristofaro A, Salmaso L. Model performance analysis and model validation in logistic regression. Statistica 2007;63:375-96.

27. Antunovic L, Sanctis RD, Cozzi L, et al. PET/CT radiomics in breast cancer: promising tool for prediction of pathological response to neoadjuvant chemotherapy. Eur J Nucl Med Mol Imaging 2019;46:1468-77.

28. Xu X, Zhang X, Tian Q, et al. Quantitative Identification of Nonmuscle-Invasive and Muscle-Invasive Bladder Carcinomas: A Multiparametric MRI Radiomics Analysis. J Magn Reson Imaging 2019;49:1489-98.

29. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med 2007;35:2052-6.

30. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011;30:11-21.

31. Vickers AJ, Cronin AM, Elkin EB, et al. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak 2008;8:53.

32. Riley RD, Hayden JA, Steyerberg EW, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med 2013;10:e1001380.

33. Zhou HY, Chen TW, Zhang XM, et al. The diameter of the originating vein determines esophageal and gastric fundic varices in portal hypertension secondary to posthepatitic cirrhosis. Clinics (Sao Paulo) 2012;67:609-14.

34. Rahimi RS, Rockey DC. Complications and outcomes in chronic liver disease. Curr Opin Gastroenterol 2011;27:204.

35. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol 2017;28:1191-206.

36. Calès P, Oberti F, Bernard-Chabert B, et al. Evaluation of Baveno recommendations for grading esophageal varices. J Hepatol 2003;39:657-9.

37. Zhu H, Shi B, Upadhyaya M, et al. Therapeutic endoscopy of localized gastric varices: pretherapy screening and posttreatment evaluation with MDCT portography. Abdominal Imaging 2010;35:15-22.

S-text-1: endoscopic study

A Fujinon EG 485 (Fujinon, Saitama, Japan) or OlymPus CV240 electronic endoscope (Olympus Optical Co. Ltd., Tokyo, Japan) was used for EV diagnosing and grading. The grade of EV were represented as mild, moderate, and severe according to the form (F), and the absence or presence of red color sign (RC), which were recorded following The General Rules for Study of Portal Hypertension (The Japan Society for Portal Hypertension, 2nd Edition, 2004) (18) and (Endoscopic diagnosis and treatment standard trial plan for esophagogastric varices, Chinese Society of Digestive Endoscopy, 2000) (17).

S-text-2

In the measurement of EVD and CSA, an image was selected at a point where the varices appeared largest, and the observers measured the short-axis diameter of the largest visible esophageal varix (38). The individual CSAs of all varices were added to yield the final CSA (10). The EVV included all varices abutting the lumen of the observation range. The DLGV was measured at 1 cm from the starting point of the portal or splenic vein opening (21).

S-text-3

The inter-observer agreements for grading of EV were classified as excellent, good, fair, or poor according to the ric(intra-class correlation coefficient, ICC), we considered a ric value of more than 0.87 to represent almost perfect agreement and values of 0.71-0.87 and 0.50-0.71 to represent good and fair agreement, respectively, values of less than 0.50 were considered to represent poor agreement (36).

S-text-4: formula

LL: $1.726 - 1.071*Correlation\_angle45\_offset7 - 1.682*Correlation\_angle45\_offset9 + 1.241*InverseDifferenceMoment\_angle135\_offset3$

LR: $1.88 - 0.0928*ClusterShade\_angle45\_offset6 + 0.000234*LongRunHighGreyLevelEmphasis\_angle0\_offset6 - 0.000271*LongRunHighGreyLevelEmphasis\_angle45\_offset6$

Nomogram: $-2.4864 + 1.2955*Radscore + 0.0583*CSA$

References

38. Wan S, Wei Y, Yu H, et al. Computed Tomographic Portography with Esophageal Variceal Measurements in the Evaluation of Esophageal Variceal Severity and Assessment of Esophageal Variceal Volume Efficacy. Acad Radiol 2019. [Epub ahead of print].
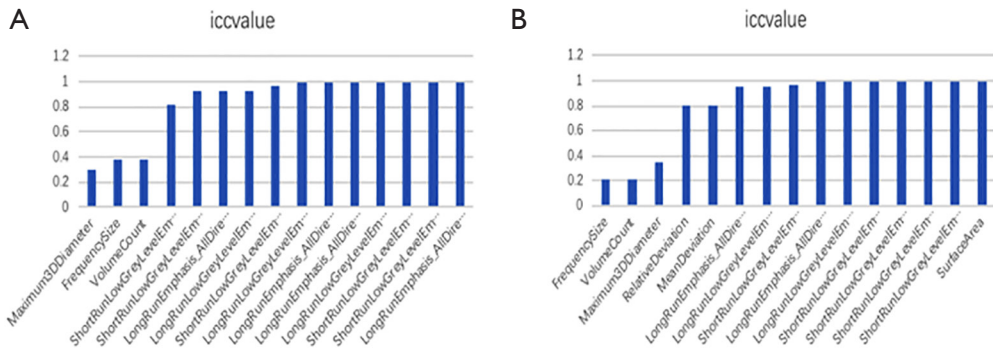
Figure S1 ICC values for inter-observer agreement assessment of LL and RL indicating excellent reliability and reproducibility of the nomogram for predicting EV severity, features with ICC values more than 0.87 (indicating excellent stability) were selected for subsequent investigation. ICC, intra-class correlation coefficient; EV, esophageal varices; LL, left liver; RL, right liver.

Table S1 The grading system for esophageal varices (Japanese Research Society for Portal Hypertension, Chinese Society of Digestive Endoscopy) and endoscopic results of our study

Grade	Endoscopic criterion		No. of patients*
	Form (F)	Red color sign (RC)	
Mild	F1	RC−	18
Moderate	F1	RC+	12
	F2	RC−	
Severe	F2	RC+	106
	F3	RC+ or RC−	

*, data is the number of patients of our result according to the grading system. F, form; RC, red color sign.
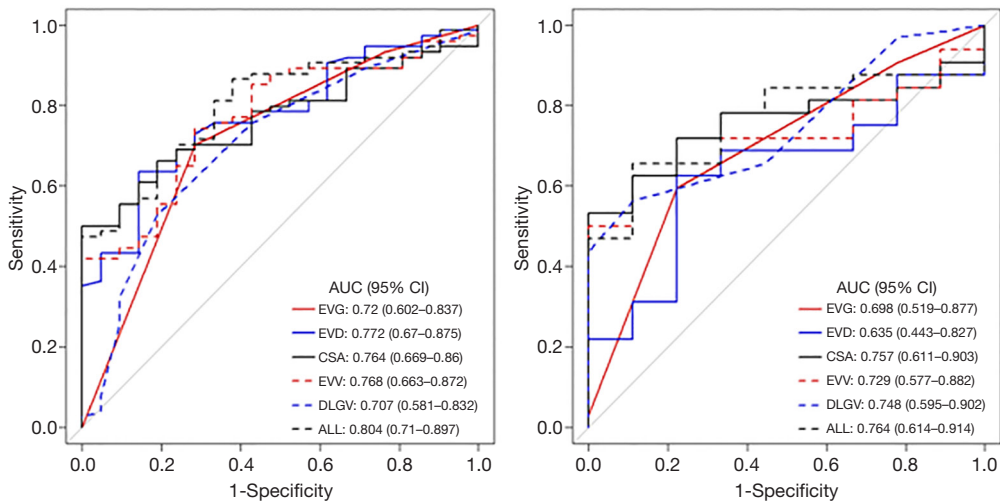

Figure S2 ROC analysis of each index in the training and validation cohorts. ROC, receiver operating characteristic.