Radiomics-Based Machine Learning Classification Strategy for Characterization of Hepatocellular Carcinoma on Contrast-Enhanced Ultrasound in High-Risk Patients with LI-RADS Category M Nodules

Lingling Li[1,#] Xiaoxin Liang[1,#] Yiwen Yu[1] Rushuang Mao[1] Jing Han[1] Chuan Peng[1,*] Jianhua Zhou[1,*]

[1] Department of Ultrasound, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, Guangdong, China

Indian J Radiol Imaging 2024;34:405–415.

Address for correspondence Jianhua Zhou, MD, PhD, Department of Ultrasound, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, 651 Dongfeng Road East, Guangzhou 510060, P.R. China (e-mail: zhoujh@sysucc.org.cn).

Chuan Peng, MD, Department of Ultrasound, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, 651 Dongfeng Road East, Guangzhou 510060, P.R. China (e-mail: pengchuan@sysucc.org.cn).

Abstract

Objective Accurate differentiation within the LI-RADS category M (LR-M) between hepatocellular carcinoma (HCC) and non-HCC malignancies (mainly intrahepatic cholangiocarcinoma [CCA] and combined hepatocellular and cholangiocarcinoma [cHCC-CCA]) is an area of active investigation. We aimed to use radiomics-based machine learning classification strategy for differentiating HCC from CCA and cHCC-CCA on contrast-enhanced ultrasound (CEUS) images in high-risk patients with LR-M nodules.

Methods A total of 159 high-risk patients with LR-M nodules (69 HCC and 90 CCA/cHCC-CCA) who underwent CEUS within 1 month before pathologic confirmation from January 2006 to December 2019 were retrospectively included (111 patients for training set and 48 for test set). The training set was used to build models, while the test set was used to compare models. For each observation, six CEUS images captured at predetermined time points (T1, peak enhancement after contrast injection; T2, 30 seconds; T3, 45 seconds; T4, 60 seconds; T5, 1–2 minutes; and T6, 2–3 minutes) were collected for tumor segmentation and selection of radiomics features, which included seven types of features: first-order statistics, shape (2D), gray-level co-occurrence matrix, gray-level size zone matrix, gray-level run length matrix,

Keywords
- contrast-enhanced ultrasound
- LI-RADS category M
- hepatocellular carcinoma
- machine learning

[#] *These authors contributed equally to this work.*
[*] *These are co-corresponding authors.*

article published online
January 17, 2024

DOI https://doi.org/10.1055/s-0043-1777993.
ISSN 0971-3026.

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India

neighboring gray tone difference matrix, and gray-level dependence matrix. Clinical data and key radiomics features were employed to develop the clinical model, radiomics signature (RS), and combined RS-clinical (RS-C) model. The RS and RS-C model were built using the machine learning framework. The diagnostic performance of these three models was calculated and compared.

Results Alpha-fetoprotein (AFP), CA19-9, enhancement pattern, and time of washout were included as independent factors for clinical model (all $p < 0.05$). Both the RS and RS-C model performed better than the clinical model in the test set (area under the curve [AUC] of 0.698 [0.571–0.812] for clinical model, 0.903 [0.830–0.970] for RS, and 0.912 [0.838–0.977] for the RS-C model; both $p < 0.05$).

Conclusions Radiomics-based machine learning classifiers may be competent for differentiating HCC from CCA and cHCC-CCA in high-risk patients with LR-M nodules.

Introduction

Hepatocellular carcinoma (HCC) is the most prevalent subtype of primary liver cancer, accounting for approximately 85 to 90% of primary liver cancers.[1] Unlike most malignant tumors, the diagnosis of HCC with typical imaging hallmarks can be confirmed by noninvasive diagnostic criteria.[2,3] Since the treatment options and prognoses are different for HCC and non-HCC malignancies (including intrahepatic cholangiocarcinoma [CCA] and combined hepatocellular and cholangiocarcinoma [cHCC-CCA]), imaging-based differentiation of HCC from non-HCC malignancies is important.

Contrast-enhanced ultrasound (CEUS) with microbubbles is able to characterize focal liver lesions (FLLs) in real time by continuous imaging. CEUS Liver Imaging Reporting and Data System (LI-RADS) was developed for standardization of reporting and data collection of imaging techniques.[4] This system assesses the likelihood of a liver observation to be HCC using a 5-point scale, ranging from LR-1 (definitely benign) to LR-5 (definitely HCC). In particular, observations that are probably or definitely malignant but not necessarily HCC are classified as an additional category, LR-M. Several studies have already confirmed that the vast majority of CCA and cHCC-CCA are characterized as LR-M.[5,6] Nevertheless, to ensure high specificity of LR-5 for the diagnosis of HCC, considerable numbers of HCC were also categorized as LR-M rather than LR-5. This, however, results in limited sensitivity of LR-5 for HCC.[7] It is worth noting that radical resection or locoregional ablative therapies are preferred over liver transplant for treatment of HCC in many Asian countries, so maximal sensitivity is anticipated when it comes to an HCC diagnostic algorithm in these areas.[8] In other words, reduced sensitivity of the LI-RADS system for HCC limits its clinical application in countries and regions that rely primarily on local treatments. In addition, the mainstay management for CEUS LR-M nodules is needle biopsy,[4] the accuracy of which is primarily limited by intratumoral heterogeneity and poor technique. A high proportion of HCC in the subset of LR-M nodules may lead to an increase in invasive procedures and medical burden. To address these issues, new strategies aimed at characterizing HCC in the subset of LR-M nodules are therefore required.

Radiomics is a novel tool that noninvasively extracts quantitative information on cancer hallmarks from images, thus constituting an image-based biomarker (named radiomics signature [RS]) for accurate diagnosis.[9,10] Most radiomic studies rely on traditional machine learning techniques in the radiomics feature selection and model building step.[11,12] To the best of our knowledge, there is no strong evidence to indicate that CEUS-based RS or combined RS-clinical (RS-C) model can be employed to characterize HCC in the subset of LR-M nodules. Therefore, the purpose of this study was to investigate whether CEUS-based RS or RS-C model could be applied to preoperatively differentiate HCC from CCA and cHCC-CCA within the CEUS LR-M category in high-risk patients.

Materials and Methods

Study Population

The study was approved by the institutional review board of Sun Yat-Sen University Cancer Center, and the requirement for written informed patient consent was waived as this study involved only retrospective analysis of previously collected data.

Patients who underwent CEUS examinations for characterizing FLLs from January 2006 to December 2019 were screened. The inclusion criteria were as follows: (1) at high risk for HCC; (2) observations diagnosed as HCC, CCA, or cHCC-CCA via biopsy and surgery resection; (3) patients who underwent CEUS within 30 days before histopathological diagnosis; and (4) if they met any of the LR-M criteria (rim arterial phase hyperenhancement [APHE], early washout, or marked washout), as described in CEUS LI-RADS v2017. A subject was excluded from the study if they fulfilled any of the following criteria: (1) previous treatment for HCC and (2) incomplete information of potential covariables, or degraded images that did not satisfy the requirements for analysis. A flowchart depicting the enrollment process and reasons for exclusion in this study is shown in ►**Fig. 1**. A total of 159 patients were included in the final analysis. The study population was then randomly divided into two sets,

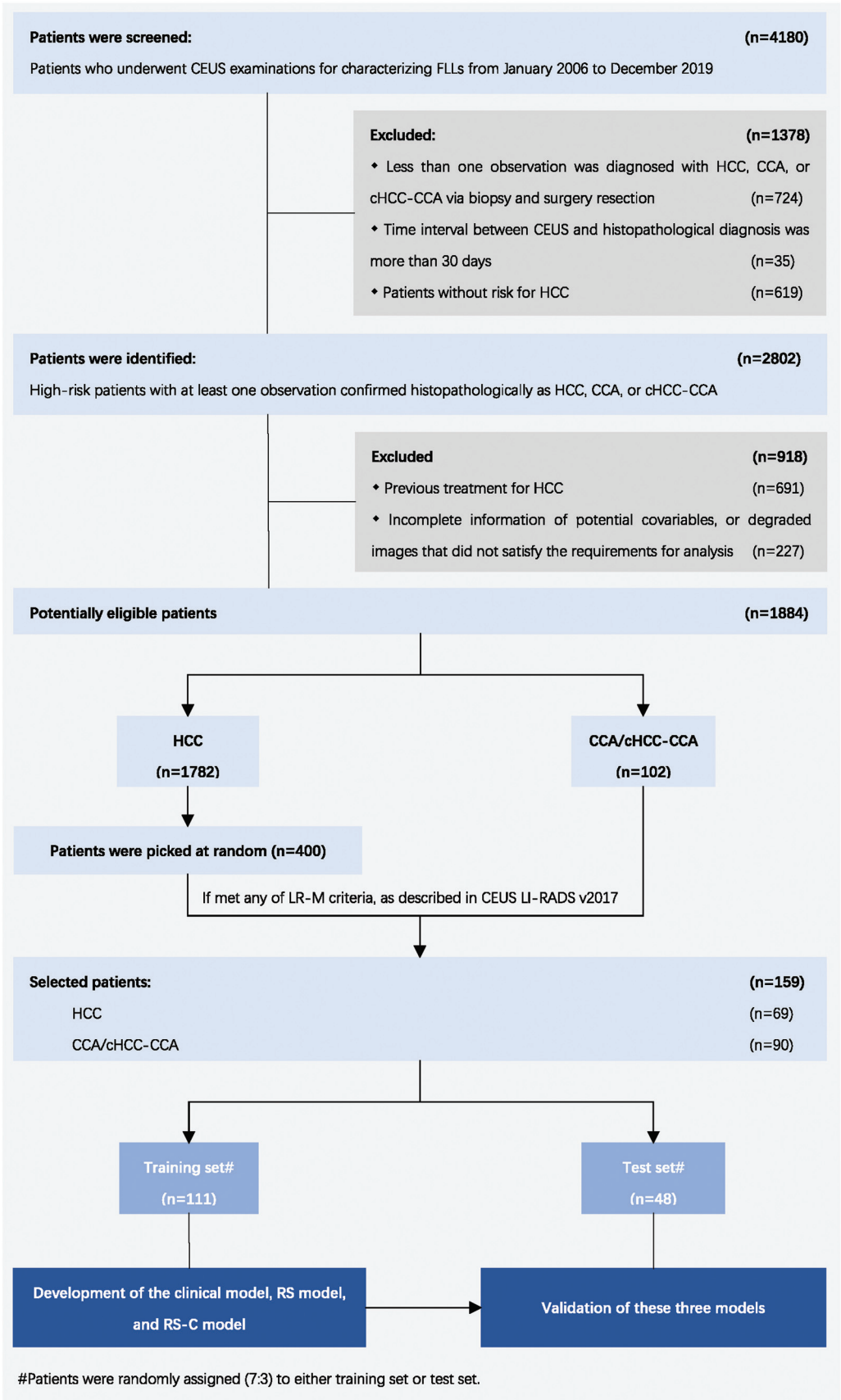

Fig. 1 The flowchart of study selection. CCA, intrahepatic cholangiocarcinoma; cHCC-CCA, combined hepatocellular and cholangiocarcinoma; CEUS, contrast-enhanced ultrasound; FLL, focal liver lesion; HCC, hepatocellular carcinoma.

including a training set ($n = 111$; mean age, 52 ± 11 years; male, 90 [81.1%]; and HCC:CCA/cHCC-CCA = 48/63) and a test set ($n = 48$; mean age, 53 ± 11 years; male, 40 [83.3%]; and HCC:CCA/cHCC-CCA = 21/27). The training set was used for feature selection, dimension reduction of radiomics features, and model building, while the test set was used for comparing the diagnostic performance of models.

Clinical Characteristics and Clinical Model

The clinical data of patients were collected from the hospital information system, which included demographic variables (age, gender, cirrhosis), imaging characteristics based on CEUS (number of lesions, nodule size, enhancement patterns, time of washout, and degree of washout), and laboratory tests (alpha-fetoprotein [AFP], and glycoprotein antigen 19-9 [CA19-9]). All these latent covariables were candidates for the clinical model. Factors that were significant at the 0.1 level from the univariate logistics regression analysis were included in the multivariable logistics regression analysis. Those with a *p*-Value less than 0.05 in the multivariable analysis were included as independent variables in the clinical model.

CEUS Examinations

All CEUS examinations including initial B-mode evaluation were performed using (1) Acuson Sequoia 512 (Siemens Medical Solutions, Mountain View, CA, United States) with a 4C1 convex array probe, (2) Acuson Sequoia system (Siemens Medical Solutions) with a 5C1 convex array probe, (3) Acuson S2000 (Siemens Medical Solutions) with a 6C1 convex array probe, and (4) Philips iU22 (Royal Philips Electronics, Amsterdam, the Netherlands) with a 5C1 convex array probe. In this study, grayscale ultrasound (US) was initially performed to localize lesions and assess their conventional US characteristics. SonoVue (Bracco, Milan, Italy) at a dose of 2.0 mL was administrated intravenously for CEUS examinations. CEUS was performed using a contrast pulse sequencing mode with a low mechanical index of 0.06 to 0.08. The focus was set at the posterior acoustic field. A timer was started immediately after injection of the suspension. Continuous imaging was recorded as a cine loop for the first approximately 70 seconds after injection, followed by intermittent imaging recorded every 20 to 30 seconds for the remaining 5 minutes after injection. These recordings were used to provide arterial phase (from 10–20 to 30–45 seconds), portal venous phase (from 30–45 seconds to 2 minutes), and late phase (from 2–5 minutes) images. In this study, images captured at six predetermined time points (T1, visually identified peak enhancement of contrast-enhancing area after contrast injection; T2, 30 seconds; T3, 45 seconds; T4, 60 seconds, T5, 1–2 minutes; and T6, 2–3 minutes) were collected for radiomics analysis.

Radiological Assessment

Two readers (R.M. and J.H., with 5 and 13 years of posttraining experience, respectively, in abdominal imaging including liver CEUS) independently reviewed all the US and CEUS images without prior knowledge of the medical and surgical history, laboratory results, computed tomography (CT) and/or magnetic resonance imaging (MRI) findings, and pathologic results. Disagreements between the two readers were resolved through a consensus assessment by a third reader (J.Z., 21 years of posttraining experience in abdominal imaging including liver CEUS). First, the readers evaluated the number, location, nodule size, and echo of lesions on grayscale US images. In patients with multiple lesions, the largest lesion with available histopathological assessment was selected as a target observation for further evaluation. Then the following features for each observation were recorded: (1) enhancement patterns (recorded as no APHE, nonrim APHE, or rim APHE), (2) specific timing of washout (also classified as early washout [<60 seconds] or late washout [≥60 seconds]), and (3) degree of washout (recorded as mild washout or marked washout [<2 minutes]). The readers also assigned each observation a CEUS LI-RADS category. According to CEUS LI-RADS v2017,[4] an observation is assigned to the LR-M category if it meets any of the following CEUS LR-M criteria: rim APHE, early washout, or marked washout.

RS and RS-C Model

Tumor Segmentation and Feature Extraction

After the CEUS images mentioned in the "CEUS examinations" section were loaded into Novo Ultrasound Kit v1.0.0.R software (GE Healthcare, UK), manual segmentation of each observation was performed by two radiologists (Y.Y. and C.P., with 3 and 8 years of posttraining experience, respectively, in abdominal imaging including liver CEUS) independently in a blinded manner using the "Segment" function. The radiomics features of region of interest (ROI) that covered the whole observation are extracted using the "USomics" module in the same software, which included seven types of features as follows: (1)first-order statistics, (2) shape (2D), (3) gray-level co-occurrence matrix (GLCM), (4) gray-level size zone matrix (GLSZM), (5) gray-level run length matrix (GLRLM), (6) neighboring gray tone difference matrix (NGTDM), and (7) gray-level dependence matrix (GLDM). Three transformations (i.e., Laplacian of gaussian [LOG], wavelet transform [WT], and local binary pattern [LBP]) were applied to the original images, yielding the derived images. All radiomics features, except for shape, were computed from both the original images and derived images. In total, 7,896 radiomics features, divided in six sets (T1–T6) of 1,316 features each, were calculated for every observation.

Feature Selection, Dimension Reduction, and Modeling

The machine learning-based approach was used in the process of feature selection, dimension reduction of radiomics features, and model building. Prior to analysis, the data were standardized using *z*-score normalization. The interreader agreements of radiomics features were assessed with intraclass correlation coefficients (ICC). A total of 572 features with good reproducibility (ICC > 0.75) were considered for inclusion in the further feature selection session. The top 20 features with the highest maximum relevance minimum

redundancy (mRMR) scores were selected. The least absolute shrinkage and selection operator (LASSO) algorithm was applied to select key radiomics features with highest contributions to classification (►**Supplementary Fig. S1**). The multistep process of feature selection and dimension reduction is shown in ►**Supplementary Fig. S2**. Logistic regression was then utilized to build the RS. The primary outcome measure (as described in the "Statistical Analysis" section) of this RS after three train–test splits were calculated to study the robustness of the RS. It, together with independent variables of the clinical model, comprises a combined RS-C model. A nomogram was constructed for presentation of the final RS-C model. The risk probability could be calculated using the nomogram.

Statistical Analysis

Characteristics were compared between the entire study population and training set, as well as the training set and test set. Continuous data are presented as mean and standard deviation, and categorical data are presented as frequencies and percentages. Groups were compared using *t*-test for continuous variables and the chi-squared test or Fisher's exact test for categorical test. Receiver operating characteristics (ROC) analyses for the clinical model, RS, and RS-C model for both the training set and the test set were performed; and true positive, false positive, false negative, true negative, area under the curve (AUC), sensitivity, specificity, and accuracy were calculated. AUC was considered the primary outcome measure, and DeLong's test was used for pairwise comparisons of AUC. Decision curve analyses were used to assess the clinical utility of the models, and calibration curves and Hosmer–Lemeshow (H-L) test for goodness of fit. Statistical tests were conducted using "SciPy" and "Statsmodels" packages, and the machine learning–based classifiers were implemented using the "Sklearn" toolkit in Python version 3.8.1. A two-tailed *p*-value of less than 0.05 indicated statistical significance, unless otherwise stated.

Results

Patient Characteristics

Demographic variables, imaging characteristics, laboratory tests, and pathological findings for the 159 patients with LR-M nodules (69 HCC and 90 CCA/cHCC-CCA) are shown in ►**Table 1**. There was no significant difference in the clinical characteristics between the entire study population and training set, as well as the training set and the test set (all $p > 0.05$).

Clinical Model

Results of the univariate and multivariate analysis are shown in ►**Table 2**. Four independent factors identified as significant in multivariate analysis were included in the clinical model. Compared to patients with CCA/cHCC-CCA, patients with HCC had higher AFP levels ($p = 0.002$) and lower CA19-9 levels ($p = 0.032$). In the subset of LR-M nodules, HCC lesions less commonly presented with rim APHE ($p = 0.016$) and washout within 30 seconds ($p = 0.015$) than CCA and cHCC-CCA.

RS and RS-C Model

An overview of the modeling process is shown in ►**Fig. 2**. The top five radiomics features with highest contribution to model discrimination are listed in ►**Table 3**. These were T3_LOG_GLRLM_ ShortRunLowGrayLevelEmphasis (RF1), T2_LOG_GLDM_LowGrayLevelEmphasis (RF2), T2_Wavelet_HL_GLDM_DependenceEntropy (RF3), T6_Wavelet_LH_GLCM_JointEntropy (RF4), and T2_Original_ Firstorder_Skewness (RF5). Significant asymmetries of these selected radiomics features between HCC and CCA/cHCC-CCA group were observed (►**Fig. 3**). RS was established based on these five radiomics features using the logistic regression algorithm. The computed formula was as follows:

$$RS = -1.016 \times RF1 + 0.853 \times RF2 - 0.977 \times RF3 - 0.84 \times RF4 - 1.109 \times RF5 - 0.59.$$

The robustness of RS was validated by calculating the AUC in different training sets and test sets after three train–test splits (►**Supplementary Table S1**).

The combined RS-C model included four independent clinical factors and RS. A nomogram constructed for presentation of the final RS-C model is shown in ►**Fig. 4**. The final classification model (i.e., RS-C model) was created from the nomogram by thresholding the nomogram output probabilities at a value of 0.359 (►**Fig. 4**). The optimal threshold was determined based on maximum Youden index. Patients whose risk score exceeds the predetermined risk threshold were diagnosed as HCC.

Performance of the Models

A range of diagnostic indexes of clinical model, RS, and RS-C model are estimated in ►**Table 4**, and results of DeLong's test are shown in ►**Supplementary Table S2**. The results of the ROC curve analysis for the clinical model, RS, and RS-C model in the training set and the test set are displayed in ►**Fig. 5A** and **B**, respectively. In the test set, the clinical model achieved an AUC of 0.698 (0.571–0.812). Both RS and RS-C model yielded better performance than the clinical model in the test set (0.903 [0.830–0.970] for RS and 0.912 [0.838–0.977] for RS-C model, both DeLong's test $p < 0.05$).

In the decision curve analyses, both RS and the RS-model resulted in higher net benefit than the clinical model (►**Fig. 5C**). As shown in ►**Fig. 5D**, the majority of the calibration curves followed the diagonal line for both the training set (H-L test; $p = 0.087$) and the test set (H-L test, $p = 0.288$), indicating reliable risk estimates of the nomogram.

Discussion

In the present study, machine learning–based strategy was utilized for classification in the subset of CEUS LR-M nodules. Compared to the clinical model (AUC: 0.698), both RS (AUC: 0.903; $p = 0.018$) and the RS-C model (AUC: 0. 912; $p = 0.003$) showed a higher discriminatory ability to correctly classify observations as HCC or CCA/cHCC-CCA within the subset of CEUS LR-M nodules in the test set.

Table 1 The main characteristics of patients and observations

Parameters	All patients ($n = 159$)	Training set ($n = 111$)	Test set ($n = 48$)	*p1*	*p2*
Age (y)[a]	52 ± 12	52 ± 11	53 ± 11	0.79	0.53
Male	130 (81.8)	90 (81.1)	40 (83.3)	> 0.99	0.91
Cirrhosis	68 (42.8)	44 (39.6)	24 (50.0)	0.70	0.30
AFP > 20 μg/L	71 (44.7)	49 (44.1)	22 (45.8)	> 0.99	0.98
CA19-9 > 35 μg/mL	53 (33.3)	35 (31.5)	18 (37.5)	0.86	0.58
Number of lesions					
One	139 (87.4)	97 (87.4)	42 (87.5)	0.50	0.91
Two	8 (5.0)	6 (5.4)	2 (4.2)		
More than two	12 (7.6)	8 (7.2)	4 (8.3)		
Nodule size					
≤20 mm	3 (1.9)	1 (0.9)	2 (4.2)	0.28	0.27
20–50 mm	81 (50.9)	55 (49.6)	26 (54.2)		
≥50 mm	75 (47.2)	55 (49.6)	20 (41.7)		
Echo demonstrated by ultrasonography					
Hypoechoic	132 (83.0)	94 (84.7)	38 (79.2)	0.94	0.46
Isoechoic	11 (6.9)	7 (6.3)	4 (8.3)		
Hyperechoic	16 (10.1)	10 (9.0)	6 (12.5)		
Enhancement patterns					
No APHE	5 (3.1)	2 (1.8)	3 (6.3)	0.93	0.71
Nonrim APHE	135 (84.9)	94 (84.7)	41 (85.4)		
Rim APHE	19 (12.0)	15 (13.5)	4 (8.3)		
Time of washout					
< 30 sec	29 (18.2)	20 (18.0)	9 (18.8)	> 0.99	> 0.99
30sec-1min	130 (81.8)	91 (82.0)	39 (81.3)		
Degree of washout (<2 min)					
Marked washout	24 (15.1)	18 (16.2)	6 (12.5)	0.94	0.72
Mild washout	135 (84.9)	93 (83.8)	42 (87.5)		
Pathologic findings					
HCC	69 (43.4)	48 (43.2)	21 (43.8)	> 0.99	> 0.99
CCA/cHCC-CCA	90 (56.6)	63 (56.9)	27 (56.3)		

Abbreviations: AFP, alpha-fetoprotein; APHE, arterial phase hyperenhancement; CA19-9, glycoprotein antigen 19-9; CCA, cholangiocarcinoma; cHCC-CCA, combined hepatocellular-cholangiocarcinoma; HCC, hepatocellular carcinoma.
[a]Data are mean ± standard deviation.
Note: *p1* = training set versus all patients; *p2* = test set versus training set.
Unless otherwise indicated, data are numbers with percentages in parentheses.

Although major success was achieved, challenges and gaps of LI-RADS remain. Limited sensitivity of LR-5 for HCC is one of the main issues. Sensitivity of LR-5 for HCC is sacrificed to maintain a high specificity, resulting in some HCC being characterized as LR-4, LR-M, or rarely LR-3.[7] It was reported that a substantial number of LR-M observations for which biopsy is usually performed to definitively diagnose were HCC.[13,14] As Chernyak stated,[15] accurate differentiation within the LR-M category between HCC and non-HCC malignancies based on noninvasive imaging or other effective means is an area of active investigation. The solutions to the above-mentioned investigation might be helpful to reduce unnecessary invasive procedures and increase sensitivity of the LR-5 category for HCC. However, the clinical model achieved poor performance for characterization of HCC within the subset of CEUS LR-M nodules, with an AUC of 0.698 (0.571–0.812) in the test set. The present study is dedicated to characterize HCC in the subset of CEUS LR-M nodules using a machine learning strategy. The RS and RS-C model performed well (AUC of 0.903 for RS and 0.912 for the RS-C model). Hence, this study is expected to achieve optimal clinical management of HCC.

In this study, five radiomics features that contribute the most to the classifier outcome were filtered out. None of

Table 2 Univariate and multivariate logistic regression analysis of independent variables for differential diagnosis of HCC and CCA/cHCC-CCA in high-risk patients with LR-M nodules

Parameters	Univariate logistic regression analysis		Multivariate logistic regression analysis		Clinical model	
	OR	*p*	OR	*p*	OR	*p*
Age (y)	0.99 (0.97–1.03)	0.943				
Male	6.00 (1.65–49.21)	0.564				
Cirrhosis	1.84 (0.85–3.98)	0.121				
AFP > 20 μg/L	4.56 (2.04–10.20)	<0.001[a]	4.17 (1.65–10.52)	0.002[a]	4.40 (1.76–10.99)	0.002[a]
CA19-9 > 35 μg/mL	0.33 (0.14–0.79)	0.013[a]	0.31 (0.10–0.92)	0.035[a]	0.30 (0.10–0.90)	0.032[a]
Number of lesions	0.89 (0.41–1.67)	0.600				
Nodule size (mm)	0.95 (0.46–1.97)	0.897				
Echo demonstrated by ultrasonography	1.96 (0.71–5.37)	0.192				
Enhancement patterns	0.07 (0.01–0.51)	0.009[a]	0.10 (0.01–0.89)	0.039[a]	0.07 (0.01–0.62)	0.016[a]
Time of washout	5.54 (1.52–20.23)	0.010[a]	4.78 (1.10–20.74)	0.036[a]	5.71 (1.40–23.24)	0.015[a]
Degree of washout (<2 min)	17.34 (2.22–135.50)	0.007[a]	6.18 (0.65–58.91)	0.114		

Abbreviations: CCA, cholangiocarcinoma; cHCC-CCA, combined hepatocellular-cholangiocarcinoma; HCC, hepatocellular carcinoma.
[a]$p < 0.05$.

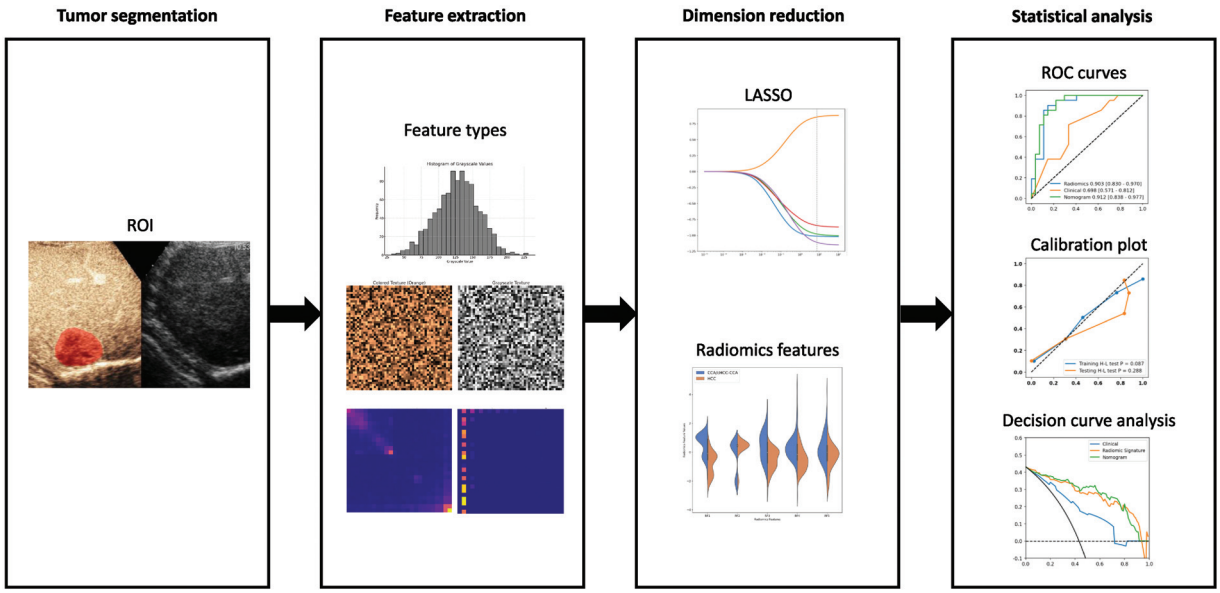

Fig. 2 Overview of modeling process. LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristics; ROI, region of interest.

these were shape features; one was first-order statistics, which assessed the distribution of grayscale pixel intensities within ROI; and four of the five features were texture features that sensitively reflected intratumoral heterogeneity, cellular density, and level of vascularization.[16–18] The above underlying phenotypic and pathophysiologic characteristics can help predict the likely lineage of origin and thus benefit differential diagnosis of HCC and CCA/cHCC-CCA. These subtle nuances from medical images, difficult to appreciate by visual inspection even for experienced radiologists, can be found by radiomics analysis. In this regard, radiomics analysis possesses its unique advantages in supporting clinical application scenarios by high-throughput extraction of numerous quantitative features.

The role of radiomics analysis in the differentiation of HCC from non-HCC lesions has recently received increasing

Table 3 The definitions of the five most contributive radiomics features

Radiomics feature	Time point	Feature type	Transformation	Definition
RF1	T3	GLRLM	LOG	Short-run low gray level emphasis
RF2	T2	GLDM	LOG	Low gray level emphasis
RF3	T2	GLDM	Wavelet_HL	Dependence entropy
RF4	T6	GLCM	Wavelet_LH	Joint entropy
RF5	T2	First order	Original	Skewness

Abbreviations: GLCM, gray level co-occurrence matrix; GLDM, gray level dependence matrix; GLRLM, gray level run length matrix; LOG, Laplacian of gaussian; RF, radiomics feature.

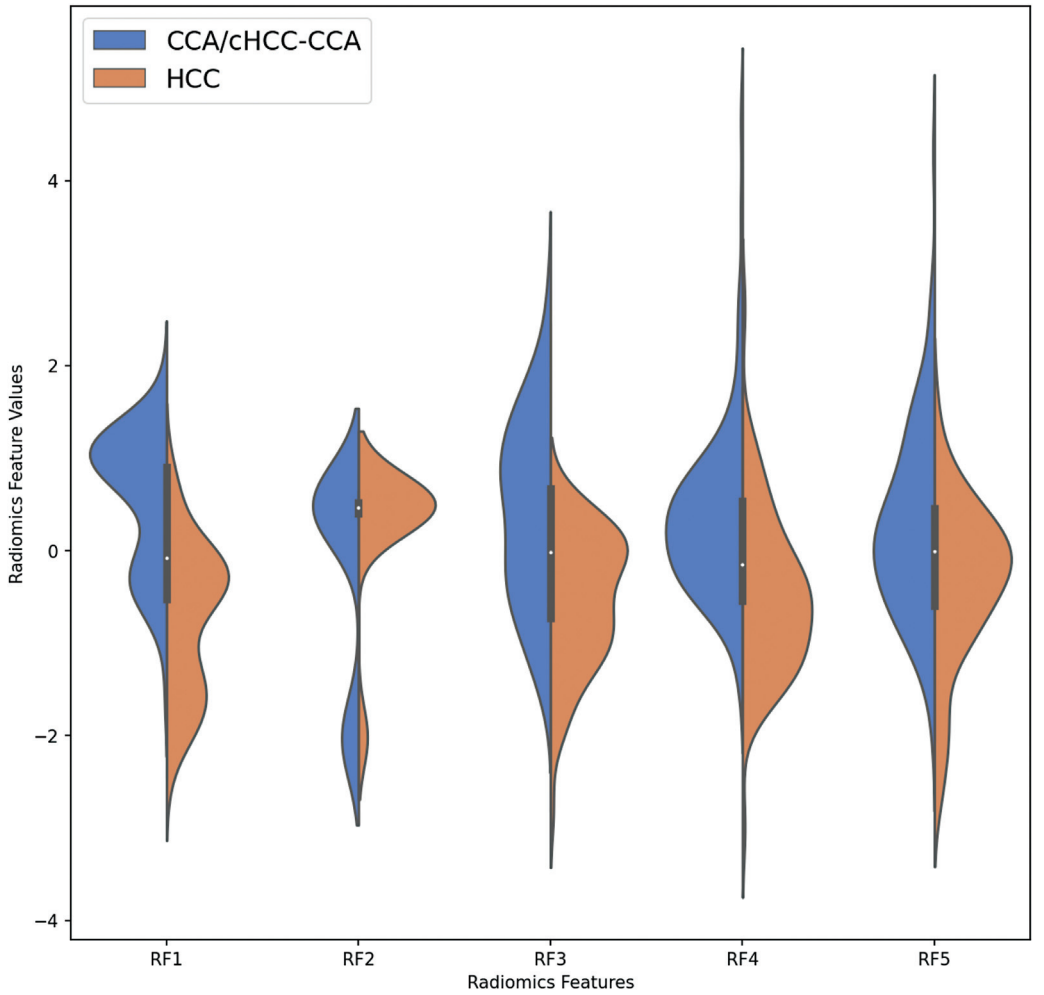

Fig. 3 Violin plot showing significant asymmetries of the five selected radiomics features between HCC and CCA/cHCC-CCA group. CCA, intrahepatic cholangiocarcinoma; cHCC-CCA, combined hepatocellular and cholangiocarcinoma; HCC, hepatocellular carcinoma; RF, radiomics feature.

attention. Mokrane et al[19] established a radiomics machine learning signature based on CECT images from 178 patients for HCC diagnosis in cirrhotic patients with indeterminate nodules. However, this model had poor performance, with an AUC of only 0.66 in the validation cohort. Another study with 668 patients (531 HCC, 48 cHCC-CCA, and 89 CCA) built a US-based radiomics model to preoperatively predict the histopathological subtypes of primary liver cancers, with an AUC of only 0.775 in the test cohort.[20] The RS and RS-C model built in this study demonstrated better performance compared to the above models (AUC: 0.903 for RS and 0.912 for RS-C model in the test set). Nevertheless, specificity of the predictions was limited (85.2% for RS and 77.8% for the RS-C model in the test set). Deep learning that employs multilayer neural networks, a new branch of machine learning, provides opportunities to analyze images at greater depths, and it is expected to further improve model discrimination.

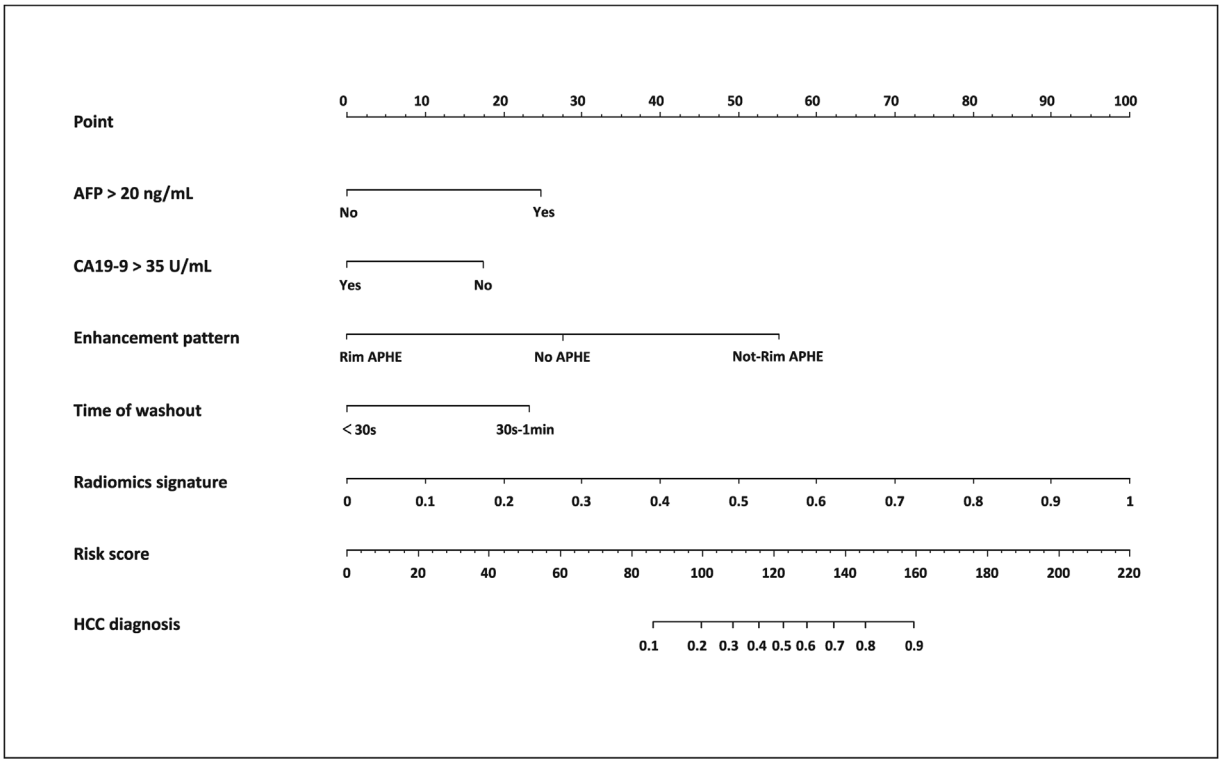

Fig. 4 A nomogram constructed for presentation of the final radiomic signature-clinical model. AFP, alpha-fetoprotein; APHE, arterial phase hyperenhancement; CA19-9, glycoprotein antigen 19-9.

Table 4 The diagnostic accuracy of the clinical model, RS, and RS-C model in the training set and test set

	TP	FP	FN	TN	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
Training set								
Clinical model	31	12	17	51	0.811 (0.747–0.873)	73.9 (65.0–81.2)	64.6 (49.4–77.4)	81.0 (72.9–88.4)
RS	45	12	3	51	0.932 (0.891–0.965)	86.5 (78.9–91.6)	93.8 (81.8–98.4)	81.0 (72.6–88.9)
RS-C model	41	4	7	59	0.952 (0.919–0.979)	90.0 (83.1–94.4)	85.4 (71.6–93.5)	93.7 (83.7–97.9)
Test set								
Clinical model	15	9	6	18	0.698 (0.571–0.812)	68.8 (54.7–80.0)	71.4 (47.7–87.8)	66.7 (46.0–82.8)
RS	19	4	2	23	0.903 (0.830–0.970)	87.5 (75.3–94.1)	90.5 (68.2–98.3)	85.2 (65.4–95.1)
RS-C model	20	6	1	21	0.912 (0.838–0.977)	85.4 (72.9–92.8)	95.2 (74.1–99.8)	77.8 (57.3–90.6)

Abbreviations: AUC, area under the curve; FN, false negative; FP, false positive; RS, radiomics signature; RS-C model, radiomics signature-clinical model; TN, true negative; TP, true positive.

There are also several limitations. First, this study was a single-center retrospective study with a small sample size. Owing to this, potential selection bias might exist. Furthers studies are required to confirm all these findings. Second, precise modeling depends upon the implementation of accurate and rapid segmentation of tumor. However, manual segmentation employed in this study is experience dependent, laborious, and time- and energy-consuming. Automatic segmentation with minimal need for user input is more efficient and more desirable. Finally, due to technical difficulty of feature extraction from cine loops, images captured at predetermined time points, instead of cine loops, were selected for radiomics analysis.

In conclusion, machine learning–based classification strategy has the potential to differentiate HCC from CCA and cHCC-CCA in high-risk patients with CEUS LR-M nodules.

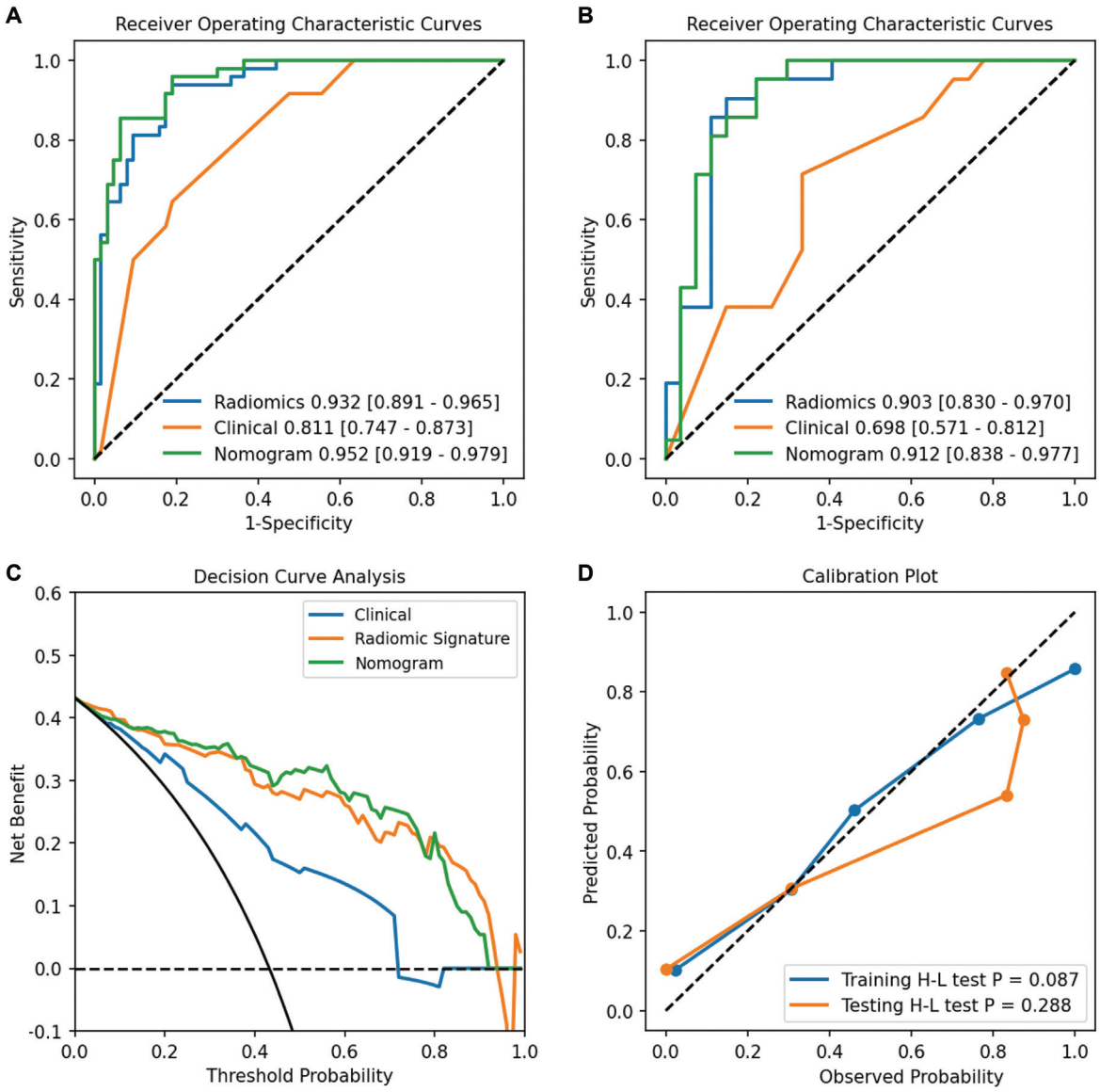

Fig. 5 Performance of models. (**A**) Receiver operating characteristics (ROC) curves for the clinical model, radiomics signature (RS), and combined RS-clinical (RS-C) model (i.e., the nomogram) in the training set. (**B**) ROC curves for the clinical model, RS, and nomogram in the test set. (**C**) Decision curve analyses for the clinical model, RS, and nomogram in the test set. (**D**) Calibration curves of the nomogram. H-L test, Hosmer–Lemeshow test.

Funding
None.

Conflict of Interest
None declared.

References

1 Llovet JM, Zucman-Rossi J, Pikarsky E, et al. Hepatocellular carcinoma. Nat Rev Dis Primers 2016;2(16018):16018
2 Heimbach JK, Kulik LM, Finn RS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology 2018;67(01): 358–380
3 European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu. European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69(01):182–236
4 American College of Radiology website. CEUS LI-RADS 2017 Core. 2017. Accessed March 22, 2022 at: www.acr.org/-/media/ACR/Files/RADS/LI-RADS/CEUS-LI-RADS-2017-Core.pdf
5 Zheng W, Li Q, Zou XB, et al. Evaluation of contrast-enhanced US LI-RADS version 2017: application on 2020 liver nodules in patients with hepatitis B infection. Radiology 2020;294(02): 299–307
6 Terzi E, Iavarone M, Pompili M, et al; CEUS LI-RADS Italy study group collaborators. Contrast ultrasound LI-RADS LR-5 identifies hepatocellular carcinoma in cirrhosis in a multicenter retrospective study of 1,006 nodules. J Hepatol 2018;68(03):485–492
7 Rodgers SK, Fetzer DT, Kono Y. Using LI-RADS with contrast-enhanced ultrasound. Clin Liver Dis (Hoboken) 2021;17(03): 154–158

8 Chernyak V, Fowler KJ, Kamaya A, et al. Liver Imaging Reporting and Data System (LI-RADS) version 2018: imaging of hepatocellular carcinoma in at-risk patients. Radiology 2018;289(03): 816–830
9 Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749–762
10 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(06):1191–1206
11 Lee SH, Park H, Ko ES. Radiomics in breast imaging from techniques to clinical applications: a review. Korean J Radiol 2020;21 (07):779–792
12 Peng J, Kang S, Ning Z, et al. Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging. Eur Radiol 2020;30(01):413–424
13 Li L, Hu Y, Han J, Li Q, Peng C, Zhou J. Clinical application of liver imaging reporting and data system for characterizing liver neoplasms: a meta-analysis. Diagnostics (Basel) 2021;11 (02):323
14 Shin J, Lee S, Bae H, et al. Contrast-enhanced ultrasound liver imaging reporting and data system for diagnosing hepatocellular carcinoma: A meta-analysis. Liver Int 2020;40(10):2345–2352
15 Chernyak V, Fowler KJ, Do RKG, et al. LI-RADS: looking back, looking forward. Radiology 2023;307(01):e222801
16 Kim JH, Ko ES, Lim Y, et al. Breast cancer heterogeneity: MR imaging texture analysis and survival outcomes. Radiology 2017; 282(03):665–675
17 Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. Radiology 2013;266(01):326–336
18 Fujima N, Homma A, Harada T, et al. The utility of MRI histogram and texture analysis for the prediction of histological diagnosis in head and neck malignancies. Cancer Imaging 2019;19(01):5
19 Mokrane FZ, Lu L, Vavasseur A, et al. Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules. Eur Radiol 2020;30 (01):558–570
20 Peng Y, Lin P, Wu L, et al. Ultrasound-based radiomics analysis for preoperatively predicting different histopathological subtypes of primary liver cancer. Front Oncol 2020;10:1646

