Discover Oncology

RESEARCH Open Access

A T2 weighted imaging-based radiomics nomogram for the classification of hepatic blood-rich lesions: hepatocellular carcinoma and benign liver lesions

Chen Yang[1†], Boburjon Turgunov[2†], Fandong Zhu[1], Yang Xia[3], Jianfeng Zhou[4] and Zhenhua Zhao[1*]

†Chen Yang and Boburjon Turgunov are co-authors.

*Correspondence:
Zhenhua Zhao
zhao2075@163.com
[1]Department of Radiology, Shaoxing People's Hospital, No. 568, North Zhongxing Road, Yuecheng District, Shaoxing City, Zhejiang, China
[2]Shaoxing University School of Medicine, Shaoxing, China
[3]Department of Radiology, Shaoxing Women and Child's Health Hospital, Shaoxing, China
[4]Department of Radiology, Zhuji Second People's Hospital, Zhuji, China

Abstract

Purpose To develop a radiomics nomogram based on T2-weighted imaging (T2WI) to distinguish hepatocellular carcinoma (HCC) from blood-rich benign liver lesions (BLLs) and to assess its value as an adjunct to conventional imaging.

Methods This study analyzed imaging and clinical data from 220 patients with pathologically confirmed HCC ($n=140$) and BLLs ($n=80$). Patients were divided into training cohort ($n=100$), test cohort ($n=44$), and validation cohort ($n=76$). We developed a radiomics model, a clinical model, and a fusion model integrating Rad-score with clinical factors, and identified the best predictive model among them in the test set, which was then compared with the diagnostic efficiency of two radiologists. The efficacy was assessed by the area under the receiver operating characteristic curve (AUC).

Results Four radiomics features and two clinical factors (age, liver cirrhosis) were selected for the construction of radiomics and clinical models, respectively. The fusion model exhibited superior discriminative ability in the training set (AUC = 0.984), test set (AUC = 0.948), and validation set (AUC = 0.955). In the validation set, a statistical difference was observed between the fusion model and the radiomics model ($P=0.001$). Furthermore, the diagnostic performance of the fusion model based on T2WI was comparable to that of conventional imaging based on enhanced MRI. In the cohort with conventional imaging errors, the AUCs of the clinical model, radiomics model, and fusion model were 0.841, 0.737, and 0.905, respectively.

Conclusions The T2WI-based radiomics nomogram effectively differentiated HCC from blood-rich BLLs, and might be an adjunct to conventional imaging.

Keywords Hepatocellular carcinoma, Liver lesion, Blood-rich, Radiomics, Magnetic resonance imaging

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

1 Background

Liver cancer, the second leading cause of cancer-related deaths in China, was recorded over 900,000 new cases worldwide in 2020, half of which will occur in China, and this alarming trend is projected to exceed one million cases by 2025 [1, 2]. Hepatocellular carcinoma (HCC) constitutes 85–90% of primary liver cancers, and its early detection is clinically critical and poses enormous challenges [2]. The detection and diagnosis of HCC are mainly based on imaging examinations with characteristic features of hyperenhancement in the arterial phase (AP) followed by hypointensity in the portal-venous phase (PVP) or delayed phase (DP) [3], but it is not specific. On the one hand, due to the emergence of complex aetiology, degree of differentiation and pathological variations, various enhancement modalities for HCC have been perceived [4–6]. Up to 40% of HCCs showed no enhancement on AP, and 40–60% of small HCC lesions showed no enhancement or a capsular appearance on PVP [7]. On the other hand, benign liver lesions (BLLs) with abundant blood supply, including hepatic hemangioma (HH), hepatic angiomyolipoma (HAML), focal nodular hyperplasia (FNH), hepatic adenoma (HA), and hepatic abscess (HB), may sometimes exhibit enhancement patterns similar to HCC. For example, FNH might lack the typical central scar or be associated with fatty liver, giving a misleading image [8, 9]. In another case, epithelioid HAML with few or no lipocytes was also compatible with HCC [10]. These dilemmas increase the risk of misdiagnosis.

Magnetic resonance imaging (MRI), with the advantages of multiple parameters and sequences, performed a significantly higher sensitivity in the diagnosis of HCC, especially for small lesions (≤ 2 cm) [11, 12]. Gatos et al. [13] distinguished well between benign lesions, HCC, and metastatic focal liver lesions with high sensitivity and specificity using T2-weighted imaging (T2WI). However, a pooled study found that LI-RADS for computed tomography or MRI to detect HCC had a sensitivity of 67% and a specificity of 93% [14]. Moreover, imaging reports were highly subjective, reducing the accuracy of diagnosis. Therefore, there is an urgent need for clinicians to seek improved diagnostic methods.

Radiomics, a burgeoning field of medical artificial intelligence research, consists of the high-throughput extraction of radiomics features from imaging data, including details of tumor phenotype and tumor microenvironment [15, 16]. In other words, radiomics is distinct from and complementary to other relevant databases [13]. According to previous research, successful applications of radiomics have been reported in differentiating a specific focal hepatic lesion from HCC [17–19]. To our knowledge, few studies have concentrated on the discrimination between HCC and hypervascular BLLs. Therefore, we attempted to establish and estimate a T2WI-based radiomics model for the classification of HCC and hypervascular BLLs. We believe that this will further advance the clinical application of personalised diagnosis and precision therapy.

2 Materials and methods

2.1 Patients

Our institutional review committee approved the retrospective study and waived the right to informed consent (registration number: No. 2021-K-Y-085-01).

Patients were retrieved from the pathology database of our center from January 2010 to September 2023, and samples were obtained from surgical resection or needle biopsy.

In detail, a total of 220 patients (including 140 HCC, 60 HH, 8 FNH, 5 HAML, and 7 HB) with contrast-enhanced MR images were included in this study. Inclusion criteria were as follows: (1) Patients with contrast-enhanced MR scans performed within 2 weeks before surgery. (2) Patients with hyperenhancing lesions, defined as the main body of the lesion being uniformly or partially (> 50%) more intense than the surrounding liver parenchyma in AP, excluding the lesion already showing hyperintensity on unenhanced T1WI [20]. (3) Patients with lesions larger than 1 cm in maximum diameter [21, 22]. Exclusion criteria were as follows: (1) HCC patients who had received radiotherapy or chemotherapy before surgery; (2) Patients with incomplete clinical information. (3) Patients with images of unsatisfactory quality that compromised the analysis. Of the 144 patients enrolled from January 2010 to October 2021, a random allocation ratio of 7:3 was employed to assign them to either the training cohort (n = 100) or the test cohort (n = 44). The remaining patients (n = 76) enrolled from November 2021 to September 2023 formed the validation cohort.

Age, gender, serum alpha-fetoprotein (AFP) level (> 400 ng/ml, ≤ 400 ng/ml), maximum diameter, and lesion location (left lobe, right lobe including caudate lobe) were obtained from clinical records.

2.2 Imaging data parameters

Before surgery or biopsy, a Siemens Verio 3.0T superconducting MR scanner equipped with a 12-channel abdominal coil was used for all patients. To minimize the impact of respiratory motion artifacts on texture features, all patients underwent imaging with the same respiratory gating technique and were instructed to hold their breath during acquisition. Gadodiamide (Omniscan, Berlin, Germany) was injected at a rate of 2.5–3.0 ml/s, with a dosage of 0.1 mmol/kg body weight, using a power injector. The main axial image parameters included a field of view (FOV) 380 mm, a slice thickness of 6 mm, a 1.2 mm interval. The main scanning sequences including a T2WI fat suppression sequence (repetition time (TR) 1200 ms, echo time (TE) 81 ms), a T1WI fat suppression sequence (TR 3.25 ms, TE 1.13 ms), a diffusion-weighted imaging (DWI) sequence (TR 7400 ms, TE 73 ms) with b values of 0 and 800 s/mm^2, and dynamic enhancement scans such as AP, PVP, lequilibrium phase (LP) and DP performed after contrast injection at 15–25 s, 60 s, 180 s and 5 min, respectively.

2.3 Traditional imaging analysis

Two abdominal radiologists who were unaware of the pathology results (a junior radiologist with about 3 years' diagnostic experience and a senior radiologist with more than 10 years' diagnostic experience) reviewed MR images (T2WI, T1WI, DWI, apparent diffusion coefficient (ADC), AP, PVP, LP and DP) along with the clinical data, and made a diagnosis for each patient. For patients with cirrhosis and other high-risk conditions, they used the Liver Imaging Reporting & Data System (LI-RADS) to assess lesions.

2.4 Image preprocessing

All images were resampled to an isotropic voxel size of 1 × 1 × 3 mm^3 using B-spline interpolation, followed by Z-score normalization based on the liver parenchyma region of interest (ROI).

2.5 MR radiomics analysis

T2WI data were used for the radiomics analysis. Junior radiologists performed semi-automatic lesion delineation using ITK-SNAP software (Version 3.8.0), with manual adjustment of ROIs in the axial, coronal, sagittal positions when necessary to exclude surrounding compressed parenchyma or perilesional edema while avoiding areas of necrosis (Fig. 1). All modified ROIs underwent subsequent verification by senior radiologists to ensure anatomical precision. Eventually, a total of 396 radiomics features were extracted from each ROI using A.K. software (Artificial Intelligent Kit, GE Medical Healthcare, Milwaukee, Wisconsin), including 19 morphological features, 42 histogram features, 180 Gy-level run-length matrix features, 11 Gy-level size-zone matrix features, and 144 Gy-level co-occurrence matrix features. Before commencing the screening, 20 lesions were randomly selected for ROI delineation by both junior and senior physicians, and features were extracted for consistency checks. The radiomics features with an intra-class correlation coefficient (ICC) > 0.75 were retained for further analysis.

2.6 Construction and evaluation of the radiomics model

This study employed a multi-stage feature selection strategy to construct a radiomics prediction model. Initially, dimensionality reduction was performed using the Max-Relevance and Min-Redundancy (mRMR) algorithm: the predictive value of each radiomics feature was evaluated by calculating its mutual information with clinical outcomes (Max-Relevance criterion), while redundancy among features was reduced by computing Pearson correlation coefficients between them (Min-Redundancy criterion). This step retained the top 50 most discriminative features with low mutual information from the originally extracted radiomics features.

To further enhance the specificity of feature selection, Least Absolute Shrinkage and Selection Operator (LASSO) regression was applied for secondary screening. The optimal regularization parameter λ was determined through 5-fold cross-validation,

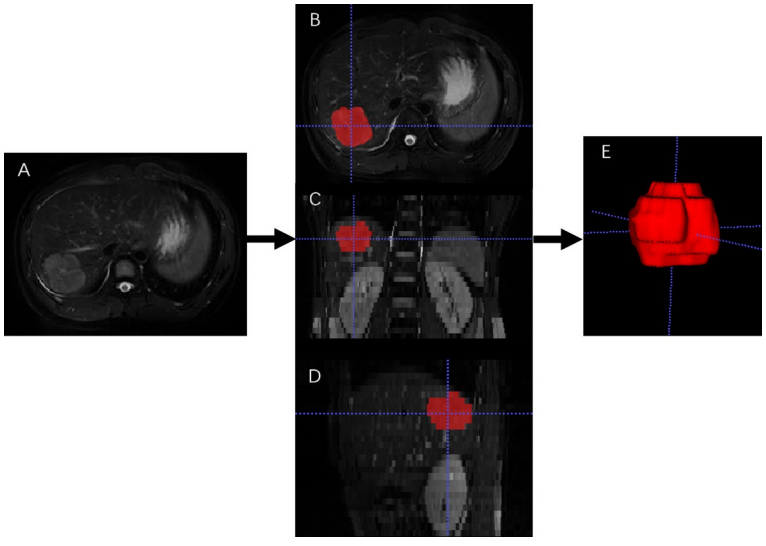


Fig. 1 An example of segmentation in a 59-year-old male patient with HCC. **A** The patient's T2W images in DICOM format were opened in ITK-snap. **B–D** The window width and window level were adjusted to perform semi-automatic segmentation of the lesion on transverse plane, and the lesion margin was adjusted in the transverse, coronal and sagittal planes. **E** A three-dimensional ROI was formed. However, sometimes the lesion margin was not correctly identified in all layers, in which case the lesion was manually outlined slice by slice

ultimately identifying features with non-zero coefficients. These selected features were then incorporated into a weighted linear combination to calculate the Rad-score: each feature was first standardized using Z-score normalization, followed by weighted summation according to the coefficients obtained from LASSO regression.

2.7 Construction and evaluation of the combined model

Clinical characteristics exhibit incremental discriminative value [23, 24]. Consequently, a combined model was developed through logistic regression, integrating the Rad-score with clinical factors that demonstrated a significant distinction between HCC and BLLs. To ensure model robustness and mitigate overfitting, 5-fold cross-validation was employed to identify the optimal model with the highest average area under the curve (AUC) in the test group. Receiver operating characteristic (ROC) curves were generated to assess the diagnostic performance of the prediction models in term of AUCs. Decision curve analysis (DCA) was applied to validate the feasibility of the combined model, taking into account clinical factors. Nomogram was constructed to visually represent the combined model's performance, aiding in the pre-surgical differentiation of HCC from hypervascular BLLs. Taking pathology as the gold standard, the performance of the optimal model would be compared with the results of two radiologists. Finally, patients misdiagnosed by radiologists were divided into a group and the machine learning models described above were used to distinguish HCC from BLLs in the group.

2.8 Statistical analysis

Clinical risk factors were analyzed using SPSS 23.0 (IBM Corp, Chicago, USA). Quantitative data following to normal distribution were presented by a T-test and expressed as $X \pm S$, or tested by a Wilcoxon test and expressed as median (upper and lower quartile). Categorical variables such as sex, lesion location, and AFP level were analyzed by the χ^2 test or F test as appropriate.

The R software (version 4.0.3, https://www.r-project.org/) was implemented for feat ure dimension reduction as well as prediction model construction and evaluation. The AUC and the nomogram were generated using the "pROC" package and the "rms" package, respectively. A Delong test was performed to compare AUC values among different prediction models with statistical significance indicated by a P-value < 0.05.

3 Results

3.1 Clinical factors of the patients

220 patients were divided into HCC (140 cases, 63.68 ± 9.93 years, range 37–87 years) and BLLs (80 cases, 51.32 ± 12.24 years, range 15–76 years). The clinical factors in the univariate logistic regression analysis were shown in Table 1.

Statistical differences in age, gender, lesion location, AFP level, liver cirrhosis and hepatitis history between HCC and BLLs ($P < 0.05$) were found in the training group, whereas no statistical difference was found in maximum diameter ($P = 0.484$). The multiple logistic regression analysis showed that only age ($P = 0.003$) and liver cirrhosis ($P = 0.047$) remained as independent predictors in the clinical model (Table 2).

Table 1 Clinical factors in univariate logistic regression analysis

Characteristics	Training group (*n* = 100)			Test group (*n* = 44)			Validation group (*n* = 76)		
	BHLs (*n* = 30)	HCC (*n* = 70)	*P*	BHLs (*n* = 13)	HCC (*n* = 31)	*P*	BHLs (*n* = 37)	HCC (*n* = 39)	*P*
Age	50.2 ± 10.9	63.9 ± 9.2	< 0.001*	48.0 (36.0, 61.0)	62.0 (51.0, 69.0)	0.020*	52.9 ± 11.8	65.9 ± 10.7	< 0.001*
Gender			0.001*			< 0.001*			< 0.001*
Male	15 (50.0%)	57 (81.4%)		4 (30.8%)	30 (96.8%)		11 (29.7%)	28 (71.8%)	
Female	15 (50.0%)	13 (18.6%)		9 (69.2%)	1 (3.2%)		26 (70.3%)	11 (28.2%)	
AFP level			0.027*			0.433			0.012*
> 400 ng/ml	0 (0.0%)	13 (18.6%)		0 (0.0%)	4 (12.9%)		0 (0.0%)	7 (17.9%)	
≤ 400ng/ml	30 (100.0%)	57 (81.4%)		13 (100.0%)	27 (87.1%)		37 (100.0%)	32 (82.1%)	
Liver cirrhosis			< 0.001*			< 0.001*			< 0.001*
Presence	1 (3.1%)	31 (96.9%)		0 (0)	17 (100.0%)		0 (0)	20 (100.0%)	
Absence	29 (42.6%)	39 (57.4%)		13(48.1%)	14 (51.9%)		37 (66.1%)	19 (33.9%)	
Hepatitis history			< 0.001*			< 0.001*			< 0.001*
Presence	0 (0)	39 (100.0%)		1 (4.3%)	22 (95.7%)		1 (3.6%)	27 (96.4%)	
Absence	30 (49.2%)	31 (50.8%)		12 (57.1%)	9 (42.9%)		36 (75.0%)	12 (25.0%)	
Maximum diameter(mm)	52.5 (39.8, 69.0)	43.5 (33.0, 62.7)	0.484	56.0 (36.0, 85.5)	39.0 (32.0, 62.0)	0.203	49.1 ± 15.0	34.8 (20.9, 51.1)	0.051
Lesion location			0.005*			0.043*			0.102
Left	17 (56.7%)	19 (27.1%)		8 (61.5%)	9 (29.0%)		15 (40.5%)	9 (23.1%)	
Right (including caudate lobe)	13 (43.3%)	51 (72.9%)		5 (38.5%)	22 (71.0%)		22 (59.5%)	30 (76.9%)	

3.2 Traditional imaging analysis

Based on the above clinical information and full-phase MRI images (including T1WI, T2WI, DWI, ADC, enhanced sequence), two radiologists misdiagnosed 42 cases (junior doctors misdiagnosed 30 cases, and senior doctors misdiagnosed 23 cases). The sensitivities of the junior radiologists in the training, test, and validation sets were 0.857, 0.935, and 0.718, respectively; specificities were 0.867, 1, and 0.865, respectively; and accuracies were 0.860, 0.955, and 0.789, respectively. For senior radiologists, sensitivities were 0.914, 1, and 0.821, respectively. Specificities were 0.867, 0.923, and 0.865, respectively. Accuracies were 0.900, 0.977, and 0.842, respectively. Table 3 listed the clinical and imaging characteristics of misdiagnosed cases.

3.3 Radiomics feature extraction

Four optimal radiomics features were filtrated to build the radiomics model. Details of these features could be found in Table 4; Fig. 2. The Rad-score of each patient was calculated using the following formula.

Rad-score = 0.2052 × Cluster Shade_angle135_offset4–1.5800 × High Intensity Large Area Emphasis (HILAE) − 1.3509 × HaralickCorrelation_All Direction_offset1_SD − 0.8757 × Long Run High Grey Level Emphasis (LRHGLE)_angle0_offset4 + 0.9363.

3.4 Evaluation of prediction models

By integrating the Rad-score with clinically independent predictors (age, liver cirrhosis), a fusion model was built. Figure 3 illustrates the ROC results for three prediction models in the training, test and validation groups. The AUC values for the fusion model reached 0.984, 0.948, 0.955 in the training, test and validation groups, respectively, while the AUCs for the clinical and radiomics models were 0.893, 0.891, 0.910 and 0.879, 0.839, 0.807, respectively. Table 4 shows that there was a statistically significant difference between the fusion model and the clinical model ($P = 0.003$), between the fusion model and the radiomics model ($P = 0.031$) in the training set, and between the fusion model and the radiomics model in the validation set ($P = 0.001$). Calibration curves for the three models were presented in Fig. 4. According to DCA, the net benefit of the fusion model was the highest among the three models (Fig. 5). A diagnostic nomogram was provided in Fig. 6. Figure 7 offered a comparison of the diagnostic performance of the fusion model, junior radiologist, and senior radiologist on the training, test, and validation sets. Finally, the patients misdiagnosed by the radiologists were divided into a group and the above machine learning models were used to distinguish HCC from BLLs (Figs. 3, 4 and 5; Table 5). The results revealed that the AUC of the clinical model, radiomics model, and fusion model were 0.841, 0.737, and 0.905, respectively.

4 Discussion

In this study, we developed three predictive models to differentiate HCC from blood-rich BLLs. The fusion model, visualized as a nomogram, combined two clinical factors and the Rad-score, which integrated four radiomics features extracted from T2WI. While the addition of clinical data did not significantly improve the performance of the radiomics model, results from the validation group suggested that it might have a positive impact. Among the three machine models, the T2WI-based fusion model emerged as the most effective predictive model. Its diagnostic performance was comparable to

Table 2 Clinical factors in multiple logistic regression analysis

Characteristics	Training cohort		
	B	*P*	OR (95%CI)
Age	0.245	0.003	1.277 (1.088, 1.499)
Liver cirrhosis	− 2.815	0.047	0.060 (0.004, 0.965)

"B" stands for the regression coefficient calculated by SPSS

Table 3 The clinical and imaging characteristics of misdiagnosed cases

Characteristics	BLLs	HCC
Number of misdiagnoses	13	29
Number of junior radiologist misdiagnoses	9	21
Number of senior radiologist misdiagnoses	10	13
Age (years)	15–70	48–79
Liver cirrhosis	1	12
Hepatitis history	2	20
AFP ≥ 400 (ng/ml)	0	3
Enhancement pattern		
Non-uniform enhancement[a]	3	4
Rapid wash-in and rapid wash-out enhancement	6	13
Gradual enhancement	4	10
Capsular enhancement	0	1
Progressive enhancement	0	1
Diffusion restriction		
Presence	10	24
Absence	3	5
Presence of a capsule	1	11
Presence of scarring or septations	4	7

[a]Due to components such as liquefaction necrosis or cystic changes, the radiologist initially misdiagnosed the result as no enhancement, when in fact the case showed enhancement (reviewed by a senior radiologist).

Table 4 Four radiomics features and P value in the training, test and validation cohorts

Radiomics features	Training cohort			Test cohort			Validation cohort		
	BLLs	HCC	*P*	BLLs	HCC	*P*	BLLs	HCC	*P*
Cluster Shade_angle135_offset4	− 0.341 (− 1.498, 0.095)	0.239 (−0.068, 0.557)	<0.001	0.137 (− 0.299, 0.728)	0.064 (− 0.480, 0.481)	0.908	− 0.198 (− 0.859, 0.564)	0.330 (0.084, 0.578)	0.007
High intensity large area emphasis	0.764 (−0.369, 2.049)	− 0.419 (− 0.737, − 0.070)	<0.001	0.319 ± 0.680	− 0.489 ± 0.372	0.001	− 0.177 (− 0.362, 0.254)	− 0.459 (− 0.507, − 0.182)	< 0.001
HaralickCorrelation_AllDirection_offset1_SD	0.052 (−0.302, 0.699)	− 0.305 (− 0.359, − 0.198)	<0.001	0.168 (− 0.347, 1.29)	− 0.317 (− 0.358, − 0.139)	0.066	0.225 (− 0.580, 1.333)	− 0.486 ± 0.318	< 0.001
Long run high grey level emphasis_angle0_offset4	0.844 ± 1.095	− 0.362 ± 0.710	<0.001	0.455 ± 0.742	− 0.112 ± 0.817	0.037	0.336 (− 0.422, 1.047)	− 0.269 (− 1.201, 0.331)	0.004

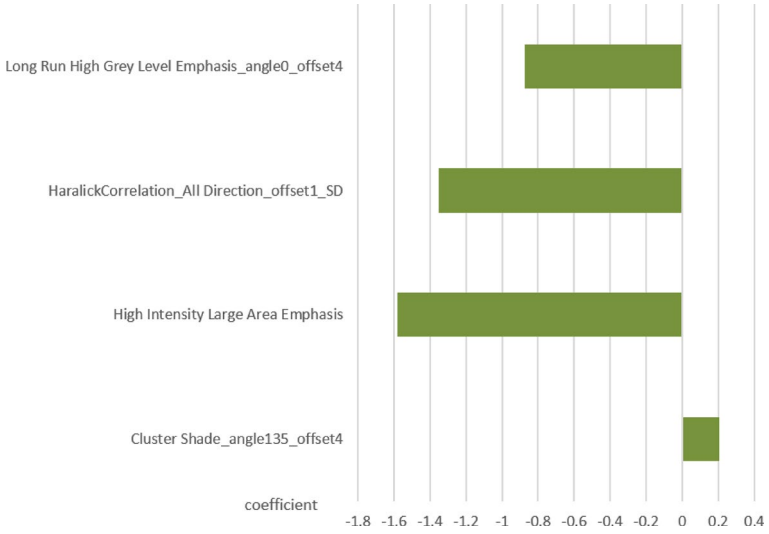


Fig. 2 The contribution of the selected parameters with their regression coefficients. ClusterShade_angle135_offset4 is positively proportional to HCC, while the other three are inversely proportional to HCC. The absolute value of the coefficient of High Intensity Large Area Emphasis is the largest, indicating that it has the greatest effect on Rad-score

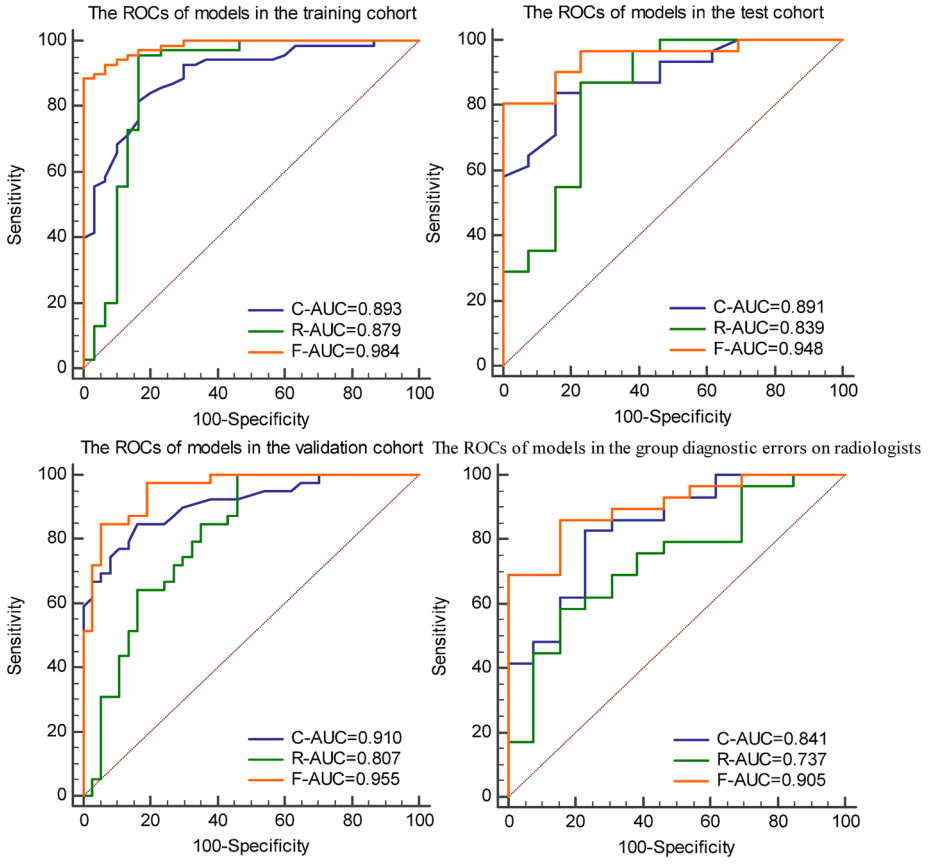


Fig. 3 The ROC curve plots of the three prediction models in the training, test and validation cohorts and in the group of radiologist diagnostic errors. C represents the clinical model, R means the radiomics model and F means the fusion model

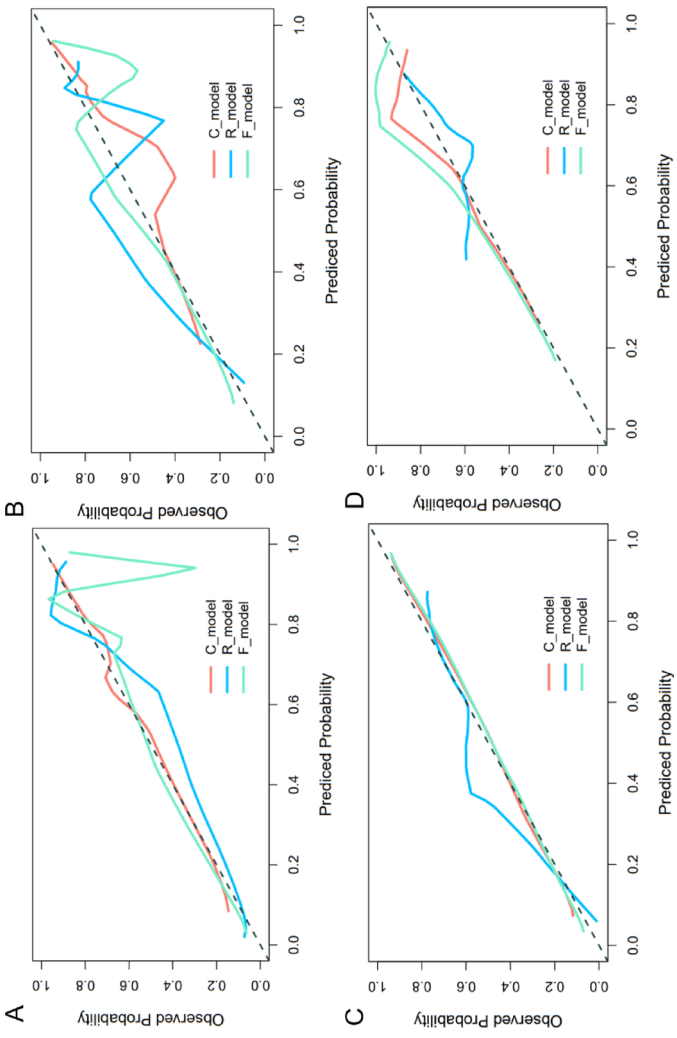


Fig. 4 Calibration curves of the nomogram developed in the training group **A**, the test group **B**, the validation group **C** and the group diagnostic errors in radiologists **D**. Diagonal dotted line = a perfect estimation by an ideal model, where the estimated outcome perfectly matches corresponds to the actual outcome. Solid line = performance of the nomogram, where a closer alignment with the diagonal dotted line represents a better estimation. In this result, the fusion model has a good calibration ability. C = clinical model, R = radiomics model, F = fusion model

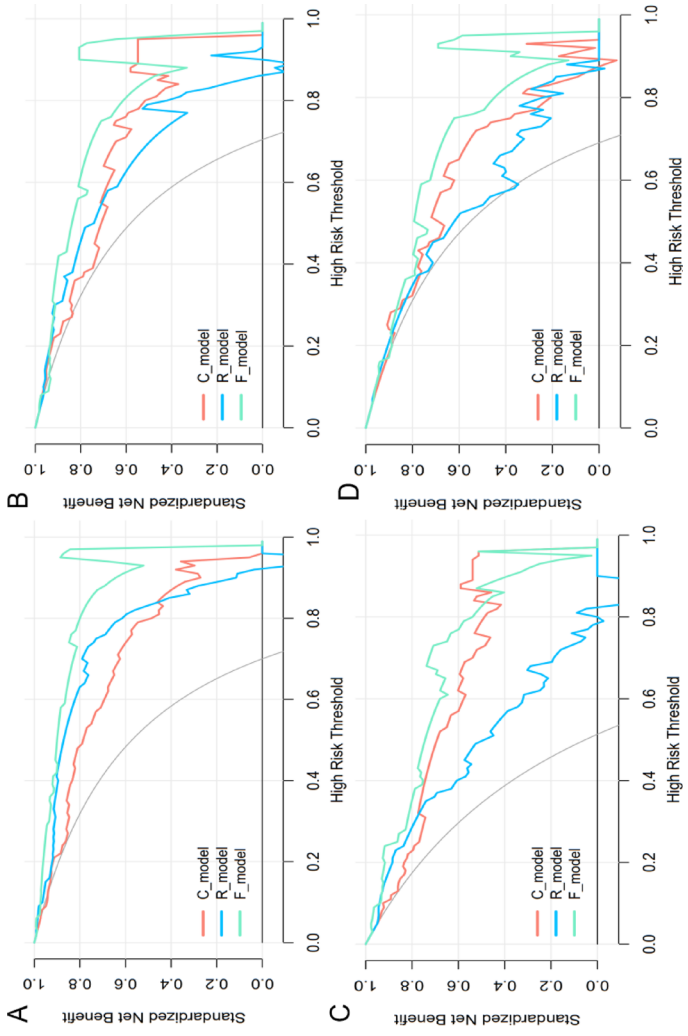


Fig. 5 Decision curves revealed the performance of three models (**A**: training group, **B**: test group, **C**: validation group, **D**: the group diagnostic errors on radiologists). From these results, the fusion model had the highest net benefit compared to the other two models. C = clinical model, R = radiomics model, F = fusion model

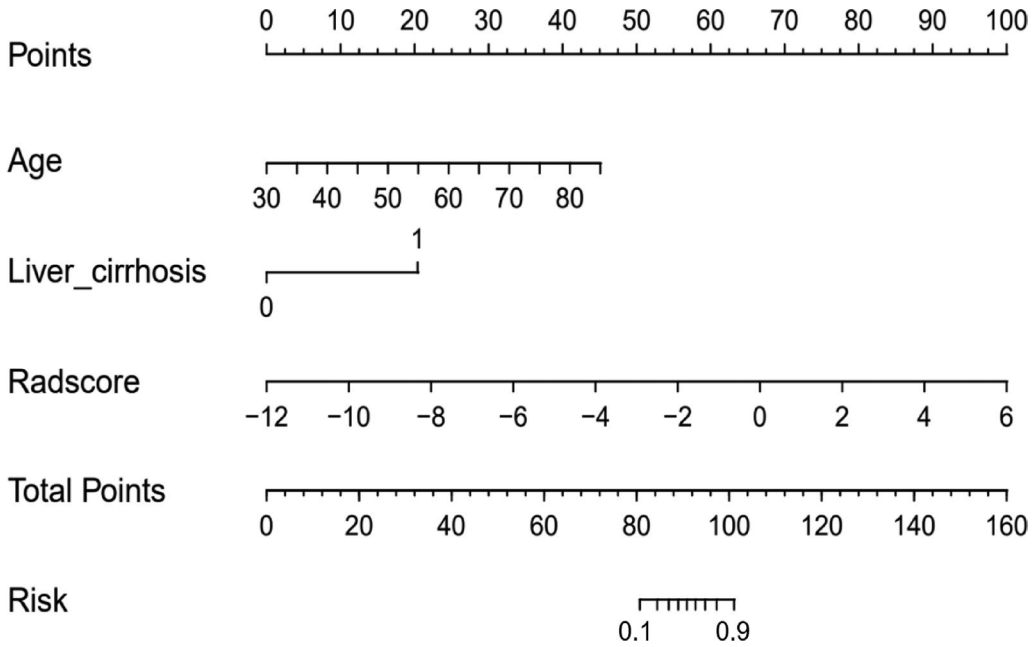


Fig. 6 The diagnostic nomogram included the Rad-score and 2 clinical features of the fusion model. Draw a line perpendicular to the axis of these points to determine how many points, the probability of HCC the patient receives for his or her Rad-score. Repeat the process for each variable. Add up the points achieved for each risk factor. Locate the final sum on the Total Point axis. Draw a line to find the patient's probability of developing HCC

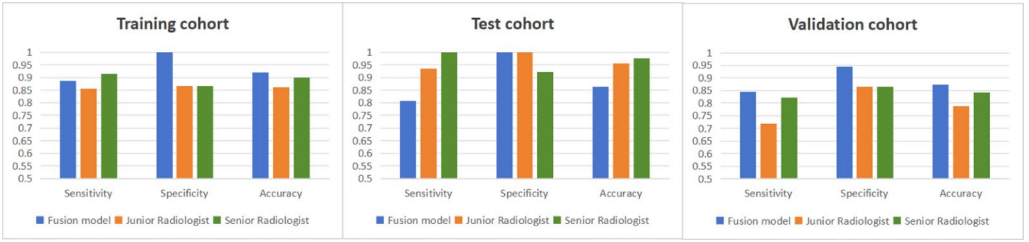


Fig. 7 Comparison of the diagnostic performance of the fusion model, the junior radiologist, and the senior radiologist on the training, test, and validation sets

Table 5 Comparison of diagnostic ability among three models for predicting HCC and BLLs

Model	AUC (95%CI)	Specificity	Sensitivity	Accuracy	*P*
Training set					
Clinical model	0.893 (95%CI 0.815–0.946)	0.833	0.814	0.820	0.003*
Radiomics signature	0.879 (95%CI 0.798–0.935)	0.833	0.957	0.920	0.031*
Fusion model	0.984 (95%CI 0.936–0.999)	1.000	0.886	0.920	Ref.
Test set					
Clinical model	0.891 (95%CI 0.760–0.965)	0.846	0.839	0.841	0.171
Radiomics signature	0.839 (95%CI 0.697–0.932)	0.769	0.871	0.841	0.126
Fusion model	0.948 (95%CI 0.836–0.992)	1.000	0.806	0.864	Ref.
Validation set					
Clinical model	0.910 (95%CI 0.821–0.963)	0.838	0.846	0.820	0.050
Radiomics signature	0.807 (95%CI 0.701–0.889)	0.541	1.000	0.776	0.001*
Fusion model	0.955 (95%CI 0.881–0.989)	0.946	0.846	0.872	Ref.
The group diagnostic errors on radiologists					
Clinical model	0.841 (95%CI 0.695–0.935)	0.769	0.828	0.810	0.212
Radiomics signature	0.737 (95%CI 0.579–0.861)	0.846	0.586	0.667	0.058
Fusion model	0.905 (95%CI 0.773–0.973)	0.846	0.862	0.857	Ref.

traditional imaging diagnoses that include more available information, such as plain, DWI, enhanced sequences, and clinical data, but it showed higher diagnostic specificity. Furthermore, the T2WI-based machine model in this study was able to classify lesions that were misdiagnosed by conventional imaging with strong diagnostic performance.

Previously, many studies have proposed several methods to discriminate HCC from various BLLs, including fp-AMLs, FNH, HH and others [17, 25]. Most of these studies employed CE- computed tomography (CT) or MRI, but their optimal model had limitations in terms of AUC, accuracy, sensitivity, and specificity. It needs to be improved. In addition, the liver lesions are usually found on CE-CT and further diagnosed on CE-MRI in clinical practice, thus we need to consider the potential harm of reusing contrast agents, especially in patients with renal insufficiency or allergy to these agents [26, 27]. T2WI, with high tissue resolution, was more sensitive in detecting cystic degeneration and necrosis. Wu et al. [28] suggested that the AUC of T2WI surpassed that of other imaging sequences in the classification of HCC and HH. Therefore, our study established "A T2WI-based radiomics nomogram" for identifying HCC and hypervascular BLLs without the need for contrast agents, and it yielded excellent diagnostic results with the AUC of 0.984, 0.948 and 0.955 in the training, test and validation sets, respectively. The clinical model also played a crucial role, achieving a relatively high AUC of 0.891 and 0.910 in the test and validation sets, respectively, whereas the AUCs of the radiomics model were 0.839 and 0.807, respectively. These results suggest that radiomics parameters could complement clinical practice to improve accurate diagnosis.

Clinical characteristics are important for the identification and diagnosis of these diseases. BLLs were predominantly found in young women without liver cirrhosis, with a mean age of 51.32 ± 12.24 years. In contrast, HCC occurred predominantly in older men with liver cirrhosis, with a mean age of 63.68 ± 9.93 years. In our result, while our univariate regression analysis indicated that the AFP level had statistical significance, this significance did not hold in multivariate regression analysis. Therefore, we decided not to include it in our clinical model. Possible reasons for this decision are as follows: (1) Serum AFP levels remain normal in 15–30% of advanced HCC [29, 30]. (2) The lesion may be too small to produce sufficient AFP, which is another possibility. (3) There may

be an influence of antiviral therapy on AFP positivity, as suggested by Qian et al. [31]. In conclusion, our clinical model, including age, liver cirrhosis, aligns with clinical differences, validating the rationality of our model.

According to the nomogram (Fig. 3D), radiomics features that improve the differentiation efficiency were also crucial references. In this study, four radiomics features belonging to texture features were extracted: Cluster Shade_angle135_offset4, Haralick Correlation_All Direction_offset1_SD, HILAE, LRHGLE_angle0_offset4. Cluster Shade, a component of the gray-level co-occurrence matrix (GLCM), groups similar objects and has been successful in discriminating between high- and low-volume cervical tumours or between benign and malignant sub-centimetre pulmonary ground-glass nodules [32, 33]. In our study, Cluster Shade_angle135_ offset4 could be a discriminator between benign and malignant liver lesions. Haralick correlation in GLCM measures the similarity of grey levels in the row or column direction of images, indicating a linear relationship of grey levels [34]. Higher values indicate a stronger correlation. Akshima and colleagues [35] confirmed the importance of Haralick correlation in distinguishing prostate cancer from normal prostate. In our study, Haralick Correlation_All Direction_offset1_SD played an important role in the identification of BLL and HCC, demonstrating strong grey correlation in all directions and a more uniform grey distribution for BLLs. HILAE is sensitive to flat areas of high intensity voxels, highlighting local heterogeneity information [36]. Coroller et al. [37] found that lymph nodes with high HILAE values showed better pathological responses. In our study, BLLs exhibited higher HILAE values, indicating finer texture and lower heterogeneity. LRHGLE measures the joint distribution of long runs and high gray level values, primarily influenced by image roughness. It has been reported that T2 imaging could discriminate normal liver tissue from HCC based on its high-intensity features [38]. In our study, this parameter effectively discriminated BLLs from HCC. Compared with benign lesions, HCC exhibited finer texture values with lower grey level on T2WI in the angle0 direction. Overall, HCC images displayed finer, more uneven textures in both two and three dimensions, as well as greater heterogeneity.

The radiomics nomogram improved the identification ability to traditional imaging based on radiologist experience with the drawback of subjective judgment. Two specialized abdominal imaging radiologists were invited to make diagnosis on enhanced MR images. Of the 140 HCC cases, 44 cases were misdiagnosed, including 12 cases with cirrhosis, 20 cases with a history of hepatitis B, and only 3 cases with AFP ≥ 400ng/ml. From imaging findings, 5 patients exhibited lesions without diffusion restriction, potentially attributable to low cell density. 4 cases demonstrated heterogeneous enhancement with relatively low enhancement intensity, potentially attributable to: scattered necrotic/cystic components within the lesions, or fibrous septa formation, or preserved sinusoidal structures in well-differentiated tumor regions. Ten cases displayed progressive enhancement patterns, while 1 case showed persistent enhancement throughout all phases, likely reflecting: abundant fibrotic stroma within the tumor (histologically confirmed collagen proportion > 40% in the persistent enhancement case), or delayed contrast retention in extracellular matrix-rich areas. Of the 80 cases of BLLs, 19 cases were misdiagnosed, including 1 case with cirrhosis and 2 cases with a history of hepatitis B. Imaging showed that 6 cases exhibited the rapid wash-in and wash-out enhancement pattern, 3 cases showed heterogeneous enhancement, 10 cases had restricted diffusion.

Therefore, atypical clinical history and imaging features may influence the diagnosis of radiologists. The junior radiologist demonstrated specificity, sensitivity, and accuracy in the ranges of 0.865–1, 0.718–0.935, and 0.789–0.955, respectively, while the senior radiologist achieved scores of 0.865–0.923, 0.821–1.000, and 0.842–0.977, respectively. The difference between the two radiologists was not statistically significant. The diagnostic performance of the T2WI-based fusion model was comparable to that of radiologists, even when additional information from enhanced sequences was available. Notably, in the validation set, the sensitivity of traditional imaging diagnosis was lower than that of radiomics models, with sensitivity, specificity, and accuracy being lower than those of the fusion model. These results suggested that compared with traditional imaging diagnostics, T2WI-based radiomics features combined with clinical features might have a lower rate of missed diagnosis and minimal subjective errors, making it a valuable valuable auxiliary tool in clinical practice. Additionally, we used a machine learning model based on T2WI to differentiate between HCC and BLLs patients who were misdiagnosed by radiologists. The three machine models demonstrated strong diagnostic efficacy, among which the combined prediction model has the highest diagnostic ability, with better sensitivity, specificity and accuracy. It further underscored the advantages of T2WI-based machine learning models for conventional imaging diagnosis.

However, this study also had the following limitations: (1) The sample size remains limited, and the distribution of benign blood-rich lesions with different pathological characteristics was uneven. The lack of pathologically confirmed hepatic adenomas in our institutional registry precluded meaningful analysis of this entity, despite its recognized imaging similarities to HCC. Increasing the total number of cases is warranted in future studies. (2) External validation is needed to further verify the accuracy and clinical applicability of the model. (3) Sample selection bias was unavoidable in this retrospective study, and prospective studies should be conducted in the future. (4) According to LIRADS (https://www.acr.org/-/media/ACR/Files/Clinical-Resources/LIRADS/Chapter-12-Technique.pdf), a higher injection rate (2.5–3.0 ml/s) of MR contrast agent does not guarantee higher contrast in MRI studies. As this is a retrospective study, we will consider reducing the contrast injection rate in the future to improve the imaging effect of liver lesions. (5) The time splitting of the datasets in this study may affect the generalization of the model, although the images and radiomics features have been preprocessed.

5 Conclusion

In conclusion, our novel T2WI-based radiomics nomogram effectively distinguishes HCC from blood-rich BLLs, offering valuable support for early clinical diagnosis and precise treatment. It is expected to reduce the reliance on MRI contrast agents and overcome the limitations of traditional imaging diagnosis.

Acknowledgements

A previous version of this manuscript was published as a preprint (https://doi.org/10.21203/rs.3.rs-1819068/v1).

Author contributions

Chen Yang and Boburjon Turgunov contributed to the work equally. Zhenhua Zhao mainly contributed to the study conception and design. Material preparation and data collection were performed by Chen Yang, Boburjon Turgunov and Fandong Zhu, and data analysis was performed by Yang Xia and Jianfeng Zhou. The first draft of the manuscript was written by Chen Yang and Boburjon Turgunov, then all authors commented on previous versions of the manuscript.

Funding

This work was supported by Key Laboratory of Functional Molecular Imaging of Tumor and Interventional Diagnosis and Treatment of Shaoxing City (No. 2020ZDSYSO1), Medicine and Health Science and Technology Plan Project in Zhejiang

Province of China (No. 2022KY1321), Shaoxing Municipal Project of Health Science, Technology Plan (No. 2020A13027), and General Plan for medical and health research in Zhejiang (No. 2021KY371).

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

Declarations

Ethics approval and consent to participate
The protocol was approved by the Medical Research Ethics Committee and the Institutional Review Board of Shaoxing People's Hospital (No. 2021-K-Y-085-01) in accordance with the Declaration of Helsinki (as revised in 2013). Due to the retrospective nature of the study, the requirement for informed consent was waived.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 2 February 2025 / Accepted: 29 May 2025
Published online: 01 July 2025

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49. https://doi.org/10.3322/caac.21660.
2. Llovet JM, Kelley RK, Villanueva A, et al. Hepatocellular carcinoma. Nat Rev Dis Primers. 2021;7(1):6. https://doi.org/10.1038/s41572-020-00240-3.
3. Osho A, Rich NE, Singal AG. Role of imaging in management of hepatocellular carcinoma: surveillance, diagnosis, and treatment response. Hepatoma Res. 2020;6:55. https://doi.org/10.20517/2394-5079.2020.42.
4. Wang SY, Yin L, Wang C, Ma MP. Atypical magnetic resonance imaging features and differential diagnosis of hepatocellular carcinoma. J Int Med Res. 2020;48(10):300060520943415. https://doi.org/10.1177/0300060520943415.
5. Zech CJ, Ba-Ssalamah A, Berg T, et al. Consensus report from the 8th international forum for liver magnetic resonance imaging. Eur Radiol. 2020;30(1):370–82. https://doi.org/10.1007/s00330-019-06369-4.
6. Kitao A, Matsui O, Yoneda N, et al. Differentiation between hepatocellular carcinoma showing hyperintensity on the hepatobiliary phase of Gadoxetic Acid-Enhanced MRI and focal nodular hyperplasia by CT and MRI. AJR Am J Roentgenol. 2018;211(2):347–57. https://doi.org/10.2214/AJR.17.19341.
7. Jones RH, Taylor AJ, Rostambeigi N, Spilseth B. Small hepatocellular carcinomas displayed as a ring enhancing mass on arterial phase MRI in the chronically diseased liver. Clin Radiol. 2017;72(11):995. https://doi.org/10.1016/j.crad.2017.07.001.
8. Lebert P, Adens-Fauquembergue M, Azahaf M, et al. MRI for characterization of benign hepatocellular tumors on hepatobiliary phase: the added value of in-phase imaging and lesion-to-liver visual signal intensity ratio. Eur Radiol. 2019;29(11):5742–51. https://doi.org/10.1007/s00330-019-06210-y.
9. Khanna M, Ramanathan S, Fasih N, et al. Current updates on the molecular genetics and magnetic resonance imaging of focal nodular hyperplasia and hepatocellular adenoma. Insights Imaging. 2015;6(3):347–62. https://doi.org/10.1007/s13244-015-0399-8.
10. Liang W, Shao J, Liu W, et al. Differentiating hepatic epithelioid Angiomyolipoma from hepatocellular carcinoma and focal nodular hyperplasia via radiomics models. Front Oncol. 2020;10:564307. https://doi.org/10.3389/fonc.2020.564307.
11. Roberts LR, Sirlin CB, Zaiem F, et al. Imaging for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis. Hepatology. 2018;67(1):401–21. https://doi.org/10.1002/hep.29487.
12. Li J, Wang J, Lei L, et al. The diagnostic performance of Gadoxetic acid disodium-enhanced magnetic resonance imaging and contrast-enhanced multi-detector computed tomography in detecting hepatocellular carcinoma: a meta-analysis of eight prospective studies. Eur Radiol. 2019;29(12):6519–28. https://doi.org/10.1007/s00330-019-06294-6.
13. Gatos I, Tsantis S, Karamesini M, et al. Focal liver lesions segmentation and classification in nonenhanced T2-weighted MRI. Med Phys. 2017;44(7):3695–705. https://doi.org/10.1002/mp.12291.
14. Li L, Hu Y, Han J, et al. Clinical application of liver imaging reporting and data system for characterizing liver neoplasms: A Meta-Analysis. Diagnostics (Basel Switzerland). 2021;11(2):323. https://doi.org/10.3390/diagnostics11020323.
15. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the Bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–62. https://doi.org/10.1038/nrclinonc.2017.141.
16. Lewis S, Hectors S, Taouli B. Radiomics of hepatocellular carcinoma. Abdom Radiol (NY). 2021;46(1):111–23. https://doi.org/10.1007/s00261-019-02378-5.
17. Zhao X, Zhou Y, Zhang Y, et al. Radiomics based on Contrast-Enhanced MRI in differentiation between Fat-Poor Angiomyolipoma and hepatocellular carcinoma in noncirrhotic liver: A multicenter analysis. Front Oncol. 2021;11:744756. https://doi.org/10.3389/fonc.2021.744756.
18. Nie P, Wang N, Pang J, et al. CT-Based radiomics nomogram: A potential tool for differentiating hepatocellular adenoma from hepatocellular carcinoma in the noncirrhotic liver. Acad Radiol. 2021;28(6):799–807. https://doi.org/10.1016/j.acra.2020.04.027.
19. Li Z, Mao Y, Huang W, et al. Texture-based classification of different single liver lesion based on SPAIR T2W MRI images. BMC Med Imaging. 2017;17(1):42. https://doi.org/10.1186/s12880-017-0212-x.
20. Kim M, Kang TW, Cha DI, et al. Identification of arterial hyperenhancement in CT and MRI in patients with hepatocellular carcinoma: value of unenhanced images. Korean J Radiol. 2019;20(2):236–45. https://doi.org/10.3348/kjr.2018.0339.

21. Liu X, Jiang H, Chen J, et al. Gadoxetic acid disodium-enhanced magnetic resonance imaging outperformed multidetector computed tomography in diagnosing small hepatocellular carcinoma: A meta-analysis. Liver Transpl. 2017;23(12):1505–18. https://doi.org/10.1002/lt.24867.
22. Song JS, Choi EJ, Hwang SB, et al. LI-RADS v2014 categorization of hepatocellular carcinoma: intraindividual comparison between gadopentetate dimeglumine-enhanced MRI and Gadoxetic acid-enhanced MRI. Eur Radiol. 2019;29(1):401–10. https://doi.org/10.1007/s00330-018-5559-z.
23. Weng SF, Reps J, Kai J, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE. 2017;12(4):e0174944. https://doi.org/10.1371/journal.pone.0174944.
24. Nie P, Yang G, Guo J, et al. A CT-based radiomics nomogram for differentiation of focal nodular hyperplasia from hepatocellular carcinoma in the non-cirrhotic liver. Cancer Imaging. 2020;20(1):20. https://doi.org/10.1186/s40644-020-00297-z.
25. Ding Z, Lin K, Fu J, et al. An MR-based radiomics model for differentiation between hepatocellular carcinoma and focal nodular hyperplasia in non-cirrhotic liver. World J Surg Oncol. 2021;19(1):181. https://doi.org/10.1186/s12957-021-02266-7.
26. Pasquini L, Napolitano A, Visconti E, et al. Gadolinium-Based contrast Agent-Related toxicities. CNS Drugs. 2018;32(3):229–40. https://doi.org/10.1007/s40263-018-0500-1.
27. Pillebout É, Martinez F. Toxicité Rénale des produits de contraste radiologiques [Contrast-induced acute kidney injury]. Nephrol Ther. 2021;17(3):191–8. https://doi.org/10.1016/j.nephro.2021.04.001.
28. Wu J, Liu A, Cui J, Chen A, et al. Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on Precontrast magnetic resonance images. BMC Med Imaging. 2019;19(1):23. https://doi.org/10.1186/s12880-019-0321-9.
29. Han LL, Lv Y, Guo H, et al. Implications of biomarkers in human hepatocellular carcinoma pathogenesis and therapy. World J Gastroenterol. 2014;20(30):10249–61. https://doi.org/10.3748/wjg.v20.i30.10249.
30. Luo P, Wu S, Yu Y, et al. Current status and perspective biomarkers in AFP negative HCC: towards screening for and diagnosing hepatocellular carcinoma at an earlier stage. Pathol Oncol Res. 2020;26(2):599–603. https://doi.org/10.1007/s12253-019-00585-5.
31. Qian X, Liu S, Long H, et al. Reappraisal of the diagnostic value of alpha-fetoprotein for surveillance of HBV-related hepatocellular carcinoma in the era of antiviral therapy. J Viral Hepat. 2021;28(1):20–9. https://doi.org/10.1111/jvh.13388.
32. Wormald BW, Doran SJ, Ind TE, et al. Radiomic features of cervical cancer on T2-and diffusion-weighted MRI: prognostic value in low-volume tumors suitable for trachelectomy. Gynecol Oncol. 2020;156(1):107–14. https://doi.org/10.1016/j.ygyno.2019.10.010.
33. Hu X, Ye W, Li Z, et al. Non-invasive evaluation for benign and malignant subcentimeter pulmonary ground-glass nodules (≤ 1 cm) based on CT texture analysis. Br J Radiol. 2020;93(1114):20190762. https://doi.org/10.1259/bjr.20190762.
34. 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. https://doi.org/10.1016/j.suronc.2018.11.013.
35. Sharma A, Kumar S, Pandey AK, et al. Haralick texture features extracted from Ga-68 PSMA PET/CT to differentiate normal prostate from prostate cancer: a feasibility study. Nucl Med Commun. 2021;42(12):1347–54. https://doi.org/10.1097/MNM.0000000000001469.
36. Yan M, Wang W. A Non-invasive method to diagnose lung adenocarcinoma. Front Oncol. 2020;10:602. https://doi.org/10.3389/fonc.2020.00602.
37. Coroller TP, Agrawal V, Huynh E, et al. Radiomic-Based pathological response prediction from primary tumors and lymph nodes in NSCLC. J Thorac Oncol. 2017;12(3):467–76. https://doi.org/10.1016/j.jtho.2016.11.2226.
38. Brancato V, Garbino N, Salvatore M, Cavaliere C. MRI-Based radiomic features help identify lesions and predict histopathological grade of hepatocellular carcinoma. Diagnostics (Basel). 2022;12(5):1085. https://doi.org/10.3390/diagnostics12051085.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.