



Multi-phase contrast-enhanced magnetic resonance image-based radiomics-combined machine learning reveals microscopic ultra-early hepatocellular carcinoma lesions

Kui Sun¹ · Liting Shi² · Jianfeng Qiu³ · Yuteng Pan³ · Ximing Wang¹ · Haiyan Wang¹

Received: 2 December 2021 / Accepted: 17 February 2022 / Published online: 1 March 2022
© The Author(s) 2022

Abstract

Purpose This study aimed to investigate whether models built from radiomics features based on multiphase contrast-enhanced MRI can identify microscopic pre-hepatocellular carcinoma lesions.

Methods We retrospectively studied 54 small hepatocellular carcinoma (SHCC, diameter < 2 cm) patients and 70 patients with hepatocellular cysts or haemangiomas from September 2018 to June 2021. For the former, two MRI scans were collected within 12 months of each other; the 2nd scan was used to confirm the diagnosis. The volumes of interest (VOIs), including SHCCs and normal liver tissues, were delineated on the 2nd scans, mapped to the 1st scans via image registration, and enrolled into the SHCC and internal-control cohorts, respectively, while those of normal liver tissues from patients with hepatocellular cysts or haemangioma were enrolled in the external-control cohort. We extracted 1132 radiomics features from each VOI and analysed their discriminability between the SHCC and internal-control cohorts for intra-group classification and the SHCC and external-control cohorts for inter-group classification. Five radial basis-function, kernel-based support vector machine (SVM) models (four corresponding single-phase models and one integrated from the four-phase MR images) were established.

Results Among the 124 subjects, the multiphase models yielded better performance on the testing set for intra-group and inter-group classification, with areas under the receiver operating characteristic curves of 0.93 (95% CI, 0.85–1.00) and 0.97 (95% CI, 0.92–1.00), accuracies of 86.67% and 94.12%, sensitivities of 87.50% and 94.12%, and specificities of 85.71% and 94.12%, respectively.

Conclusion The combined multiphase MRI-based radiomics feature model revealed microscopic pre-hepatocellular carcinoma lesions.

Keywords Radiomics · Magnetic resonance imaging · Hepatocellular carcinoma · Machine learning

This article is part of the Topical Collection on Advanced Image Analyses (Radiomics and Artificial Intelligence).

Kui Sun and Liting Shi should be considered co-first authors.

✉ Ximing Wang
wxming369@163.com

✉ Haiyan Wang
whyott@163.com

¹ Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jing Wu Road, No. 324, Jinan 250021, China

² Department of Radiology, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian 271016, China

³ Medical Science and Technology Innovation Center, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan 250000, China

Introduction

According to the World Cancer Report published by the International Agency for Research on Cancer (IARC), hepatocellular carcinoma (HCC), the third leading cause of cancer-related mortality worldwide, accounts for approximately 80% of primary liver cancers. In 2020, there were an estimated 410,038 new cases and 391,152 deaths from HCC in China [1]. Cirrhosis secondary to hepatitis B virus (HBV) infection is a major risk factor for HCC and affects the majority of middle-aged and elderly Chinese men [2]. Additional main risk factors for HCC observed worldwide include external exposure to toxins (aflatoxin exposure, alcohol consumption) and hepatitis C virus (HCV) infection. However, the early manifestations of the process of HCC development from dysplastic nodules (DNs) to small

hepatocellular carcinoma (SHCC), which is defined by the very early HCC criteria of the Barcelona Clinic Liver Cancer (BCLC) criteria as a single HCC nodule less than 2 cm in diameter, are difficult to detect clinically. Thus, a large number of Chinese patients are diagnosed at an advanced stage and receive only palliative care [3]. The 5-year survival rate for HCC is only 33% in all races [4], but in China, it drops to 14.1% [5]. Therefore, it is specifically essential to find a new measurement to detect these changes in the early stage for diagnosis and intervention.

Medical imaging technologies, including contrast-enhanced computed tomography (CE-CT), contrast-enhanced magnetic resonance imaging (CE-MRI), and contrast-enhanced ultrasound imaging (CE-US), can achieve satisfactory diagnostic results for HCC. If the imaging profile on dynamic MR is specific for HCC (intense contrast-enhanced agent uptake in the arterial phase followed by extracellular contrast wash-out in the venous and/or delayed phase), a diagnosis can be made even without histological confirmation [6–8]. However, it is still difficult to identify very early lesions by medical imaging, even for an experienced radiologist. Although these lesions may have already become malignant, they remain microscopic on dynamic MRI and lack typical radiological markers [9].

The term “radiomics” was first proposed by Lambin et al. [10]. It is a field that focuses on improving image analysis and the high-throughput extraction of a vast number of quantitative features from medical images. The underlying basis of radiomics is that the analysis of these quantitative features can provide more and better information than physicians can by visually analysing these images. It has been reported that radiomics features of the tumour area present significant predictive efficacy in the classification of HCC [11–15]. Li Yang et al. [16] and Huang et al. [17] showed the best areas under the receiver operating characteristic curve (AUROCs) of 0.861 and 0.784 in the validation cohort, respectively. However, these previous studies were conducted on the assumption that HCC lesions could be observed on MR images, and few studies have focused on early or very early HCC lesions with no imaging changes.

Hence, we aimed to develop a multiphase MRI-based radiomics model to evaluate the risk of microscopic pre-HCC lesions for HCC and/or other liver-related disease patients.

Materials and methods

Patient cohort

The institutional review board of our institution approved this retrospective study. The requirement for informed consent was waived.

From September 2018 to June 2021, we retrieved the data of hospitalized patients screened for HCC due to cirrhosis secondary to HBV infection from Shandong Provincial Hospital affiliated to Shandong First Medical University.

The inclusion criteria for SHCC patients were as follows: (1) patients were diagnosed with SHCC by a radiologist with more than ten years of experience. (2) Patients received two MR examinations, of which SHCC was not apparent on the former but was diagnosed on the latter. The interval between the two examinations was not more than one year. (3) Patients had never been treated with transcatheter arterial chemoembolization (TACE) in the newly developed SHCC lesion. (4) Patients' MR imaging at the newly developed SHCC lesion site was consistent with the ‘wash-in and wash-out’ phenomenon described in the guidelines. (5) Patients had complete T1-weighted, arterial-phase, portal venous-phase, and delayed-phase MR images at both time-points. The SHCC lesions on the first MR scan of eligible SHCC patients were included in the SHCC cohort; a normal volume of interest (VOI) from the first MR scan of each eligible SHCC patient was included in the internal-control cohort. Moreover, we also set up an external-control cohort to ensure the robustness of the research findings. The inclusion criteria are as follows: (1) patients were diagnosed with hepatic cyst or haemangioma by a radiologist with more than ten years of experience. (2) There were one to three lesions, and each one measured less than 2.3 cm in diameter. (3) Patients had no history of hepatitis, drug-induced liver damage, or alcohol abuse. (4) Patients had complete T1-weighted, arterial-phase, portal venous-phase, and delayed-phase MR images. The exclusion criteria were as follows: (1) the tumour outline was unclear on MR images in the SHCC cohort. (2) The quality of the MR images was poor in the internal- or external-control cohort. Figure 1 shows the whole experimental design.

MR image acquisition and registration and delineation of volumes of interest (VOIs)

The MR image acquisition parameters are detailed in Table 1. The MR images were retrieved from the picture archiving and communication (PACS) system, including non-contrast-enhanced T1-weighted (T1WI), arterial phase, portal venous phase, and delayed phase images. In patients with SHCC, we collected two MR scans within an interval of less than 12 months. Imaging characteristics consistent with SHCC lesions could be detected in the 2nd scan but no visible changes could be observed by the naked eye in the exact lesion location in the 1st scan.

To ensure accurate VOI delineation, we used the Elastix toolbox [18] in the 3D slicer open-source software [19] version 4.1.1 (<http://www.slicer.org>) to register the images successively. The arterial phase images of the 2nd

Fig. 1 The flow chart of the whole experiment includes several steps of data acquisition, registration, outlining, feature extraction and selection, model construction, and prediction

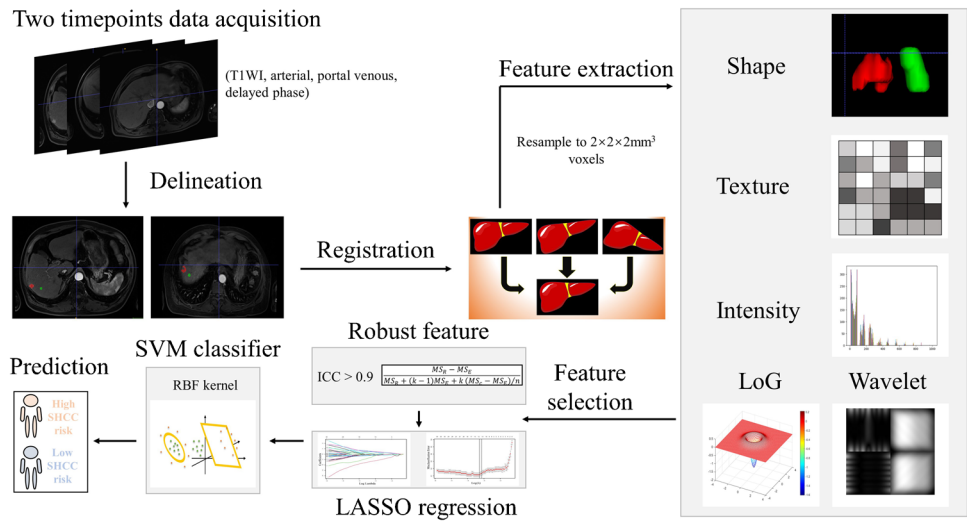


Table 1 The detail of the MR image acquisition

Parameters	Siemens 3.0 T Skyra (TIM Systems, Siemens Medical Solutions, Erlangen, Germany)	Siemens 3.0 T Prisma (TIM Systems, Siemens Medical Solutions, Erlangen, Germany)	Siemens 3.0 T Verio (TIM Systems, Siemens Medical Solutions, Erlangen, Germany)
Scanner	Siemens 3.0 T Skyra (TIM Systems, Siemens Medical Solutions, Erlangen, Germany)	Siemens 3.0 T Prisma (TIM Systems, Siemens Medical Solutions, Erlangen, Germany)	Siemens 3.0 T Verio (TIM Systems, Siemens Medical Solutions, Erlangen, Germany)
TR (ms)	4.31 (median)	3.92 (median)	3.92 (median)
TE (ms)	2.03 (median)	1.39 (median)	1.39 (median)
Slice Thickness (mm)	3 (median)	3 (median)	3 (median)
Dose of Gd-EOB-DTPA MR contrast agent	0.1 mmol/kg	13 ml	13 ml
Injection rate of Gd-EOB-DTPA MR contrast agent	2.0 ml/s	2.5 ml/s	2.5 ml/s
Acquisition time of the arterial phase	14 s	19 s	19 s
Acquisition time of the portal venous phase	26–30 s	25–32 s	60 s
Acquisition time of the delayed phase	60 s	3–5 min	3–5 min

scan were used as a template, and the arterial and other phases of the 1st scan were the targets to be successively registered.

Then, a radiologist with 20 years of experience manually delineated the SHCC tumour VOIs on the template images using ITK-SNAP open-source software version 3.8.0 (Yushkevich P and Gerig G). The tumour delineation covered the entire SHCC tumour lesion in all slices. For the internal-control cohort, the radiologist delineated the normal liver VOIs of these patients on MR images. A normal liver VOI was defined as a lack of imaging abnormality changes associated with SHCC on ten successive levels on both scans. The VOI for the external-control cohort was delineated identically as for the internal-control cohort. Consequently, we obtained 68 SHCC VOIs, 54

internal-control VOIs, and 70 external-control VOIs on the MR images from each phase.

Feature extraction

We performed a rescaling operation to normalize the MR images. The resampled voxel sizes were set to 2 × 2 × 2 mm³ to standardize the slice thickness. We mapped the VOIs onto the registered image to extract radiomics features using the PyRadiomics package [20] version 3.0.1. The radiomics features were generated from the original, wavelet-filtered, and Laplacian of Gaussian (LoG)-filtered images. The features included shape, intensity (‘First-order statistics’), and texture. Texture features included grey-level cooccurrence matrix (GLCM), grey-level size zone matrix (GLSZM),

grey-level run length matrix (GLRLM), neighbouring grey-tone difference matrix (NGTDM), and grey-level dependence matrix (GLDM) features.

Feature robustness and reproducibility

The robustness and reproducibility of the features were assessed with the intraclass correlation coefficient (ICC) [21–23]. Thirty patients from the SHCC cohorts were chosen randomly for VOI re-delineation by another experienced radiologist two weeks after the first delineation. Features with an ICC coefficient greater than 0.9 were retained and considered to have excellent robustness and reproducibility. Moreover, we concatenated the four-phase MR image features to evaluate whether the joint-phase (all-phase) features provided better discriminability than the single-phase features.

Feature selection

We implemented two types of analysis: intra-group classification and inter-group classification. The intra-group classification involved the 68 SHCC VOIs and 54 internal-control VOIs, and the inter-group classification involved the 68 SHCC VOIs and 70 external-control VOIs.

We randomly divided the intra-group classification ($n = 122$) and the inter-group classification ($n = 138$) data into a training set ($n = 92$ for intra-group classification, $n = 104$ for inter-group classification) and testing set ($n = 30$ for intra-group classification, $n = 34$ for inter-group classification) at a 3:1 ratio. These datasets came from the four sets of single-phase features and the single set of all-phase features. In total, there were ten datasets.

The least absolute shrinkage and selection operator (LASSO) regression model was used for feature selection. Employing regularization, the LASSO regression model adjusts the penalty coefficient value (λ), compresses most of the coefficients to zero, and retains the values with nonzero coefficients. Consequently, the retained features are nonredundant and sparse, potentially preventing the classifier from overfitting. By adjusting the parameter (λ) with the training set, the optimal features were screened via the minimum criteria with tenfold cross-validation.

Classifier and assessment of the performance of different models

We chose the radial basis-function, kernel-based support vector machine (RBF-SVM) as the classifier. First, the most valuable features filtered by the LASSO regression model from the four single phase-based and single combined-phase MR image features were used to train the corresponding classification models with the training set (intra-group and

inter-group classification). The optimal parameters of the SVM models were selected via tenfold cross-validation. Second, the efficacy of all models was tested with the testing set (intra-group and inter-group classification). Five models were established based on the RBF-SVM classifier: four according to the individual image phases (T1WI, arterial phase, portal venous phase, and delayed phase), and an all-phase model according to the integration of the four single-phase MR image features. The area under the receiver operating characteristic curve (AUROC), area under the precision-recall curve (AUPRC), sensitivity, specificity, and accuracy were calculated to evaluate classifier performance.

Statistical analysis

The SPSS version 26.0 (<http://www.ibm.com/>, IBM) and R version 4.0.1 (<https://www.r-project.org/>) statistical software was used for statistical analysis. The chi-square test was used to assess categorical data. The Shapiro–Wilk test was used to assess the normality of continuous data. The t test was used if the continuous data conformed to a normal distribution; otherwise, the Mann–Whitney U test was used. Normally distributed data are described as the mean \pm SD; otherwise, the median (IQR) was used. Differences in AUC values between different models on the testing set were estimated using the Delong test. A two-sided $p < 0.05$ was regarded as significant.

Results

Baseline characteristics

A total of 124 patients were enrolled in this study, including 54 SHCC patients (ranging from 36 to 73 years of age, with a mean of $56.9 \pm (9.35)$ years) and 70 hepatic cyst or haemangioma patients (ranging from 36 to 73 years of age, with a mean of $48.5 \pm (14.88)$ years). In the SHCC group, 4 patients were women and 50 were men. In the hepatic cyst or haemangioma group, 30 patients were women and 40 were men. There were statistically significant differences in age ($p < 0.001$ [t value: 3.65]) and gender ($p < 0.001$ [χ^2 value: 16.99]) between the two groups. The interval between the two MRI scans was $4.81 \pm (2.67)$ months. Details of the VOI numbers, volumes, and diameters are provided in supplementary Table 2.

Optimal radiomics signature

We extracted 1132 radiomics features from the VOIs at each phase, consisting of 14 shape, 18 intensity, 68 texture, 688 wavelet, and 344 LoG features. Through the first round of

screening using the ICC, 179, 325, 277, and 266 features remained among the T1WI and arterial, portal venous, and delayed phase image features, respectively (Fig. 1a of the supplementary material).

Subsequently, after filtration with LASSO regression, the T1WI, arterial phase, portal venous phase, delayed phase, and all-phase models were built with 9, 6, 6, 23, and 36 optimal radiomics features, respectively (Fig. 1b1, b2 of the supplementary material). The retained features of each model were considered the radiomics signatures in the intra-group classification. The all-phase model consisted of 8 T1W and 12 arterial phase, 8 portal venous phase, and 8 delayed phase image features (Fig. 1c1 of the supplementary material).

In addition, 16, 13, 12, 10, and 18 optimal radiomics features (Fig. 1b3, b4 of the supplementary material) were selected by LASSO regression in the models built from the T1WI, arterial phase, portal venous phase, delayed phase, and all-phase imaging features in the inter-group classification, respectively. The all-phase model was constructed from 7 T1W, 3 arterial phase, 1 portal venous phase, and 7 delayed phase imaging features (Fig. 1c2 of the supplementary material). The optimal radiomics signature of each model is detailed in Table 1 of the supplementary material.

Best performance of the models in intra-group and inter-group classification

For intra-group classification, the performance of the all-phase model was significantly greater than that of the single-phase classifiers. The optimal performance was achieved with an AUROC of 1.00 (95% CI, 1.00–1.00) with the training set and 0.93 (95% CI, 0.85–1.00) with the testing set (Fig. 2) and corresponding AUPRCs of 1.00 and 0.94 (Fig. 3), respectively. Basic first-order statistics (‘Range’ and ‘Maximum’) and high-dimensional texture features (Grey-Level Co-occurrence Matrix [‘DifferenceAverage’] and Grey-Level Size Zone Matrix [‘ZoneVariance’]) contributed to the model construction.

Likewise, the all-phase model achieved good performance in inter-group classification with the 18 optimal radiomics features. The RBF-SVM classifier achieved an AUROC of 0.99 (0.99–1.00) with the training set and 0.97 (95% CI, 0.92–1.00) on the testing set (Fig. 2) and corresponding AUPRCs of 1.00 and 0.98 (Fig. 3), respectively. In model construction, the most important contributing features were first-order statistics (‘Maximum’ and ‘Variance’) and high-dimensional texture features (Grey-Level

Fig. 2 Comparison of the AUROC on the training set (a) and testing set (b) of the different models from the intra-group classification, and training set (c), testing set (d) from the inter-group classification

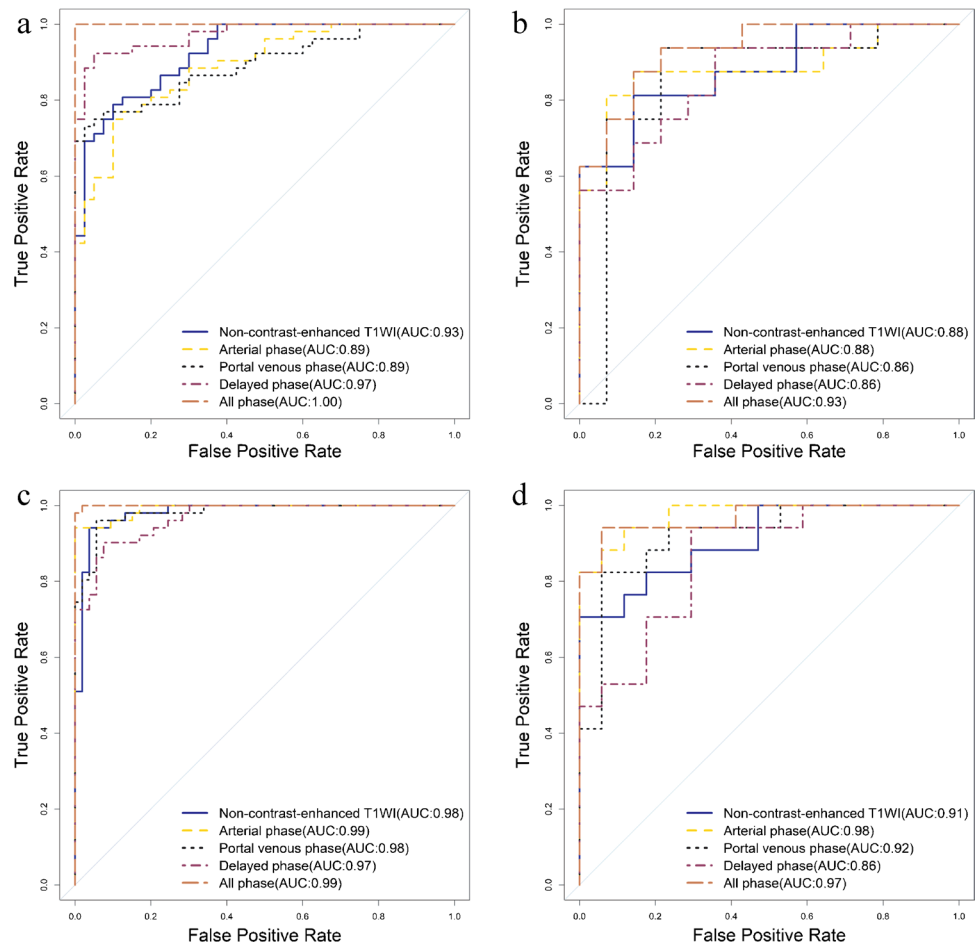
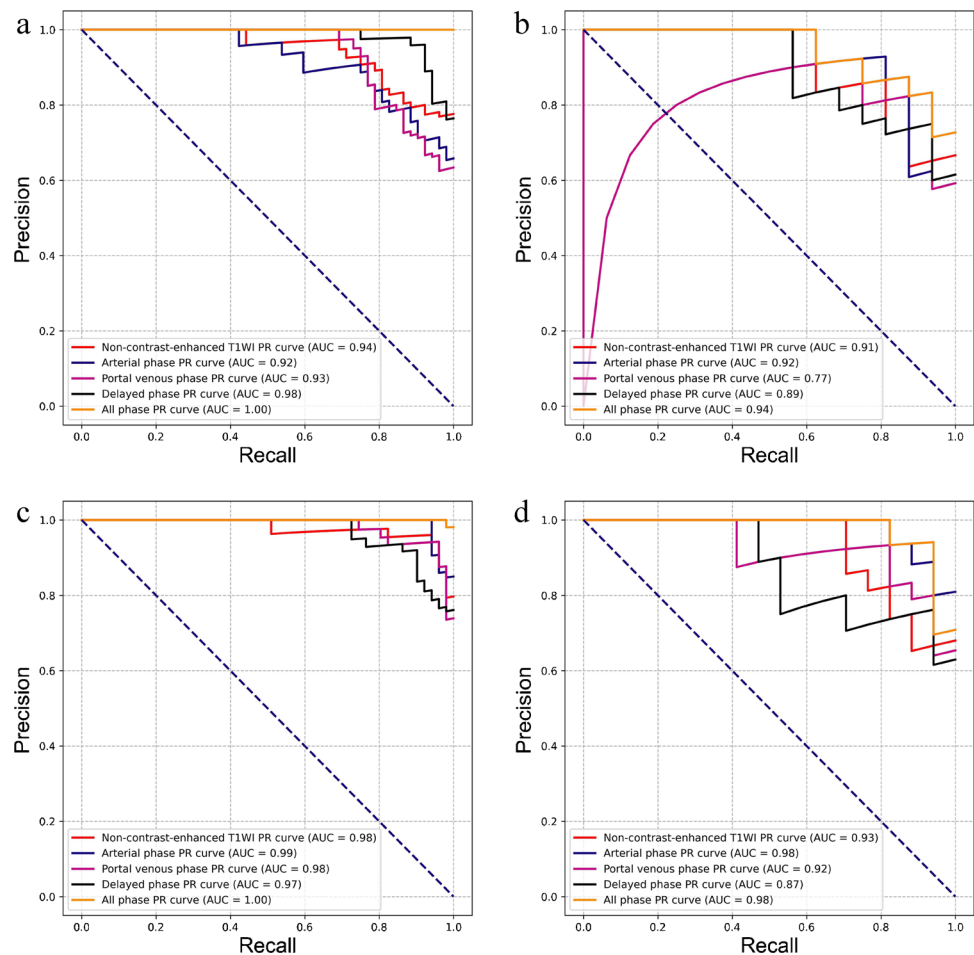


Fig. 3 Comparison of the AUPR on the training set (a) and testing set (b) of the different models from the intra-group classification, and training set (c), testing set (d) from the inter-group classification



Size Zone Matrix [‘ZoneVariance’] and Grey-Level Dependence Matrix [‘HighGrayLevelEmphasis’]). The performance details of the different models are shown in Table 2.

Differences in the AUROCs between intra-group and inter-group classification

With the testing set, we compared the performance of the same model in intra-group and inter-group classification according to the AUROC values. The Delong test revealed no differences between the values for the two classifications, $p > 0.05$ (Table 3). In addition, the top 3 radiomics features from the weighted ranking of the all-phase model were analysed to explore whether there were significant differences between normal tissue and undetectable SHCC lesions (Table 4). The results showed a significant difference ($p < 0.05$) in the values of the radiomics features between normal tissue area and undetectable SHCC lesions (Fig. 4). Figure 5 depicts expression heatmaps for the top two weighted features of the all-phase model for intra-group and inter-group classification. The heatmaps reveal that the heterogeneity within the tumour cannot be

reflected by MRI images, increasing the interpretability of the radiomics features.

Discussion

In this study, we aimed to develop a multiphase MRI-based radiomics model to evaluate the risk of microscopic lesions or high-risk nodules in HCC patients with radiologically undetectable lesions. This model showed clearly excellent performance in both intra-group and inter-group classification. These results are clinically significant in that such models can help physicians find potential malignant lesions for HCC and/or other liver-related disease patients in the very early stage.

In a majority of Chinese HCC patients, the disease tends to be secondary to HBV infection [24, 25], and a large proportion of patients are generally identified at an advanced stage [26]. Cirrhosis increases the risk of SHCC/HCC [27]. The transformation of cirrhotic nodules into HCC is a continuous and complex process [28], following the development of low-grade dysplastic nodules (LGDNs)

Table 2 The performance details of different models

Group									
Data set	Intra-					Inter-			
	Model	AUC (95%CI)	Acc	Sen	Spe	AUC (95%CI)	Acc	Sen	Spe
Training set	T1WI	0.93(0.877–0.976)	83.70%	78.85%	90.00%	0.98(0.958–1)	95.19%	94.12%	96.23%
	Arterial phase	0.89(0.815–0.951)	81.52%	75.00%	90.00%	0.99(0.981–1)	97.12%	94.12%	100%
	Portal venous phase	0.89(0.829–0.957)	83.70%	73.08%	97.50%	0.98(0.961–1)	95.19%	96.08%	94.34%
	Delayed phase	0.97(0.946–0.999)	93.48%	92.31%	95.00%	0.97(0.939–0.994)	91.35%	90.20%	92.45%
	All-phase	1 (1–1)	100%	100%	100%	0.99(0.999–1)	99.04%	100%	98.11%
Testing set	T1WI	0.88(0.759–1)	83.33%	81.25%	85.71%	0.91(0.816–1)	85.29%	70.59%	100%
	Arterial phase	0.88(0.754–1)	86.67%	81.25%	92.86%	0.98(0.937–1)	91.18%	82.35%	100%
	Portal venous phase	0.86(0.695–1)	86.67%	93.75%	78.57%	0.92(0.827–1)	88.24%	82.35%	94.12%
	Delayed phase	0.86(0.731–0.992)	80.00%	93.75%	64.29%	0.86(0.740–0.983)	82.35%	94.12%	70.59%
	All-phase	0.93(0.85–1)	86.67%	87.50%	85.71%	0.97(0.916–1)	94.12%	94.12%	94.12%

Abbreviation: *intra-*, intra-group classification; *inter-*, inter-group classification; *CI*, confident interval; *AUC*, the area under the curve; *Acc*, accuracy; *Sen*, sensitivity; *Spe*, specificity

Table 3 Comparison of the AUROC between inter- and intra-group classification on the testing set

Model name (inter- vs. intra-)	AUROC values	<i>p</i> values
T1W1 vs. T1W1	0.91 vs. 0.88	0.698
Arterial phase vs. arterial phase	0.98 vs. 0.88	0.194
Portal venous phase vs. portal venous phase	0.92 vs. 0.86	0.510
Delayed phase vs. delayed phase	0.86 vs. 0.86	0.999
All-phase vs. all-phase	0.97 vs. 0.93	0.478

Abbreviations: *inter-*, inter-group classification; *intra-*, intra-group classification; *AUROC*, the area under the receiver operator characteristic curve

to high-grade dysplastic nodules (HGDNs) to HCC [29]. Among them, DN is considered to be a precancerous lesion of SHCC/HCC [30–35]. DNs usually present with nontypical manifestations of influence on gadoteric acid (GD-EoB-DTPA)-enhanced MRI [36, 37]. According to a study by Eremites SC et al. [38], LGDNs result in only a 4% increase in arterial blood supply, HGDNs, an approximately 17–32%

increase, and HCC, an approximately 94% increase. Therefore, a large proportion of DNs presents with uniform T1 and T2 signals on MRI and does not show enhancement in the arterial phase [39], which makes it difficult to distinguish [40, 41]. The diagnosis of SHCC on MR depends on the increase in unmatched arterioles, the decrease in blood supply to the portal vein, the deposition of iron and lipids, and the changes in the formation of envelopes. From our experience, even before the radiologist can recognize SHCC/HCC lesions on MRI by the naked eye, the precancerous lesion may already exist. During the development of DNs or SHCC, few clinical symptoms are present, which can be fatal. Therefore, early inspection, detection, intervention, and treatment are particularly important.

To date, there have been several radiomics studies on hepatocellular carcinoma. Li Yang et al. [16] achieved an AUROC of 0.861 with the validation cohort of an MRI-based model incorporating significant clinical radiological factors and a fusion radiomics signature obtained from hepatobiliary phase (HBP) images. Huang et al. [17] reported that radiomics feature models based on CE-MR images

Table 4 The top 3 radiomics features weight of inter- and intra-group classification on all-phase model

Model	Filter	Feature class	Feature	Image phase	Weight
All-phase model (intra-)	Wavelet (HLL)	GLDM	Large Dependence High Gray Level Emphasis	Arterial	11.0103
	LoG ($\sigma=2$ mm)	First order	Range	T1WI	9.6618
	Wavelet (LHL)	First order	Maximum	Delayed	7.9649
All-phase model (inter-)	Original	First order	Maximum	Delayed	7.2267
	Wavelet (LLH)	GLSZM	Zone Variance	Arterial	3.8969
	LoG ($\sigma=4$ mm)	First order	Variance	Delayed	2.8917

Abbreviations: *intra-*, intra-group classification; *inter-*, inter-group classification; *LoG*, Laplacian of Gaussian; *GLDM*, gray level dependence matrix; *GLSZM*, gray level size zone matrix

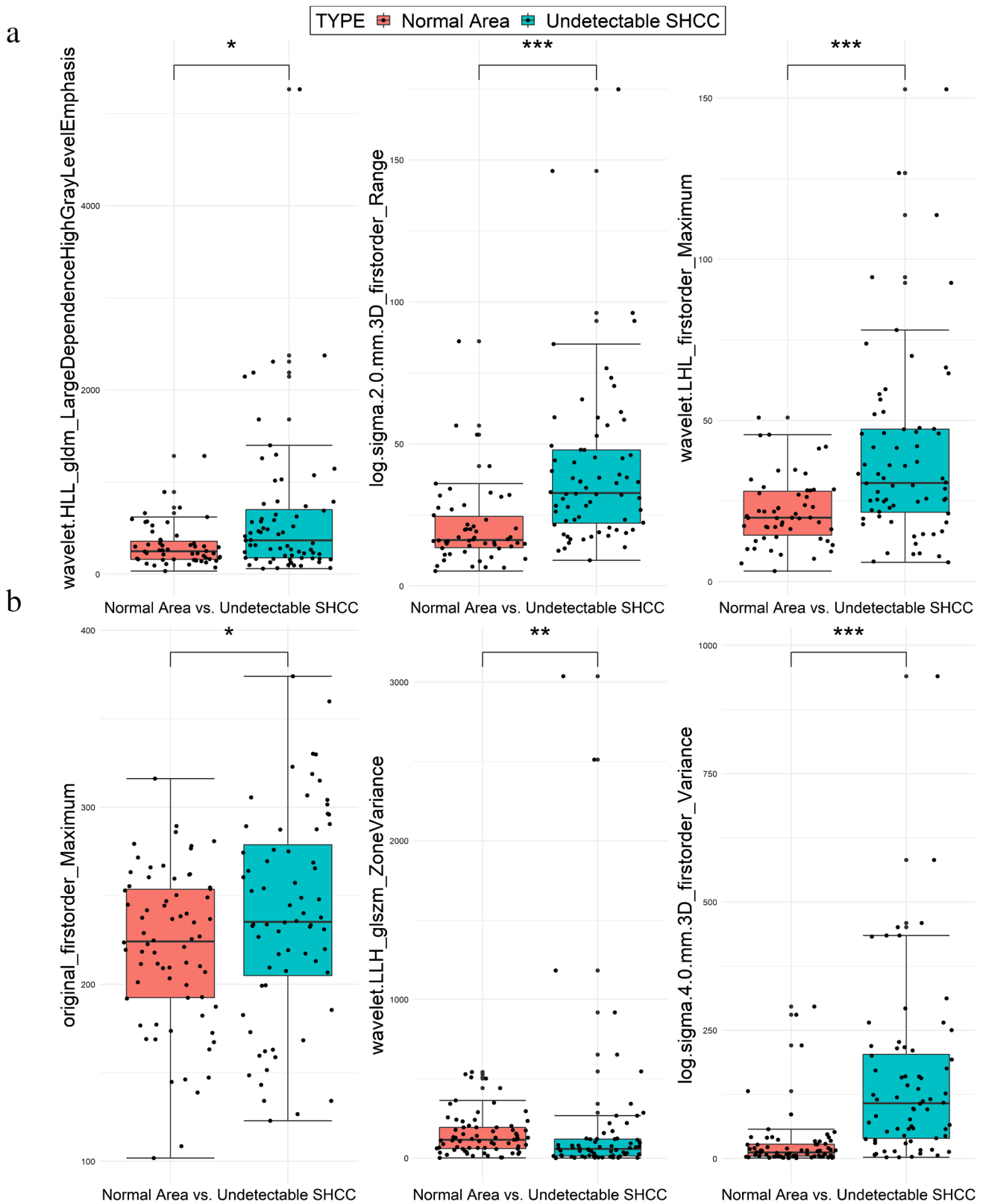


Fig. 4 Statistical analysis of the top 3 radiomics features a weight of intra-group classification (a) and inter-group classification (b) on the all-phase model

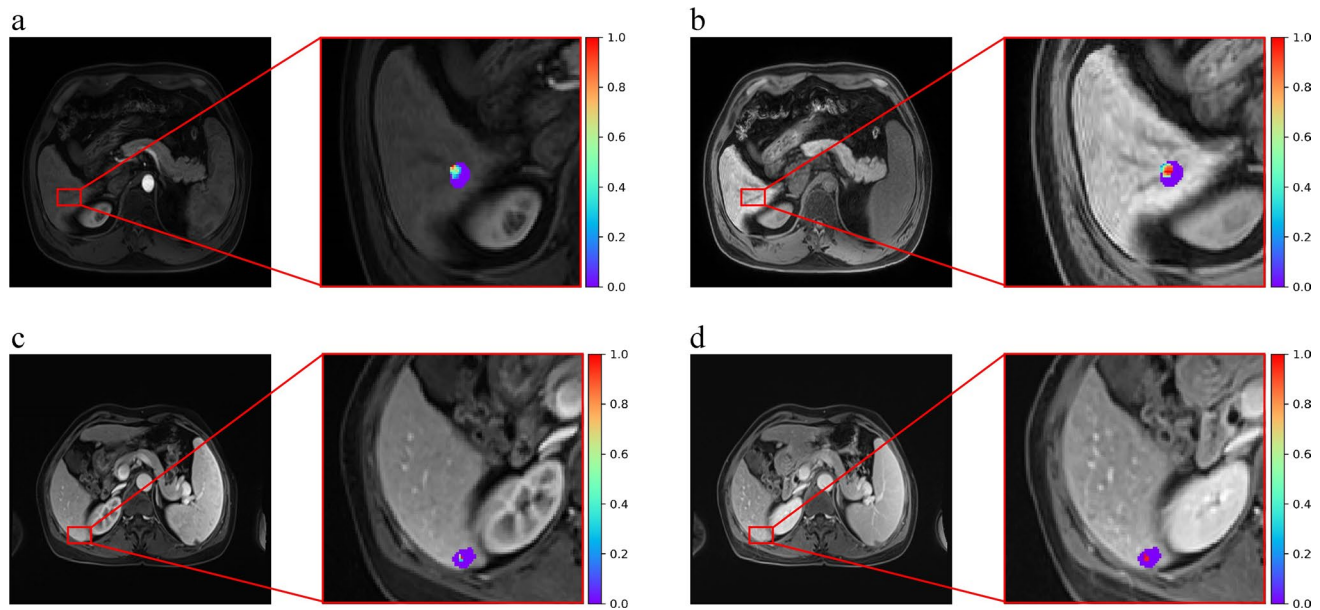


Fig. 5 Feature visualization of the gray level dependence matrix (GLDM) of arterial phase MRI by wavelet filter (a). Feature visualization of the first order Range of T1WI phase MRI by Laplacian of Gaussian (LoG; $\sigma=2$ mm) filter (b). Feature visualization of the gray

level size zone matrix (GLSZM) of arterial phase MRI by wavelet filter (c). Feature visualization of the first order Maximum of delayed phase MRI by none filter (d)

had favourable performance in predicting HCC, with mean AUROCs of 0.712, 0.784, 0.771, and 0.774 when constructed from arterial phase, portal venous phase, delayed phase, and hepatobiliary phase features, respectively. Many studies have been based on analyses of lesions that have already changed on visual imaging [42–47]. This makes sense if the lesion is detected before it can be observed by the naked eye.

In the current study, we included both an internal-control and an external-control cohort. If we had only set up an internal-control cohort, there could have been biases in the experimental results because we would not have been able to guarantee that the selected normal VOIs would not be disturbed by the information from underlying malignant lesions (although we attempted to avoid this as much as possible). Hence, we also selected a group of normal liver VOIs from patients diagnosed with hepatic haemangiomas or cysts as the external-control cohort to ensure a more stable experiment. The experimental results also confirm the existence of biases in one aspect; the AUROC for inter-group classification was generally better than that for intra-group classification with the testing set. However, this difference was not significant, with $p > 0.05$ obtained with the Delong test. This result indicates that despite the existence of biases, these models can still exert a predictive efficacy on the corresponding populations. The goal of the intra-group classification analysis is to screen HCC patients early to determine whether new high-risk lesions have developed, while the purpose of inter-group classification

analysis is to determine whether there are risky lesions related to HCC in the screening of non-HCC populations. We invited two radiologists (Mj.X. and Lq.C.) with five years of experience to perform a reader test with the testing set (Fig. 2 of the supplementary material). In intra-group classification, the AUROCs obtained by the two radiologists were 0.62 and 0.75, respectively. In the inter-group classification, the AUROCs obtained by the two radiologists were 0.71 and 0.73, respectively. When lesions are still in the microscopic stage, very early diagnosis can be greatly challenging for radiologists. Our model can assist radiologists in diagnosing these lesions ultra-early, producing great benefit to the patients.

Radiomics features related to image transformation are highly important for revealing information on very small lesions. The most common such transformation is the application of filters, as the image information obtained by different filter transformations may be different. According to Zhang et al. [48], the AUROC of a model constructed from filter-free radiomics features was 0.728 with the validation cohort. In a study of Zhao et al. [45], the radiomics features extracted from images processed by the wavelet filter were not included; only the original features and the features extracted from images transformed by the LoG filter were included. Among them, most of the features for constructing the optimal model were LoG features, comprising approximately ~75%, and the AUROC was 0.771 with the validation cohort. Throughout our research, in the best models for intra-group and inter-group classification,

the features extracted from images subjected to LoG filtering and wavelet transformation accounted for ~97% and ~78% of the model construction, and the AUROCs were 0.93 and 0.97 with the testing set, respectively. The information contained in the original images is not sufficient to explain all the phenomena and meet clinical needs. For further image analysis and research, some edge detection methods are often applied, such as LoG filters and wavelet transforms. The LoG is composed of a Gaussian kernel and Laplacian kernel; the latter is sensitive to areas with rapidly changing intensities, highlighting specific texture information in the original texture image and enhancing edges. A Gaussian smoothing filter is usually needed to smooth the image before the Laplacian operation to reduce the susceptibility to noise. Wavelet transformation produces good local characteristics. When the scale of the wavelet function is large, the anti-noise ability is strong, and when the scale of the wavelet function is small, the ability to extract image details is strong. Therefore, a balance between suppressing noise and extracting image edge details can be achieved. We recognized that some features reflecting tumour heterogeneity and microenvironment [7] were intensity (first-order) and texture (GLCM, GLSZM, etc.) features, consistent with other studies [22, 49–52]. Depending on the organ being imaged and the type of imaging modality, the first-order statistics may or may not have been the same across all applications. In intra-group classification, the top three features used to construct the best model were ‘Large Dependence High Grey Level’, ‘Range’, and ‘Maximum’ (Table 4). We found that the median values of these features in the lesion areas were generally higher than that in the normal areas [367.65 (175.13, 712.37) vs. 247.59 (159.17, 364.83), 32.73 (21.89, 47.89) vs. 16.17 (13.31, 24.88), and 30.61 (21.39, 47.29) vs. 19.83 (13.81, 28.25)]. In inter-group classification, the top three features used to construct the best model were ‘Maximum’, ‘Zone Variance’, and ‘Variance’ (Table 4). The average value of the ‘Maximum’ feature of the lesion areas was greater than that of the normal area [238.23 ± 58.92 vs. 220.63 ± 44.07]. The median of the ‘Variance’ feature value of the lesion areas was greater than that of the normal area [107.86 (38.94, 205.54) vs. 12.99 (5.01, 28.06)], but the ‘Zone Variance’ value was lower than that of the normal area [57.72 (59.56, 193.61) vs. 115.79 (12.76, 119.92)]. ‘Entropy’ and ‘Uniformity’ are two commonly used features computed in medical imaging. In a liver study [53], researchers showed that the ‘Total Entropy’ of the liver of healthy people was higher than that of patients with liver metastases. We calculated ‘Sum Entropy’ instead of ‘Total Entropy’ and found that the results corresponded to the above conclusions [1.39 (1.17, 1.64) vs. 1.44 (1.38, 1.50)]. Griethuysen et al. [18]

stated that ‘Zone Entropy’ measures the uncertainty/randomness in the distribution of zone sizes and grey levels. The higher values obtained in the lesion indicates more heterogeneity [3.41 (2.58, 4.05) vs. 2.00 (1.58, 2.58)] than in the normal area.

Limitations

The limitations of this study are as follows: first, the contour of the tumour area relied on manual delineation by an experienced radiologist, which required considerable time and energy expenditures. Second, this study relies on powerful registration algorithms, which can directly affect the accuracy of segmentation. For this reason, we asked the experienced radiologists to double-check the registration effect to minimize the impact of registration uncertainty on the experimental results. Third, a small number of samples were included, and this was a single-centre, retrospective study. The results of this study thus reflect only the patients at the centre and are not representative of the general population. Therefore, multicentre, large sample, and prospective studies are needed to further improve the results of this study.

Indeed, the essence of our experiment is hindsight. When applied to reality, we cannot define the VOI on the first MRI scan because there are no abnormalities when observed by the naked eye. However, the essence of research is to solve practical problems. The clinical transformations of this research in the future could entail the following: first, we need a software algorithm to automatically or semiautomatically identify and segment the liver. Second, a sliding convolution kernel of a specific size, e.g., 5×5 or 7×7 , would extract features from left to right sequentially for the whole liver or a specific liver segment and input these quantitative data into our model to obtain the predicted probability. These probabilities can be visualized with a heatmap, which conveniently displayed which areas are at high risk. This could help clinicians make decisions and intervene to achieve precision medicine.

Conclusions

Although new lesions in SHCC patients cannot be observed on MR imaging, a combination of radiomics features and machine learning algorithms can be sensitive to underlying abnormalities that cannot be detected by the naked eye.

In our study, we reported that the optimal model, based on the integration of radiomics features from four phases of MR images, could achieve excellent performance in evaluating microscopic pre-hepatocellular carcinoma lesions.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00259-022-05742-8>.

Author contribution Hy.W, Xm.W, and Jf.Q were the guarantor. K.S was involved in **statistical analysis**, data collection, and entry. K.S and Lt.S were involved in radiomics and machine learning analysis. All other members were involved in study conduct, manuscript draft, and revision.

Funding This study has received funding from Taishan Scholar Project of Shandong Province (81871354 and 81571672) and NSFC and Academic Promotion Programme of Shandong First Medical University (2019QL023) for Xm.W. The Cancer Prevention and Treatment fund of Shandong Province Natural Science Foundation (ZR2019LZL009) for Hy.W. Taishan Scholar Project of Shandong Province (TS201712065) and Academic Promotion Programme of Shandong First Medical University (2019QL009) and Science and Technology Funding from Jinan (2020GXRC018) for Jf.Q.

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

Declarations

Ethics approval The institutional review board approved this study.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

References

1. Wild C, Weiderpass E, Stewart B. World cancer report cancer research for cancer prevention. International Agency for Research on Cancer; 2020.
2. Yang T, Xing H, Wang G, et al. A novel online calculator based on serum biomarkers to detect hepatocellular carcinoma among patients with hepatitis B. *Clin Chem*. 2019;65(12):1543–53.
3. Liu X, Qin S. Immune checkpoint inhibitors in hepatocellular carcinoma: opportunities and challenges. *Oncologist*. 2019;24(Suppl 1):S3–10.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7–30.
5. Allemani C, Matsuda T, Carlo VD, et al. Global surveillance of trends in cancer survival: analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers during 2000–2014 from 322 population-based registries in 71 countries (CONCORD-3). *The Lancet*. 2018;391(10125):1023–75.
6. Ayuso C, Rimola J, Vilana R, et al. Diagnosis and staging of hepatocellular carcinoma (HCC): current guidelines. *Eur J Radiol*. 2018;101:72–81.
7. Vietti Violi N, Lewis S, Hectors S, et al. Radiological diagnosis and characterization of HCC. *Hepatocellular Carcinoma*. 2019;71–92.
8. Hartke J, Johnson M, Ghabril M. The diagnosis and treatment of hepatocellular carcinoma. *Semin Diagn Pathol*. 2016;34(2):153–9.
9. Sakamoto M. Early HCC: diagnosis and molecular markers. *J Gastroenterol*. 2009;44(19):108–11.
10. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441–6.
11. Tang A, Bashir MR, Corwin MT, et al. Evidence supporting LI-RADS major features for CT- and MR imaging-based diagnosis of hepatocellular carcinoma: a systematic review. *Radiol*. 2018;286(1):29–48.
12. Wu M, Tan H, Gao F, et al. Predicting the grade of hepatocellular carcinoma based on non-contrast-enhanced MRI radiomics signature. *Eur Radiol*. 2019;29(6):2802–11.
13. Zhou W, Zhang L, Wang K, et al. Malignancy characterization of hepatocellular carcinomas based on texture analysis of contrast-enhanced MR images. *J Magn Reson Imaging Jmri*. 2017;45(5):1476–84.
14. Hectors SJ, Wagner M, Bane O, et al. Quantification of hepatocellular carcinoma heterogeneity with multiparametric magnetic resonance imaging. *Sci Rep*. 2017;7(1):2452.
15. Zhou Y, He L, Huang Y, et al. CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. *Abdom Radiol*. 2017;42(6):1695–704.
16. Yang L, Gu D, Wei J, et al. A radiomics nomogram for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Liver Cancer*. 2019;8(5):373–86.
17. Huang X, Long L, Wei J, et al. Radiomics for diagnosis of dual-phenotype hepatocellular carcinoma using Gd-EOB-DTPA-enhanced MRI and patient prognosis. *J Cancer Res Clin Oncol*. 2019;145(12):2995–3003.
18. Klein S, Staring M, Murphy K, Viergever MA, Pluim J. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29(1):196–205.
19. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. 2012;30(9):1323–41.
20. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Can Res*. 2017;77(21):e104–7.
21. Hu Y, Xie C, Yang H, et al. Assessment of intratumoral and peritumoral computed tomography radiomics for predicting pathological complete response to neoadjuvant chemoradiation in patients with esophageal squamous cell carcinoma. *JAMA network open*. 2020;3(9):e2015927.
22. Zhu WS, Shi SY, Yang ZH, Song C, Shen J. Radiomics model based on preoperative gadoteric acid-enhanced MRI for predicting liver failure. *World J Gastroenterol*. 2020;26(11):1208–20.
23. Wang JC, Fu R, Tao XW, et al. A radiomics-based model on non-contrast CT for predicting cirrhosis: make the most of image data. *Biomark Res*. 2020;8:47.
24. Chen Z, Xie H, Hu M, et al. Recent progress in treatment of hepatocellular carcinoma. *Am J Cancer Res*. 2020;10(9):2993–3036.
25. Zhu ZX, Huang JW, Liao MH, Zeng Y. Treatment strategy for hepatocellular carcinoma in China: radiofrequency ablation versus liver resection. *Jpn J Clin Oncol*. 2016;46(12):1075–80.
26. De Stefano F, Chacon E, Turcios L, Marti F, Gedaly R. Novel biomarkers in hepatocellular carcinoma. *Dig Liver Dis*. 2018;50(11):1115–23.

27. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018;391(10127):1301–14.
28. Kudo M. Multistep human hepatocarcinogenesis: correlation of imaging with pathology. *J Gastroenterol*. 2009;44(Suppl 19):112–8.
29. Kitao A, Zen Y, Matsui O, Toshifumi Gabata A, Nakanuma Y. Hepatocarcinogenesis: multistep changes of drainage vessels at CT during arterial portography and hepatic arteriography–radiologic-pathologic correlation. *Radiol*. 2009;252(2):605–14.
30. Lee M, Kim K, Kim SY, et al. Genomic structures of dysplastic nodule and concurrent hepatocellular carcinoma. *Hum Pathol*. 2018;81:37–46.
31. Quaia E, De Paoli L, Pizzolato R, et al. Predictors of dysplastic nodule diagnosis in patients with liver cirrhosis on unenhanced and gadobenate dimeglumine-enhanced MRI with dynamic and hepatobiliary phase. *AJR Am J Roentgenol*. 2013;200(3):553–62.
32. Castelli G, Pelosi E, Testa U. Liver cancer: molecular characterization, clonal evolution and cancer stem cells. *Cancers*. 2017;9(9):127.
33. Theise ND, Park YN, Kojiro M. Dysplastic nodules and hepatocarcinogenesis. *Clin Liver Dis*. 2002;6(2):497–512.
34. Chou CT, Chou JM, Chang TA, et al. Differentiation between dysplastic nodule and early-stage hepatocellular carcinoma: the utility of conventional MR imaging. *World J Gastroenterol*. 2013;19(42):7433–9.
35. Kahng YS, Lee YS, Kim BK, Park WS, Lee JY, Kang CS. Loss of heterozygosity of chromosome 8p and 11p in the dysplastic nodule and hepatocellular carcinoma. *J Gastroenterol Hepatol*. 2003;18(4):430–6.
36. Okada M, Imai Y, Kim T, et al. Comparison of enhancement patterns of histologically confirmed hepatocellular carcinoma between gadoxetate- and ferucarbotran-enhanced magnetic resonance imaging. *J Magn Reson Imaging*. 2010;32(4):903–13.
37. Kudo M. The 2008 Okuda lecture: Management of hepatocellular carcinoma: From surveillance to molecular targeted therapy. *J Gastroenterol Hepatol*. 2010;25(3):439–52.
38. Efremidis SC, Hytiroglou P, Matsui O. Enhancement patterns and signal-intensity characteristics of small hepatocellular carcinoma in cirrhosis: pathologic basis and diagnostic challenges. *Eur Radiol*. 2007;17(11):2969–82.
39. Hong HS, Hua SK, Kim MJ, Becker JD, Mitchell DG, Kanematsu M. Single breath-hold multiarterial dynamic MRI of the liver at 3T using a 3D fat-suppressed keyhole technique. *J Magn Reson Imaging*. 2008;28(2):396–402.
40. Lee MH, Kim SH, Park MJ, Park CK, Rhim H. Gadoteric acid-enhanced hepatobiliary phase MRI and high-b-value diffusion-weighted imaging to distinguish well-differentiated hepatocellular carcinomas from benign nodules in patients with chronic liver disease. *AJR Am J Roentgenol*. 2011;197(5):868–75.
41. Kogita S, Imai Y, Okada M, et al. Gd-EOB-DTPA-enhanced magnetic resonance images of hepatocellular carcinoma: correlation with histological grading and portal blood flow. *Eur Radiol*. 2010;20(10):2405–13.
42. Liu F, Liu D, Wang K, et al. Deep learning radiomics based on contrast-enhanced ultrasound might optimize curative treatments for very-early or early-stage hepatocellular carcinoma patients. *Liver Cancer*. 2020;9(4):397–413.
43. Ning P, Gao F, Hai J, et al. Application of CT radiomics in prediction of early recurrence in hepatocellular carcinoma. *Abdom Radiol*. 2020;45(1):64–72.
44. Yuan C, Wang Z, Gu D, et al. Prediction early recurrence of hepatocellular carcinoma eligible for curative ablation using a Radiomics nomogram. *Cancer Imaging : Off Publ Int Cancer Imaging Soc*. 2019;19(1):21.
45. Zhao Y, Wu J, Zhang Q, et al. Radiomics analysis based on multiparametric mri for predicting early recurrence in hepatocellular carcinoma after partial hepatectomy. *J magn reson imaging : JMRI*. 2021;53(4):1066–79.
46. Shan QY, Hu HT, Feng ST, et al. CT-based peritumoral radiomics signatures to predict early recurrence in hepatocellular carcinoma after curative tumor resection or ablation. *Cancer imaging : Off Publ Int Cancer Imaging Soc*. 2019;19(1):11.
47. Kim S, Shin J, Kim D-Y, Choi GH, Kim M-J, Choi J-Y. Radiomics on gadoteric acid-enhanced magnetic resonance imaging for prediction of postoperative early and late recurrence of single hepatocellular carcinoma. *Clin Cancer Res*. 2019;25(13):3847–55.
48. Zhang Z, Jiang H, Chen J, et al. Hepatocellular carcinoma: radiomics nomogram on gadoteric acid-enhanced MR imaging for early postoperative recurrence prediction. *Cancer imaging : Off Publ Int Cancer Imaging Soc*. 2019;19(1):22.
49. Guo D, Gu D, Wang H, et al. Radiomics analysis enables recurrence prediction for hepatocellular carcinoma after liver transplantation. *Eur J Radiol*. 2019;117:33–40.
50. Wu J, Liu A, Cui J, Chen A, Xie L. Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images. *BMC Med Imaging*. 2019;19(1):23.
51. Wang W, Gu D, Wei J, et al. A radiomics-based biomarker for cytokeratin 19 status of hepatocellular carcinoma with gadoteric acid-enhanced MRI. *Eur Radiol*. 2020;30(5):3004–14.
52. Jiang YQ, Cao SE, Cao S, et al. Preoperative identification of microvascular invasion in hepatocellular carcinoma by XGBoost and deep learning. *J Cancer Res Clin Oncol*. 2021;147(3):821–33.
53. Brinkmann G, Melchert UH, Lalk G, et al. The total entropy for evaluating 31P-magnetic resonance spectra of the liver in healthy volunteers and patients with metastases. *Invest Radiol*. 1997;32(2):100–4.

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