European Radiology (2019) 29:4648–4659
https://doi.org/10.1007/s00330-018-5935-8

MAGNETIC RESONANCE

Preoperative prediction of microvascular invasion in hepatocellular cancer: a radiomics model using Gd-EOB-DTPA-enhanced MRI

Shi-Ting Feng[1] · Yingmei Jia[1] · Bing Liao[2] · Bingsheng Huang[3] · Qian Zhou[4] · Xin Li[5] · Kaikai Wei[1] · Lili Chen[2] · Bin Li[4] · Wei Wang[6] · Shuling Chen[6] · Xiaofang He[7] · Haibo Wang[4] · Sui Peng[4,8] · Ze-Bin Chen[9] · Mimi Tang[8] · Zhihang Chen[9] · Yang Hou[10] · Zhenwei Peng[11] · Ming Kuang[6,9]

Received: 5 July 2018 / Revised: 1 November 2018 / Accepted: 29 November 2018 / Published online: 28 January 2019

Abstract

Objectives Preoperative prediction of microvascular invasion (MVI) in patients with hepatocellular cancer (HCC) is important for surgery strategy making. We aimed to develop and validate a combined intratumoural and peritumoural radiomics model based on gadolinium-ethoxybenzyl-diethylenetriamine (Gd-EOB-DTPA)-enhanced magnetic resonance imaging (MRI) for preoperative prediction of MVI in primary HCC patients.

Methods This study included a training cohort of 110 HCC patients and a validating cohort of 50 HCC patients. All the patients underwent preoperative Gd-EOB-DTPA-enhanced MRI examination and curative hepatectomy. The volumes of interest (VOIs) around the hepatic lesions including intratumoural and peritumoural regions were manually delineated in the hepatobiliary phase of MRI images, from which quantitative features were extracted and analysed. In the training cohort, machine-learning method was applied for dimensionality reduction and selection of the extracted features.

Results The proportion of MVI-positive patients was 38.2% and 40.0% in the training and validation cohort, respectively. Supervised machine learning selected ten features to establish a predictive model for MVI. The area under the receiver operating characteristic curve (AUC), sensitivity, specificity of the combined intratumoural and peritumoural radiomics model in the training and validation cohort were 0.85 (95% confidence interval (CI), 0.77–0.93), 88.2%, 76.2%, and 0.83 (95% CI, 0.71–0.95), 90.0%, 75.0%, respectively.

Conclusions We evaluate quantitative Gd-EOB-DTPA-enhanced MRI image features of both intratumoural and peritumoural regions and provide an effective radiomics-based model for the prediction of MVI in HCC patients, and may therefore help clinicians make precise decisions regarding treatment before the surgery.

Shi-Ting Feng and Yingmei Jia are co-first authors.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5935-8) contains supplementary material, which is available to authorized users.

✉ Zhenwei Peng
pzhenw@mail.sysu.edu.cn

✉ Ming Kuang
kuangminda@hotmail.com

[1] Department of Radiology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[2] Department of Pathology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[3] National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China

[4] Clinical Trials Unit, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[5] GE Healthcare, Shanghai, China

[6] Department of Medical Ultrasonics, Division of Interventional Ultrasound, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[7] Department of Radiation Oncology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[8] Department of Gastroenterology and Hepatology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[9] Department of Liver Surgery, The First Affiliated Hospital of Sun Yat-sen University, 58 Zhong Shan Road 2, Guangzhou 510080, China

[10] Jinan University, Guangzhou, China

[11] Department of Oncology, The First Affiliated Hospital of Sun Yat-sen University, 58 Zhong Shan Road 2, Guangzhou 510080, China

Key Points

- *An effective radiomics model for prediction of microvascular invasion in HCC patients is established.*
- *The radiomics model is superior to the radiologist in prediction of MVI.*
- *The radiomics model can help clinicians in pretreatment decision making.*

Keywords Hepatocellular cancer · Radiomics · Magnetic resonance imaging · Gd-EOB-DTPA

Abbreviations

3D VIBE	Three-dimensional volume interpolated breath-hold test
AFP	Alpha-fetoprotein
AIC	Akaike information criterion
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUC	Area under receiver operating characteristic curve
CI	Confidence interval
CT	Computed tomography
FLASH	Fast low angle shot
FS	Fat suppression
Gd-EOB-DTPA	Gadolinium-ethoxybenzyl-diethylenetriamine
GGT	Gamma-glutamyltransferase
HASTE	Half-Fourier single-shot turbo spin-echo
HBP	Hepatobiliary phase
HCC	Hepatocellular cancer
ICC	Intra-class correlation coefficient
KW	Kruskal-Wallis
LASSO	Least absolute shrinkage and selection operator
MRI	Magnetic resonance imaging
MVI	Microvascular invasion
NPV	Negative predictive value
PPV	Positive predictive value
Rad score	Radiomics score
RFS	Recurrence-free survival
ROI	Region of interest
TSE	Turbo spin-echo
VOI	Volume of interest

Introduction

Liver resection and transplantation are considered as first-line curative treatments for hepatocellular cancer (HCC). Unfortunately, approximately 50% of patients recur within 2 years after curative hepatectomy [1, 2]. Microvascular invasion (MVI) is one of the well-known potential predictors for early recurrence of HCC [3, 4]. MVI is defined as the invasion of tumour cells within a vascular space lined by endothelium that is visible only on microscopy without macroscopic tumour invasion [5]. The presence of MVI indicates aggressive behaviour of HCC and poor survival outcomes [3].

The diagnosis of MVI is crucial for decision of surgery strategies. Absence of MVI is required for liver transplantation in HCC patients according to the new Milan criteria [6, 7]. If the presence of MVI can be accurately predicted preoperatively, anatomic resection with expanding resection margin should be performed even for a small tumour, while tumour resection might be enough for large tumours in cases with absence of MVI [8]. Besides surgical decision making, for patients diagnosed with MVI preoperatively, additional adjuvant therapies after surgery are preferred [9]. Thus, an accurate preoperative prediction of MVI plays an important role in treatment strategy making of HCC.

Currently, the diagnosis of MVI is determined only on the postoperative histologic examination. Due to high heterogeneity of HCC, no stable serological or genomic predictors of MVI have been found so far [10]. In recent years, several studies have found that some image features of computed tomography (CT) and magnetic resonance imaging (MRI) were predictive of MVI [11, 12]. Some studies showed that Gd-EOB-DTPA-enhanced MRI has a high value in the prediction of presence of MVI in HCC [11, 13]. In the study of Lee et al [11], three MR imaging features, including arterial peritumoural enhancement, non-smooth tumour margin, and peritumoural hypointensity on hepatobiliary phase (HBP), were independently associated with MVI [11]. However, in all these studies, imaging features were extracted visually. Although visual images provide some valuable feature information, their limitation in visual image grey scales restricts potential application in reflecting valuable microcosmic image features. Radiomics is a rapidly growing discipline based on quantitative image analysis to reflect image textures and morphology of tumours by grey value. It can obtain two-dimensional image features, as well as high-dimensional image features by extracting quantitative image features through a computer algorithm and thus can extract far more features than manual extraction by experts [14]. Several studies on skin cancer, glioma, and breast cancer showed that unviewable radiomics features were closely related to pathological microscopic structures. These features include texture features (a series of matrix transformations, such as grey-level co-occurrence metrics, run-length metrics, to reflect the high-order information of region of interest (ROI)), filter-transformed features (be used to obtain a series of target features, such as LOG transformation, Gaussian transformation),

wavelet features (the characteristics of the ROI through different resolution angles) [15–17].

Most previous radiomics studies highlighted the intratumoural region whereas the proximal liver parenchyma (distance from the tumour margin ≤ 2 cm) possessed main valuable information to diagnose MVI. Recently, a breast cancer radiomics study [17] tried to extract radiomics features from the intratumoural and peritumoural region simultaneously, finding that a combined radiomics method could produce generalisable and robust unsupervised clusters. This method was better than the extraction from the intratumoural region alone, suggesting that peritumoural radiomics added value in predicting treatment efficacy in comparison to use of intratumoural radiomics alone.

To the best of our knowledge, no literature has reported to apply a combined intratumoural and peritumoural radiomics technique to preoperative prediction of MVI of HCC. Therefore, we developed and validated radiomics models based on gadolinium-ethoxybenzyl-diethylenetriamine (Gd-EOB-DTPA)-enhanced MRI for preoperative prediction of MVI in patients with HCC through intratumoural radiomics and the combined intratumoural and peritumoural radiomics method, respectively.

Materials and methods

Patients

Patients who underwent Gd-EOB-DTPA-enhanced MRI examination before surgery between January 2013 and August 2017 were consecutively included in this study according to the inclusion and exclusion criteria. Only the patients who met all the following criteria were included: (1) resectable solitary HCC lesion or multiple HCC lesions within one liver lobe; (2) no macroscopic vascular invasion; (3) received Gd-EOB-DTPA-enhanced MRI of the liver within 1 month before surgery; (4) underwent curative hepatectomy; and (5) pathological confirmation of primary HCC. Exclusion criteria included: (1) received other anti-tumour therapies before surgery; (2) incomplete clinical or pathological information. Patients were randomly allocated to training and validation cohorts in a ratio of 7:3. Our Institutional Ethic Review Board has approved the current study, following the Declaration of Helsinki.

Pathological examination

Information about tumour size, number, colour, and capsule condition were obtained grossly, and information about histological type, differentiation grade, lymphocyte infiltration, MVI, satellite nodules, and chronic liver disease were obtained under microscope. MVI was defined as presence of tumour emboli in a vascular space lined by endothelial cells on microscopy [5]. Two pathologists with more than 10 years of experience in HCC pathology reviewed all the specimen slices independently, without knowing the patient's clinical data.

MR imaging acquisition

MRI examination was performed by using a 3.0-T system (Siemens Healthineers) in all patients. Eight-channel phased array coil was used, and the scanning scale covered from the top to the lower edge of the liver. The MR scan sequence included: half-Fourier single-shot turbo spin-echo (HASTE) sequence, fast low angle shot (FLASH) T1WI in/out of phase sequence imaging, FLASH T1WI fat suppression (FS) sequence axial imaging, and turbo spin-echo (TSE) T2WI navigation trigger axial imaging. The images in arterial phase, portal phase, and hepatic venous phase were obtained by performing three-dimensional volume interpolated breath-hold test (3D VIBE) T1WI (FS) sequence during suspended respiration at 30–35 s, 65–70 s, and 100–120 s, respectively, after as bolus injection of Gd-EOB-DTPA (Primovist®, 0.1 mL/kg body weight) with a flow rate of 1 mL/s. Additional hepatobiliary phase images were obtained at 20 min after injection. The specific parameters of each scan sequence are shown in Table 1.

Volume of interest identification and segmentation

The volumes of interest (VOIs) were delineated around the liver lesions outline for 3D volume area as indicated in hepatobiliary phase of MRI images by three independent radiologists with more than 5 years of experience with ITK-Snap software (open source software; www.itk-snap.org) (Supplementary Fig. 1) [18]. The reproducibility of the three radiologists to delineate VOI was assessed for further feature selection.

The intratumoural region was defined as the area within radiologist-annotated tumour boundaries. Then, the annotated intratumoural region was dilated at a radius of 1 cm by topology algorithm, generating the combined intratumoural and peritumoural region (dilated distance was analysed as dilated pixel counts multiplied by pixel size).

Radiomics feature extraction

The MR image features of all patients were extracted and analysed by the A.K. software version 2.0.0 (house-made software; Analysis-Kit, GE Healthcare). A total of 1044 imaging features were extracted, including four kinds of features (Fig. 1): 42 grey-level histogram (concerned with properties of individual pixels and they describe the distribution of voxel intensities within the image through commonly used and basic metrics), 972 transformed matrix texture (namely, texture features), 15 wavelet transformed texture, and 15 filter transformed texture.

Table 1 MRI scan sequence and parameters

Name	Orientation	Category	TR (ms)	TE (ms)	FOV (mm)	Matrix	Reverse angle	Band width	Seam thickness (mm)	Fat saturation	Breath-hold
T1WI IN/OUT PHASE	TRA	FLASH	200	2.2/1.1	328 × 350	192 × 256	65°	930/977	6	Yes/no	Yes
T1WI-FS	TRA	FLASH	235	2.2	328 × 350	240 × 320	70°	822	6	Yes	Yes
Contrast-enhanced											
T1WI	TRA	VIBE	3.3	1.2	328 × 350	128 × 256	13°	501	2	Yes	Yes
T2WI	TRA	TSE	2000	75	328 × 350	240 × 320	150°	672	6	Yes	No
T1WI (HBP)	TRA/COR	VIBE	3.3	1.2	328 × 350	154 × 256	13°	501	2	Yes	Yes

TRA, transverse axial; *COR*, coronal; *SAG*, sagittal; *FLASH*, fast low angle shot; *VIBE*, volume interpolated breath-hold test; *TSE*, turbo spin-echo; *FS*, fat suppression; *HBP*, hepatobiliary phase

Statistical analysis

Independent sample *t* test or Kruskal-Wallis (KW) nonparametric rank sum test was used to compare clinical characteristics between MVI-positive and MVI-negative groups in the training and validation cohort for continuous variables, while chi-square test or Fisher exact test for categorical variables. The reproducibility of three radiologists to delineate VOI was assessed by the intra-class correlation coefficient (ICC). ICC ≥ 0.8 indicated high consistency, 0.5–0.79 middle, and < 0.5 low [19].

The predictive model was established by combining the 1044 extracted features with pathologically diagnosed MVI outcomes through statistical methods of feature de-redundancy, supervised feature dimensionality reduction, and machine-learning-based classification. The least absolute shrinkage and selection operator (LASSO) logistic regression model was applied to dimensionality reduction through bootstrap. A 10-fold cross validation was used to select the minimum value of λ. Iteration was used, and frequencies of each features were ranked decreasingly. In order to avoid overfitting, the top 50 features were selected to establish the initial model. Finally, given collinearity of 50 features initially selected, logistic regression analysis was performed. Optimal model was obtained through Akaike information criterion (AIC) and area under the receiver operating characteristic curve (AUC). AUC and the corresponding 95% confidence interval (CI) were obtained from receiver operating characteristic curve (ROC) analysis, as well as the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the optimal cutoff value.

To compare the diagnostic performance of experts and radiomics model, univariable logistic regression models were used to evaluate the relationship between three MRI features extracted by Lee's study [11]. The association between the

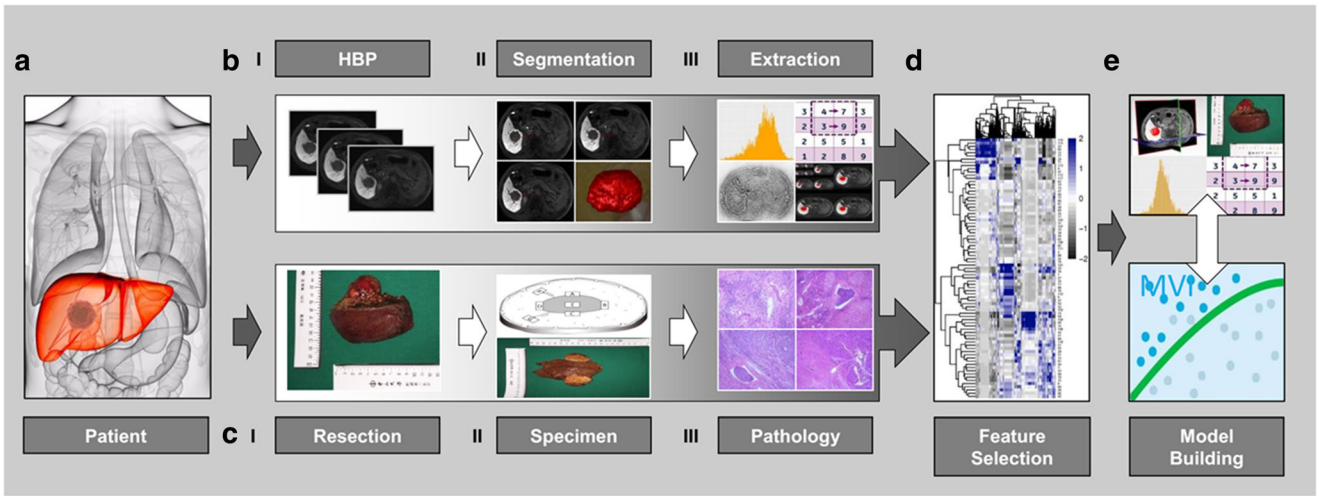

Fig. 1 Flow chart of the study. (**a**) Collecting HCC patients who met inclusion and exclusion criteria; (**b**) Extracting radiomics features: (I) obtaining Gd-EOB-DTPA-enhanced MRI images in the hepatobiliary phase; (II) segmentation: The volume of interest was delineated by experienced radiologists and three-dimensional images were formed; (III) extracting four kinds of quantitative features by software; (**c**) pathologic examination: (I) obtaining gross specimens of tumour tissue; (II) pathologic specimens; (III) pathologic diagnosis; (**d**) data cleaning and dimensions reduction; (**e**) establishing the model for predicting microvascular invasion by machine learning. Features through dimension reduction were applied to establish the model by machine learning and obtain quantitative radiomics score to predict microvascular invasion. HBP, hepatobiliary phase

Table 2 Baseline clinical characteristics of the training and validation cohort

	Total ($n = 160$)	Training ($n = 110$)	Validation ($n = 50$)	p values
MVI (%)	62 (38.8)	42 (38.2)	20 (40.0)	0.827
Age (years), mean ± SD	54.8 ± 11.4	54.6 ± 11.2	55.3 ± 12.0	0.700
Gender				0.821
Male	146 (91.3)	100 (90.9)	46 (92.0)	
Female	14 (8.8)	10 (9.1)	4 (8.0)	
ALT, U/L	34.0 (21.5, 50.5)	34.0 (20.0, 53.0)	33.5 (24.0, 48.0)	0.668
AST, U/L	32.0 (25.0, 46.5)	31.5 (25.0, 48.0)	33.0 (26.0, 46.0)	0.714
GGT, U/L	58.0 (33.0, 118.5)	58.0 (33.0, 119.0)	59.0 (34.0, 108.0)	0.925
Platelet count (10^9/L), mean ± SD	164.4 ± 66.9	162.0 ± 68.0	169.7 ± 64.7	0.501
AFP (μg/L)	25.6 (4.9, 235.8)	29.3 (5.8, 309.7)	18.5 (3.4, 185.8)	0.177
AFP group				0.613
≤ 20 μg/L	74 (46.3)	48 (43.6)	26 (52.0)	
20–400 μg/L	51 (31.9)	37 (33.6)	14 (28.0)	
> 400 μg/L	35 (21.9)	25 (22.7)	10 (20.0)	
Tumour number	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	0.812
Tumour number group				0.722
1	115 (71.9)	80 (72.7)	35 (70.0)	
≥ 2	45 (28.1)	30 (27.3)	15 (30.0)	
Tumour size (cm)	3.9 (2.7, 5.5)	3.8 (2.7, 5.0)	4.3 (2.7, 6.0)	0.528
Tumour size group				0.035
< 5 cm	111 (69.4)	82 (74.5)	29 (58.0)	
≥ 5 cm	49 (30.6)	28 (25.5)	21 (42.0)	
HBsAg				0.386
Negative	26 (16.3)	16 (14.5)	10 (20.0)	
Positive	134 (83.8)	94 (85.5)	40 (80.0)	
Child-Pugh class				0.553
A	157 (98.1)	107 (97.3)	50 (100.0)	
B	3 (1.9)	3 (2.7)	0 (0.0)	
BCLC stage				0.179
0	17 (10.6)	9 (8.2)	8 (16.0)	
A	105 (65.6)	77 (70.0)	28 (56.0)	
B	35 (21.9)	23 (20.9)	12 (24.0)	
C	3 (1.9)	1 (0.9)	2 (4.0)	

Continuous variables are presented as median (inter-quartile range, IQR) unless noted otherwise. Categorical variables are presented as n (%)

SD, standard deviation; *MVI*, microvascular invasion; *ALT*, alanine aminotransferase; *AST*, aspartate aminotransferase; *GGT*, gamma-glutamyltransferase; *AFP*, alpha-fetoprotein; *HBsAg*, hepatitis B surface antigen; *BCLC*, Barcelona Clinic Liver Cancer

clinical variables of the training cohort with MVI was assessed by univariable logistic regression analysis. And those variables with p value < 0.05 were included into the multivariable logistic regression analysis along with Rad score to further assess their independent association with MVI.

A heatmap analysis was performed to present associations between radiomics features and histopathologic features. It can visually represent data values in defined shades of colour. The radiomics score (Rad score) of each patient was calculated from the linear combination of the selected features multiplied by the coefficients estimated by the optimal logistic model.

Spearman correlation analysis was used to evaluate the associations between radiomics and histopathologic features.

The goodness of fit of the predictive model was evaluated by Hosmer-Lemeshow Wald chi-square test.

All statistical analyses were performed by R software version 3.2.3 (Bell Laboratories; https://cran.r-project.org/bin/windows/base/old/3.2.3). A two-sided p value was considered statistically significant if less than 0.05.

Results

Baseline characteristics

One hundred sixty patients were collected, including 110 patients in the training cohort and 50 in the validation cohort (Supplementary Fig. 2). Among all 160 patients, MVI was pathologically diagnosed in 62 patients (38.8%). The clinical characteristics of both cohorts are listed in Table 2. Baseline characteristics were not significantly different between both cohorts (Supplementary Table 1). Larger tumour size was detected in MVI-positive patients compared to that in MVI-negative patients in the training cohort, while higher serum AFP level and larger tumour size were found in MVI-positive patients compared to those in MVI-negative patients in the validation cohort. No significant differences were found between MVI-positive and MVI-negative patients in both cohorts in terms of other characteristics.

Interobserver and intraobserver reproducibility of radiomics feature extraction

The interobserver ICC was ≥ 0.8, 0.5–0.79, < 0.5 for 82%, 10% and 8% of the features, respectively. The intraobserver ICC was ≥ 0.8, 0.5–0.79, < 0.5 for 85%, 14% and 1% of the features, respectively.

Radiomics model establishment

The heatmap was used to show the correlation coefficient matrix among four kinds of features (Fig. 2a). To reduce dependency and redundancy, we used LASSO in logistic regression model to reduce dimensions of these 1044 features, and 50 features were selected. Logistic model was used and, finally, ten features were selected to establish the final model: kurtosis, percentile 10, percentile 75, percentile 80, ClusterShade_angle0_offset1 (ClusterShade), GLCMEntrophy_angle45_offset9 (GLCMEntrop), ShortRunEmphasis_AllDirection_offset2_SD (SRE), ShortRunLowGreyLevelEmphasis_AllDirection_offset5_SD (SRLGLE), HighGreyLevelRunEmphasis_AllDirection_offset8_SD (HGLRE), and LongRunHighGreyLevelEmphasis_AllDirection_offset8_SD (LRHGLE). Correlation coefficient matrix of these ten features were mostly negative but with low correlation coefficients (Fig. 2b), thus, were suitable for establishing a model for predicting MVI.

Hosmer-Lemeshow test showed the model established by these ten features was predictive in MVI (Hosmer-Lemeshow Wald chi-square = 10.188, $df = 8$, $p = 0.252$). Rad score was defined as a score resulted by the regression coefficients of these ten features multiplied by the value of corresponding feature (Formula see Supplementary Material). Each patient's Rad score in the training cohort was shown as bar chart in Fig. 3a, suggesting that MVI-positive patients presented higher score while MVI-negative patients presented lower score.

Validation of radiomics model

ROC curves of the training and validation cohorts were shown in Fig. 4a and b, respectively. The AUC, sensitivity, and specificity of our model in predicting MVI were 0.850 (95% CI,

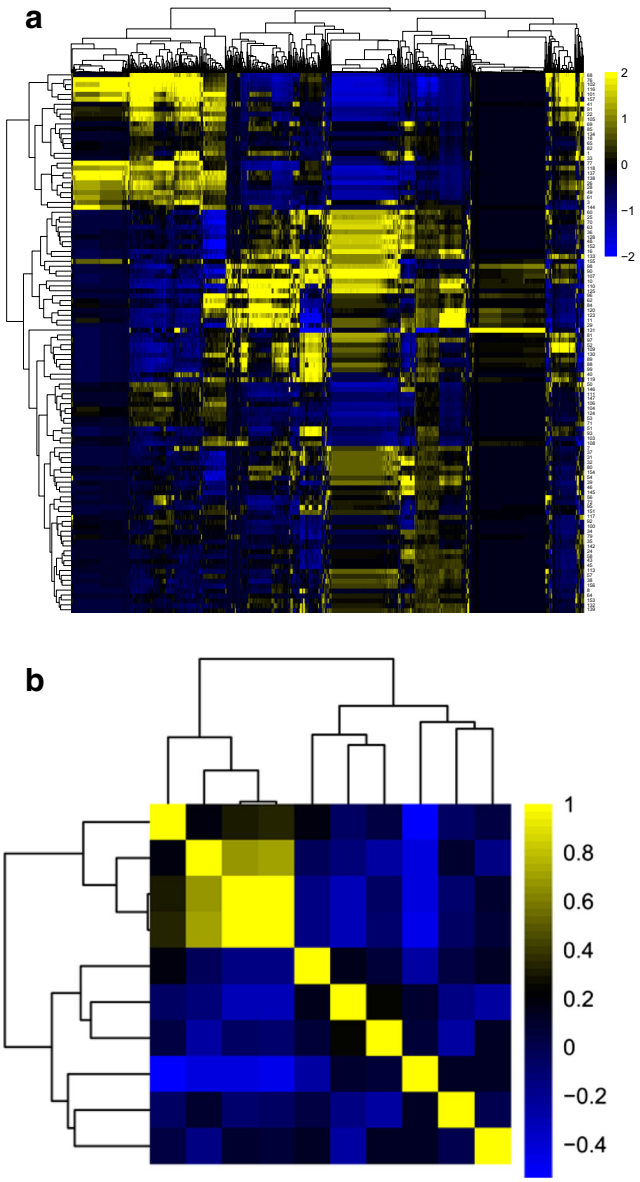

Fig. 2 Radiomics heatmaps. (**a**) Heatmap depicting correlation coefficients matrix of 1044 features in the training cohort. Unsupervised clustering analysis was used. Yellow was used to represent positive correlation and blue to represent negative correlation. (**b**) Heatmap depicting correlation coefficients matrix of ten selected features in the training cohort. Unsupervised clustering analysis was used. Yellow was used to represent positive correlation and blue to represent negative correlation

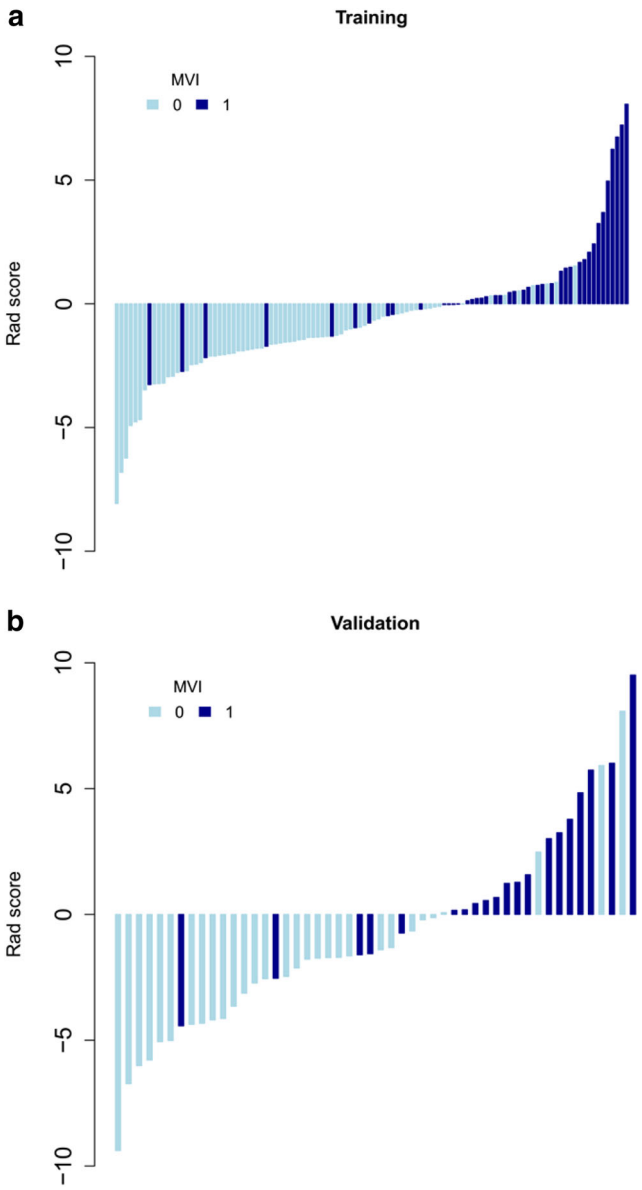

Fig. 3 Rad scores for each patient in the training and validation cohort. (**a**) Rad score for each patient in the training cohort. Light blue bars show scores for MVI-negative patients. Dark blue bars show scores for MVI-positive patients. (**b**) Rad score for each patient in the validation cohort. Light Blue bars show scores for MVI-negative patients. Dark blue bars show scores for MVI-positive patients. MVI, microvascular invasion; Rad score, radiomics score

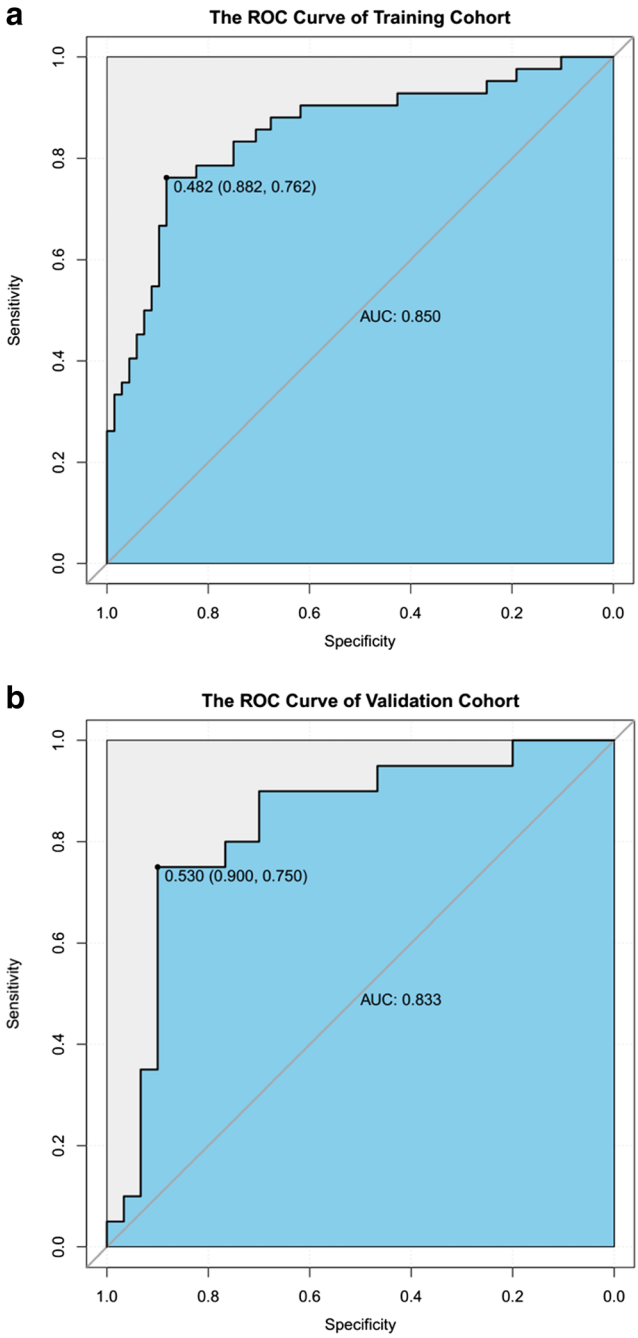

Fig. 4 Receiver operating characteristic curves (ROC) of the training and validation cohort. AUC, area under the receiver operating characteristic curve

0.77–0.93), 76.2%, and 88.2%, respectively, in the training cohort while 0.833 (95% CI, 0.71–0.95), 73.7%, and 90.0%, respectively, in the validation cohort.

MVI-positive patients had significantly higher Rad score in the training and validation cohorts (training cohort, 0.40 (− 0.04, 1.63) in MVI-positive vs. − 1.53 (− 2.19, − 0.48) in MVI-negative, $p < 0.001$; validation cohort, 0.96 (− 0.06, 3.39) vs. − 2.3 (− 4.3, − 1.35), $p < 0.001$) (Table 3, Fig. 3a, b).

Association of radiomics features with clinical data

In the training cohort, there was statistical difference between MVI-positive and MVI-negative group in terms of six of the selected ten features in univariable logistic regression model (Supplementary Table 2). Univariable analysis of the ten selected features and other clinical characteristics (Fig. 5) showed that eight features were associated with clinical characteristics listed

in Table 1. Five features (percentile 10, ClusterShade, LRHGLE, GLCMEntrophy, and percentile 75) were associated with BCLC stage. Seven features (kurtosis, percentile 10, percentile 80, ClusterShade, LRHGLE, GLCMEntrophy, and percentile 75) were associated with tumour size. In addition, percentile 10, percentile 75, percentile 80, and LRHGLE were associated with several clinical characteristics.

Association of radiomics features with histopathologic features

Three histopathologic features were analysed, including Edmondson grade, satellite nodules, and lymphocyte infiltration. Heatmap results showed that kurtosis, GLCMEntropy, and LRHGLE were associated with Edmondson grade ($p < 0.05$). Only Kurtosis and SRE were associated with satellite lesions ($p < 0.05$), and GLCMEntropy and HGLRE were associated with lymphocyte infiltration ($p < 0.05$). The other radiomics features were not associated with three histopathologic features (Fig. 6).

Comparison of diagnostic performance of experts and radiomics model

An abdominal radiologist with 17 years of experience in abdominal MR imaging performed image analysis according to these three findings extracted from preoperative MRI and established a model consists of these three features and MVI in the way Lee's [11] study adopted to predict MVI. Based on 160 patients in our study, when two of these three findings were combined, AUC, sensitivity, specificity, and accuracy were 0.57, 45.2%, 67.3%, and 58.8%, respectively. When these three findings were combined, AUC, sensitivity, specificity, and accuracy were 0.47, 19.3%, 83.7%, and 58.8%, respectively (Table 4). The radiomics model, of which AUC, sensitivity, specificity, and accuracy were 0.83, 90.0%, 75.0%, and 84.0%, respectively, seems superior to the radiologist in predicting MVI (Table 4).

Decision curve analysis used to predict MVI by radiomics model and experts was shown in Supplementary Fig. 3. It showed that if the threshold probability is larger than 20%,

Table 3 Association between radiomics features with MVI from binary logistic regression analysis

Intercept and features	Training			Validation		
	Estimate	OR (95% CI)	*p* values	Estimate	OR (95% CI)	*p* values
Multivariable analysis using selected features						
Intercept	− 0.64	0.53 (0.30, 0.88)	0.017*	− 0.26	0.77 (0.33, 1.79)	0.534
Kurtosis	1.46	4.29 (1.51, 15.65)	0.014*	1.35	3.87 (1.38, 15.6)	0.026*
Percentile 10	− 2.43	0.09 (0.01, 0.38)	0.004*	− 2.31	0.1 (0.02, 0.43)	0.006*
Percentile 75	21.30	1.79E + 09 (55.56, 5.18E + 18)	0.034*	10.10	2.43E + 05 (0.00, 1.03E + 16)	0.398
Percentile 80	− 19.83	0.00 (0.00, 0.03)	0.037*	− 7.85	0.00 (0.00, 5.94E + 5)	0.491
ClusterShade	1.02	2.78 (0.99, 8.87)	0.064	0.74	2.09 (0.69, 10.88)	0.257
GLCMEntropy	0.67	1.96 (1.05, 3.89)	0.041*	0.03	1.03 (0.37, 2.89)	0.953
SRLGLE	0.82	2.27 (1.13, 5.22)	0.033*	0.32	1.37 (0.94, 2.06)	0.099
HGLRE	− 0.48	0.62 (0.32, 1.09)	0.117	− 0.29	0.75 (0.26, 1.95)	0.557
LRHGLE	− 0.52	0.59 (0.28,1.17)	0.148	0.06	1.07 (0.46,2.35)	0.875
SRE	− 0.52	0.59 (0.32, 1.03)	0.079	− 0.08	0.92 (0.41, 2.00)	0.841
Radiomics score						
	1.00	2.72 (1.85, 4.37)	< 0.001*	0.34	1.41 (1.15, 1.82)	0.003*
AUC						
	0.85 (0.77, 0.93)		< 0.001#	0.83 (0.71, 0.95)		< 0.001#

'Kurtosis' is a measure of the 'peakedness' of the distribution of values in the image ROI, which can be used to describe the concentration degree of image brightness information. The percentile, p%, of a distribution is defined as that value of the brightness. Cluster shade is a measure of the skewness and uniformity of the GLCM. SRE is a measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more fine textural textures. GLCMEntrop measures the average amount of information required to encode the image values. SRLGLE measures the joint distribution of shorter run lengths with lower grey-level values. LRHGLRE measures the joint distribution of long run lengths with higher grey-level values. HGLRE measures the distribution of the higher grey-level values, with a higher value indicating a greater proportion of higher grey-level values and size zones in the image

OR, odds ratio; *AUC*, area under receiver operating characteristic curve; *ClusterShade*, ClusterShade_angle0_offset1; *GLCMEntrop*, GLCMEntrophy_angle45_offset9; *SRE*, ShortRunEmphasis_AllDirection_offset2_SD; *SRLGLE*, ShortRunLowGreyLevelEmphasis_AllDirection_offset5_SD; *HGLRE*, HighGreyLevelRunEmphasis_AllDirection_offset8_SD; *LRHGLE*, LongRunHighGreyLevelEmphasis_AllDirection_offset8_SD

*Statistically significant results from multivariate logistic regression analysis. # Statistically significant results from ROC analysis. Estimate is partial regression coefficient in logistic model

in other words, the estimated MVI-positive probability of a patient > 20%, the radiomics model would make more patients benefit from accurately predicting MVI, compared with the model based on imaging features extracted by experts.

Discussion

In this current study, we established a radiomics model predicting MVI preoperatively by extracting radiomics features from the intratumoural and peritumoural regions of Gd-EOB-DTPA-enhanced MRI. Our results showed that the AUC, sensitivity, and specificity of the combined intratumoural and peritumoural radiomics model were 0.83, 90%, and 75%, respectively. To the best of our knowledge, this is the first study to establish an MRI radiomics model for MVI prediction of HCC so far. The main reasons that the sensitivity and specificity of more than 75% were achieved in our model might be explained as follows. Firstly, radiomics requires accurate discrimination of lesion boundaries, and Gd-EOB-DTPA-enhanced MRI images were used in our radiomics model. It is known that more than half of HCCs have invaded capsule or completely lack them, and the infiltrative growth of the tumour further enhances the difficulty of lesion segmentation [20]. Gd-EOB-DTPA is a new hepatobiliary-specific MRI contrast agent [21]. The difference in signal between tumour tissues and surrounding liver parenchyma is more dominant in the hepatobiliary phase of Gd-EOB-DTPA-enhanced MRI than conventional contrast agents [22–24], which makes the boundaries of tumours clearer to delineate. Secondly, radiomics has the advantages of stable calculation, high repeatability, indefatigability, and no interference of human subjectivity [25, 26]. When we validated the efficacy of the features reported in the study of Lee et al [11] via visual inspection in our cohorts, the accuracy and sensitivity were much lower than those of our predictive model. This finding demonstrates the lower sensitivity, higher discrepancy, and poorer generalisability of extracting imaging features manually. Thirdly, the delineation of all the tumour slices in hepatobiliary phase extracted almost the whole tumour characteristics including the three-dimensional features (e.g. shape and smoothness), making the features more stable and representative than 2-dimensional regions of interest [27].

In our study, imaging features of intratumoural and peritumoural region were extracted simultaneously to establish the model. A previous study [5] has shown that more than 85% of MVI was found in peritumoural region within 1-cm distant from tumour boundaries. That was why the defined peritumoural region in our study was obtained by dilating the annotated intratumoural region at a radius of 1 cm. Lee et al [11] found three imaging features in Gd-EOB-DTPA highly suggestive of MVI. Among these features, arterial peritumoural enhancement and peritumoural hypointensity on hepatobiliary phase were both peritumoural features. The former was probably due to the local haemodynamic change of peritumoural region when MVI was present. The latter was probably attributed to the decreased intake of Gd-EOB-DTPA on hepatobiliary phase due to hepatic dysfunction induced by ischaemia. In light of these research findings, it was reasonable to come up with the idea that radiomics features of the peritumoural region were of great importance for the preoperative prediction of MVI. This is the first study to establish combined intratumoural and peritumoural radiomics model to predict MVI in HCC so far. Moreover, our study has

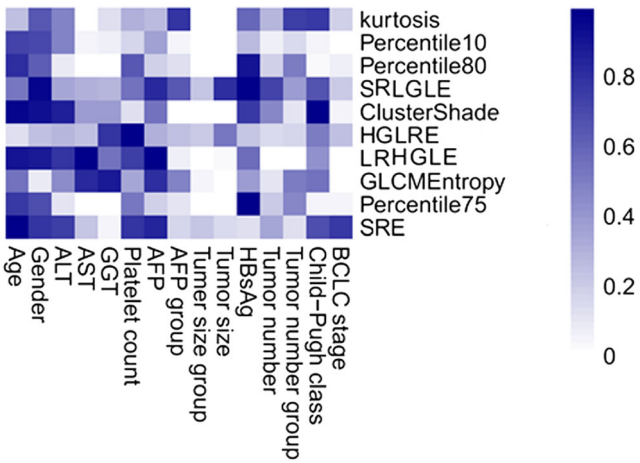

Fig. 5 Map of associations between ten features selected and clinical characteristics. Blue was used to represent high *p* values and white to represent low *p* values

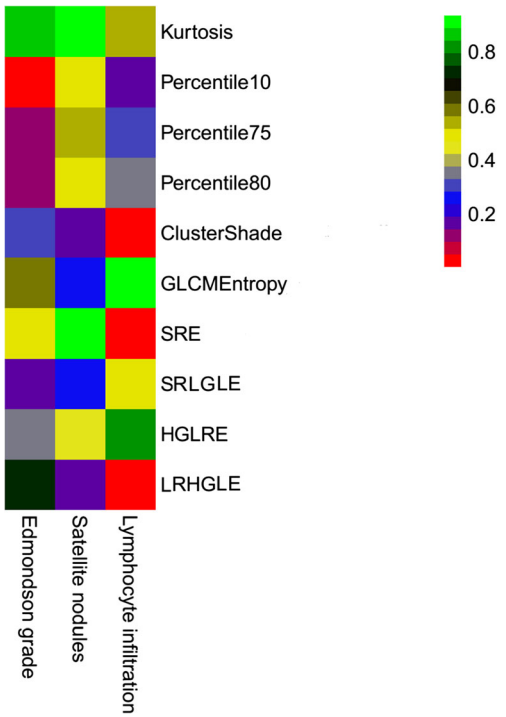

Fig. 6 Heatmap of associations between ten radiomics features and clinical characteristics. Red was used to represent low *p* values and green to represent high *p* values

demonstrated that the combined model was superior to intratumoural radiomics model in preoperative prediction of MVI in HCC.

Previous studies have found that radiomics features were closely related to tumour microscopic structure and biological behaviour [28–32]. Our study discovered ten radiomics quantitative features associated with MVI of HCC, which have not been reported before. Texture features are important markers of intratumoural homogeneity. Of the ten radiomics features associated with MVI in our study, four were histogram-based features (kurtosis, percentile 10, percentile 75, percentile 80) and others were matrix-based features (ClusterShade, GLCMEntrophy, SRE, SRLGLE, HGLRE, LRHGLE). Histogram-based features were first-order statistic, which mainly relied on the statistics of the intensity information (or brightness information) of the intratumour and peritumour, and then investigated the overall distribution of the intensity information of the intratumour and peritumour. For example, 'kurtosis' is a measure of the 'peakedness' of the distribution of values in the image ROI, which can be used to describe the concentration degree of image brightness information. A higher kurtosis implies that the mass of distribution is concentrated towards the tail(s) rather than towards the mean. The percentile (%) of a distribution is defined as that value of the brightness. Matrix-based features were second-order statistic, which can be used to describe the complexity of the intratumour and peritumour, changes in hierarchy, and the degree of thickness of the texture. For example, cluster shade is a measure of the skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean. SRE is a measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more finer textures. GLCMEntrop measures the average amount of information required to encode the image values. SRLGLE measures the joint distribution of shorter run lengths with lower grey-level values. The larger the value, the more complex the image, and the smaller the image grey value. LRHGLE measures the joint distribution of long run lengths with higher grey-level values. HGLRE measures the distribution of the higher grey-level values, with a higher value indicating a greater proportion of higher grey-level values and size zones in the image. On one hand, tumours with high-grade

Table 4 Comparison of diagnostic performance of three MRI imaging features and the radiomics model for MVI

MRI features	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
Study of Lee et al ($n = 197$)[#]					
Arterial peritumoural enhancement	54.0 (34/63)	88.1 (118/134)	77.2 (152/197)	68.0 (34/50)	80.3 (118/147)
Non-smooth tumour margin	69.8 (44/63)	68.7 (92/134)	69.0 (136/197)	51.2 (44/86)	82.9 (92/111)
Peritumoural hypointensity on HBP	31.7 (20/63)	91.8 (123/134)	72.6 (143/197)	64.5 (20/31)	74.1 (123/166)
Combination of any two findings	52.4 (33/63)	92.5 (124/134)	79.7 (157/197)	76.7 (33/43)	80.5 (124/154)
Combination of all three findings	19.0 (12/63)	99.3 (133/134)	73.6 (145/197)	92.3 (12/13)	72.3 (133/184)
Our patients included with the same criteria as the study of Lee et al ($n = 82$)					
Arterial peritumoural enhancement	21.4 (6/28)	81.5 (44/54)	61.0 (50/82)	37.5 (6/16)	66.7 (44/66)
Non-smooth tumour margin	67.9 (19/28)	37.0 (20/54)	47.6 (39/82)	35.8 (19/53)	69.0 (20/29)
Peritumoural hypointensity on HBP	14.3 (4/28)	85.2 (46/54)	61.0 (50/82)	33.3 (4/12)	65.7 (46/70)
Combination of any two findings	28.6 (8/28)	75.9 (41/54)	59.8 (49/82)	38.1 (8/21)	67.2 (41/61)
Combination of all three findings	7.1 (2/28)	90.7 (49/54)	62.2 (51/82)	28.6 (2/7)	65.3 (49/75)
All the patients in the present study ($n = 160$)					
Arterial peritumoural enhancement	33.9 (21/62)	76.5 (75/98)	60.0 (96/160)	47.7 (21/44)	64.7 (75/116)
Non-smooth tumour margin	79.0 (49/62)	30.6 (30/98)	49.4 (79/160)	41.9 (49/117)	69.8 (30/43)
Peritumoural hypointensity on HBP	35.5 (22/62)	73.5 (72/98)	58.8 (94/160)	45.8 (22/48)	64.3 (72/112)
Combination of any two findings	45.2 (28/62)	67.3 (66/98)	58.8 (94/160)	46.7 (28/60)	66.0 (66/100)
Combination of all three findings	19.3 (12/62)	83.7 (82/98)	58.8 (94/160)	42.9 (12/28)	62.1 (82/132)
All thepatients in the present study (n=160)					
Radiomics model (training, $n = 110$)	88.2 (60/68)	76.2 (32/42)	83.6 (92/110)	85.7 (60/70)	80.0 (32/40)
Radiomics model (validation, $n = 50$)	90.0 (27/30)	75.0 (15/20)	84.0 (42/50)	84.4 (27/32)	83.3 (15/18)

Sensitivity measures the proportion of true positives that are correctly identified. Specificity measures the proportion of true negatives that are correctly identified. Accuracy is the proportion of true results (both true positives and true negatives) among the total number of cases examined. PPV and NPV are the proportions of positive and negative results in statistics and diagnostic tests that are true positive and true negative results, respectively. [#] Data from the paper of Lee et al

PPV, positive predictive value; *NPV*, negative predictive value

malignancy may have greater heterogeneity and larger difference between cells, manifesting as grey-level nonuniformity and mixed signals in intratumoural regions. On the other hand, high-grade malignant tumours with faster growing speed would be more prone to necrosis due to poorer intratumoural blood supply, which leads to worse uniformity and more mixed signals in intratumoural regions [33]. In addition, high-grade malignant tumours would be more prone to invade capsules, and infiltrate grow, which leads to less uniformity and more mixed signals in peritumoural regions. As high-grade tumour is tending to present MVI, the texture features in MVI-positive HCC are less uniform than that of MVI-negative HCC, and the histogram features in MVI-positive HCC have more mixed signals than that of MVI-negative HCC in intratumoural and peritumoural regions. Therefore, the biological function similarity may influence microscopic pathologic similarity and subsequent similarity in radiomics features, which were identified in our study as valuable predictive features for MVI.

There are several limitations to this study. Firstly, the sample size is still limited compared with the relatively large number of variables. A large-scale clinical study enrolling more patients would definitely help validate and improve its applicability as an effective tool for predicting MVI in the decision-making strategy of HCC management. Secondly, our validation cohort was from the same centre as the training cohort, which restricted us to assess the generalisability of our findings to other centres and settings. Thirdly, in our study, normalisation of the signal intensities on MR images has not been performed. MR signals are a relative value rather than an absolute value, which may influence interpretation of signal characteristics on MVI prediction. Fourthly, the correlation of the actual site of MVI on histological specimen and VOI on MRI has not been performed.

In conclusion, this model showed high accuracy and sensitivity in both the training and validation cohorts, indicating its good representativeness and stability. Thus, it may be useful in preoperative individual prediction of MVI and assist clinicians in pretreatment decision making.

Funding This study was funded by the National Natural Science Foundation of China (81571750), National Natural Science Foundation of China (81771908), National Natural Science Foundation of China (81770608), Natural Science Foundation of Guangdong Province (2015A030311039), Guangzhou Science and Technology Program key projects (201704020215), and the Kelin Outstanding Young Scientist of the First Affiliated Hospital of Sun Yet-sen University (2017).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Ming Kuang.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry Two of the authors (Qian Zhou and Bin Li) have significant statistical expertise.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- diagnostic or prognostic study
- performed at one institution

References

1. Poon RT, Fan ST, Ng IO, Lo CM, Liu CL, Wong J (2000) Different risk factors and prognosis for early and late intrahepatic recurrence after resection of hepatocellular carcinoma. Cancer 89:500–507
2. Imamura H, Matsuyama Y, Tanaka E et al (2003) Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. J Hepatol 38:200–207
3. Lim KC, Chow PK, Allen JC et al (2011) Microvascular invasion is a better predictor of tumour recurrence and overall survival following surgical resection for hepatocellular carcinoma compared to the Milan criteria. Ann Surg 254:108–113
4. Zhou YM, Yang JM, Li B et al (2010) Risk factors for early recurrence of small hepatocellular carcinoma after curative resection. Hepatobiliary Pancreat Dis Int 9:33–37
5. Roayaie S, Blume IN, Thung SN et al (2009) A system of classifying microvascular invasion to predict outcome after resection in patients with hepatocellular carcinoma. Gastroenterology 137:850–855
6. Mazzaferro V, Llovet JM, Miceli R et al (2009) Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. Lancet Oncol 10:35–43
7. Iwatsuki S, Dvorchik I, Marsh JW et al (2000) Liver transplantation for hepatocellular carcinoma: a proposal of a prognostic scoring system. J Am Coll Surg 191:389–394
8. Shi M, Guo RP, Lin XJ et al (2007) Partial hepatectomy with wide versus narrow resection margin for solitary hepatocellular carcinoma: a prospective randomized trial. Ann Surg 245:36–43
9. Sun JJ, Wang K, Zhang CZ et al (2016) Postoperative adjuvant transcatheter arterial chemoembolization after R0 hepatectomy improves outcomes of patients who have hepatocellular carcinoma with microvascular invasion. Ann Surg Oncol 23:1344–1351
10. Cucchetti A, Piscaglia F, Grigioni AD et al (2010) Preoperative prediction of hepatocellular carcinoma tumour grade and microvascular invasion by means of artificial neural network: a pilot study. J Hepatol 52:880–888
11. Lee S, Kim SH, Lee JE, Sinn DH, Park CK (2017) Preoperative gadoxetic acid-enhanced MRI for predicting microvascular invasion in patients with single hepatocellular carcinoma. J Hepatol 67:526–534
12. Banerjee S, Wang DS, Kim HJ et al (2015) A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. Hepatology 62:792–800
13. Huang M, Liao B, Xu et al (2018) Prediction of microvascular invasion in hepatocellular carcinoma: preoperative Gd-EOB-

DTPA-dynamic enhanced MRI and histopathological correlation. Contrast Media Mol Imaging 2018:9674565
14. Kumar V, Gu Y, Basu S et al (2012) Radiomics: the process and the challenges. Magn Reson Imaging 30:1234–1248
15. Li Y, Liu X, Xu K et al (2018) MRI features can predict EGFR expression in lower grade gliomas: a voxel-based radiomic analysis. Eur Radiol 28:356–362
16. Esteva A, Kuprel B, Novoa RA et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542: 115–118
17. Braman NM, Etesami M, Prasanna P et al (2017) Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. Breast Cancer Res 19:57
18. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31:1116–1128
19. Leijenaar RT, Carvalho S, Velazquez ER et al (2013) Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability. Acta Oncol 52:1391–1397
20. Le TN, Bao PT, Huynh HT (2016) Liver tumor segmentation from MR images using 3D fast marching algorithm and single hidden layer feedforward neural network. Biomed Res Int 2016:3219068
21. Clément O, Mühler A, Vexler V, Berthezène Y, Brasch RC (1992) Gadolinium-ethoxybenzyl-DTPA, a new liver-specific magnetic resonance contrast agent. Kinetic and enhancement patterns in normal and cholestatic rats. Invest Radiol 27:612–619
22. Hamm B, Staks T, Mühler A et al (1995) Phase I clinical evaluation of Gd-EOB-DTPA as a hepatobiliary MR contrast agent: safety, pharmacokinetics, and MR imaging. Radiology 195:785–792
23. Joo I, Lee JM (2016) Recent advances in the imaging diagnosis of hepatocellular carcinoma: value of gadoxetic acid-enhanced MRI. Liver Cancer 5:67–87
24. Choi JW, Lee JM, Kim SJ et al (2013) Hepatocellular carcinoma: imaging patterns on gadoxetic acid-enhanced MR images and their value as an imaging biomarker. Radiology 267:776–786
25. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577
26. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
27. Lundsgaard Hansen M, Fallentin E, Axelsen T et al (2016) Interobserver and intraobserver reproducibility with volume dynamic contrast enhanced computed tomography (DCE-CT) in gastroesophageal junction cancer. Diagnostics (Basel). https://doi.org/10.3390/diagnostics6010008
28. Grossmann P, Stringfield O, El-Hachem N et al (2017) Defining the biological basis of radiomic phenotypes in lung cancer. Elife. https://doi.org/10.7554/eLife.23421
29. Fox MJ, Gibbs P, Pickles MD (2016) Minkowski functionals: an MRI texture analysis tool for determination of the aggressiveness of breast cancer. J Magn Reson Imaging 43:903–910
30. Hayano K, Tian F, Kambadakone AR et al (2015) Texture analysis of non-contrast-enhanced computed tomography for assessing angiogenesis and survival of soft tissue sarcoma. J Comput Assist Tomogr 39:607–612
31. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA (2013) Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. Radiology 266:326–336
32. Segal E, Sirlin CB, Ooi C et al (2007) Decoding global gene expression programs in liver cancer by noninvasive imaging. Nat Biotechnol 25:675–680
33. Zhou W, Zhang L, Wang K et al (2017) Malignancy characterization of hepatocellular carcinomas based on texture analysis of contrast-enhanced MR images. J Magn Reson Imaging 45:1476–1484