European Radiology (2019) 29:2890–2901
https://doi.org/10.1007/s00330-018-5797-0

ULTRASOUND

Ultrasound-based radiomics score: a potential biomarker for the prediction of microvascular invasion in hepatocellular carcinoma

Hang-tong Hu[1] · Zhu Wang[1] · Xiao-wen Huang[2] · Shu-ling Chen[1] · Xin Zheng[1] · Si-min Ruan[1] · Xiao-yan Xie[1] · Ming-de Lu[3] · Jie Yu[4] · Jie Tian[5] · Ping Liang[4] · Wei Wang[1] · Ming Kuang[3,6]

Received: 15 June 2018 / Revised: 24 August 2018 / Accepted: 24 September 2018 / Published online: 12 November 2018

Abstract

Purpose To develop an ultrasound (US)-based radiomics score for preoperative prediction of microvascular invasion (MVI) in hepatocellular carcinoma (HCC).

Methods Between January 1, 2012, and October 31, 2017, a total of 482 HCC patients who underwent contrast-enhanced ultrasound (CEUS) were retrospectively reviewed. The study population was divided into a training cohort ($n = 341$) and a validation cohort ($n = 141$) based on a cutoff time of January 1, 2016. Radiomics features were extracted from the grayscale US images of HCC. After features selection, a radiomics score was developed from the training cohort. The incremental value of the radiomics score to the clinic-pathological factors for MVI prediction was assessed in the validation cohort with respect to discrimination, calibration, and clinical usefulness.

Results The US-based radiomics score consisted of six selected features. Multivariate logistic regression analysis showed that the radiomics score, alpha-fetoprotein (AFP), and tumor size were independent predictors of MVI. The radiomics nomogram (based on the three factors) showed better performance for MVI detection (area under the curve [AUC] 0.731[0.647, 0.815] than the clinical nomogram (based on AFP and tumor size) (0.634 [0.543, 0.724]) ($p = 0.015$). Both nomograms showed good calibration. Decision curve analysis demonstrated that in terms of clinical usefulness, the radiomics nomogram outperformed the clinical nomogram.

Conclusion The US-based radiomics score was an independent predictor of MVI in HCC. Combining the radiomics score with clinical factors improved the prediction efficacy.

Key points

- *Radiomics can be applied in US images.*
- *US-based radiomics score was an independent predictor of MVI.*
- *Radiomics nomogram incorporated with the radiomics score showed good performance for MVI prediction.*

Hang-tong Hu and Zhu Wang contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5797-0) contains supplementary material, which is available to authorized users.

✉ Wei Wang
wangw73@mail.sysu.edu.cn

✉ Ming Kuang
kuangm@mail.sysu.edu.cn

[1] Department of Medical Ultrasonics, Institute of Diagnostic and Interventional Ultrasound, The First Affiliated Hospital of Sun Yat-sen University, No. 58 Zhongshan Road 2, Guangzhou, People's Republic of China

[2] Department of Ultrasonography, Zhongshan Hospital of Traditional Chinese Medicine, Affiliated to Guangzhou University of Chinese Medicine, Zhongshan, People's Republic of China

[3] Division of Interventional Ultrasound and Department of Liver Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, People's Republic of China

[4] Department of interventional Ultrasound, Chinese PLA General Hospital, Beijing 100853, People's Republic of China

[5] Key Laboratory of Molecular Imaging, Chinese Academy of Sciences, Beijing, People's Republic of China

[6] Department of Hepatobiliary Surgery, The First Affiliated Hospital of Sun Yat-sen University, No. 58 Zhongshan Road 2, Guangzhou 510080, Guangdong, People's Republic of China

Keywords Nomogram · Radiomics · Ultrasonography

Abbreviations

AFP	Alpha-fetoprotein
ALB	Albumin
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUC	Area under the curve
CEUS	Contrast-enhanced ultrasound
CT	Computed tomography
DCA	Decision curve analysis
HBsAg	Hepatitis B virus surface antigen
HBV-DNA	Hepatitis B virus DNA load
HCC	Hepatocellular carcinoma
ICC	Interclass correlation coefficient
LASSO	Least absolute shrinkage and selection operator
MRI	Magnetic resonance imaging
MVI	Microvascular invasion
PLT	Platelet
PT	Prothrombin time
ROC	Receiver operating curve
ROI	Region of interest
US	Ultrasound

Introduction

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death worldwide, with an incidence-to-mortality ratio close to 1.0 [1]. Liver resection and transplantation are widely considered curative treatments and provide the best outcomes [2]. However, the high rate of postsurgical recurrence leads to poor overall survival [3–5]. Precise detection of high-risk factors of recurrence before surgery may provide great help in reducing HCC relapse after surgery.

Microvascular invasion (MVI), which is defined as the invasion of tumor cells within a vascular space lined by endothelium, has been widely demonstrated to be a predictor of early recurrence of HCC [6, 7]. The presence of MVI indicates aggressive behavior of the HCC and poor survival outcomes [7]. However, in contrast to macrovascular invasion, MVI is not detectable with preoperative imaging techniques [1, 6]. Preoperative biopsy has also been demonstrated to be unreliable for detecting MVI due to the sampling error [8]. Thus, MVI is diagnosed only after surgery and is associated with a high recurrence rate [9]. Preoperative detection of MVI is necessary for appropriately selecting surgical candidates before liver transplantation and hepatic resection [10], which may improve the overall survival outcomes.

Previous studies have attempted to use preoperative imaging to detect MVI [11]. Imaging features, such as tumor dimension, non-smooth tumor margins, and peritumoral enhancement, identified on computed tomography (CT) or magnetic resonance imaging (MRI) have been previously suggested to be predictors of MVI [12–15]. However, in all these studies, these imaging features were visible to the naked eye. Although some valuable feature information can be detected in this manner, it is limited in terms of visual image grayscale, which restricts its potential for identifying valuable microcosmic image features that serve clinical practice. Radiomics is a rapidly growing discipline based on quantitative image analysis that reflects image textures and the morphology of tumors by gray values [16, 17]. This technique can detect high-dimensional image features by extracting quantitative image features with a computer algorithm; thus, it can extract far more features than manual feature extraction performed by experts [18–20].

To the best of our knowledge, no studies have yet reported the value of ultrasound (US) or contrast-enhanced ultrasound (CEUS) for detecting MVI. Compared with CT/MR, US is radiation-free, readily available, and easy to use for performing liver examinations. US/CEUS may become an alternative technique that provides additional information for detecting MVI if potential predictors can be identified. In the present study, we evaluated the imaging features of HCC using CEUS, and the radiomics technique was applied to build radiomics scores from the US images. A consequent radiomics nomogram was developed for the preoperative prediction of MVI.

Materials and methods

Study population

The study was approved by the ethical committee of the First Affiliated Hospital of Sun Yat-sen University and complied with the Declaration of Helsinki. Informed consent was waived for this retrospective research. Between January 1, 2012, and October 31, 2017, consecutive patients who had undergone liver resection for primary HCC were retrospectively reviewed. The clinic-pathological data and CEUS images were collected from the data system of our hospital. The inclusion criteria were as follows: (1) primary HCC confirmed after surgery, (2) solitary tumor, and (3) CEUS performed 2 weeks before surgery. The exclusion criteria were as follows: (1) macrovascular invasion (preoperative radiological evidence of macroscopic portal or hepatic vein invasion) [7], (2) preoperative anticancer therapy (radiotherapy, chemotherapy, etc.), (3) target lesion not completely visible on the US images, (4) incomplete clinic-pathological data or CEUS

images, or (5) insufficient pathological samples for the assessment of MVI [21].

Figure 1 displays the flow chart of the study population. Enrolled patients who underwent surgery between January 1, 2012, and December 31, 2015, were included as the training cohort, and those who underwent surgery between January 1, 2016, and October 31, 2017, were included as the validation cohort. In total, the training and validation cohorts comprised 341 patients (293 males and 48 females; mean age, 53.5 ± 10.7 years; range, 26 to 84 years) and 141 patients (124 males and 17 females; mean age, 53.4 ± 11.8 years; range, 18 to 83 years), respectively.

Clinic-pathological information

Baseline clinical data, including age, gender, family history of HCC, and serological results, were obtained from medical records. Data regarding serological biomarkers, including platelets (PLT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin (ALB), bilirubin, prothrombin time (PT), alpha-fetoprotein (AFP), hepatitis B virus surface antigen (HBsAg), and hepatitis B virus DNA load (HBV-DNA), were obtained within 1 month before surgery.

Pathological information included the presence of MVI, Edmondson-Steiner classification of HCC lesions, and the presence of liver cirrhosis. Six blocks of tissue were sampled from specific sites to confirm the presence of MVI. Four blocks of tissue were sampled from the junction of the tumor and the adjacent liver parenchyma. In addition, two other blocks of tissues were sampled from the tumor area and the proximal liver parenchyma (no more than 1 cm from the margins of the tumor) [21]. MVI was defined as the presence of tumor emboli in a vascular space lined by endothelium that was visible only on microscopy. Two pathologists (B.L. and L.L.C.) with at least 10 years of experience in pathology reviewed all the specimen slices. Both investigators were unaware of the clinical and imaging information of each patient. In cases of discordance, a consensus reading was performed, and the consensus data were used for following analysis.

US examination

The US equipment and settings used in the present study are detailed in the Supplement (Appendix Table, Table 1A). All the US examinations were performed by one of the four experienced radiologists who had more than 10 years of experience in liver CEUS. First, the entire liver was scanned with grayscale US. For the target lesion, the largest diameter, echo, shape (regular or irregular), boundary (well or poorly defined), and margin (smooth or not) were evaluated and recorded. Then, the transducer was fixed at the largest cross section of the lesion. A volume of 2.4 ml of SonoVue®

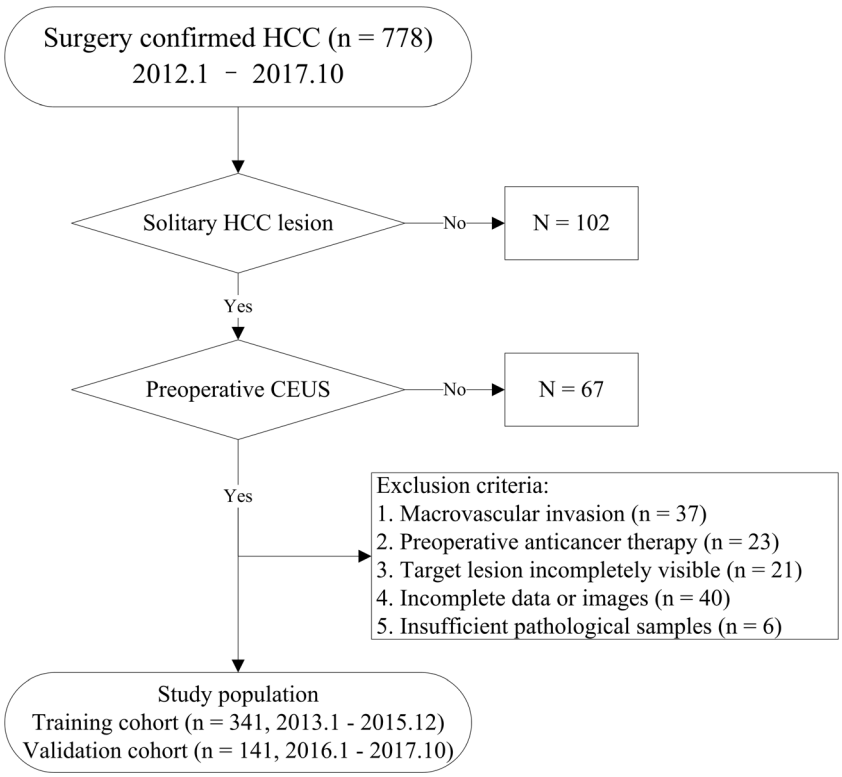

Fig. 1 Flow chart of enrolling the study population

(Bracco) was injected intravenously via the antecubital vein in bolus fashion, followed immediately by a flush with 5 ml of 0.9% normal saline solution. The targeted lesion was continuously observed, and at least 4-min digital cine clips were stored on the hard disk. CEUS features (including the enhanced level, enhanced homogeneity, enhanced boundary, and feeding artery) were evaluated and recorded [22, 23]. All the digital cine clips of the study population were retrospectively reviewed by two investigators (W.W. and Z.W.), each of whom has more than 8 years of experience in evaluating liver CEUS scans. Neither of the investigators was involved in the US examinations, and both were unaware of the clinical and other imaging information of the patients. They were asked to evaluate and record the imaging features of all the patients. In cases of discordance, a consensus reading was performed, and the consensus data were used for the following analysis.

Radiomics score

We built a radiomics score for each lesion using the radiomics technique. For each lesion, a region of interest (ROI) around the tumor outline was delineated on the grayscale US image of the largest cross section using the A.K. software (Artificial Intelligence Kit, version 1.1, GE Healthcare) (Fig. 2). Then, a total of 1044 features were extracted from the ROI using the A.K. software. Least absolute shrinkage and selection operator (LASSO) regression was performed to select features [24]. Finally, the radiomics score formula was generated using the selected features. The Supplement presents the detailed process of building the radiomics score formula (Appendix A1).

Two independent radiologists (X.W.H. and H.T.H., with at least 5 years of experience performing US examinations) were initially required to delineate ROIs on 100 randomly chosen images. Then, the inter-observer reproducibility was assessed, and the delineation for the remaining images was completed by the first radiologist.

Development and validation of MVI-predicting nomograms

Univariate and multivariate logistic regression analyses were performed in the training cohort to analyze the significant factors for MVI. Factors significant at the 0.10 level in the univariate analysis were included in the multivariate logistic regression analysis. In addition, p values below 0.05 were considered significant in the multivariate analysis. To demonstrate the incremental value of the radiomics score to the clinical risk factors for MVI estimation, both a radiomics nomogram and a clinical nomogram were developed. The radiomics nomogram incorporated the radiomics score and the independent clinical risk factors based on the multivariate analysis. The clinical nomogram incorporated only the independent clinical risk factors. The performances of the two nomograms were evaluated and compared with respect to discrimination, calibration, and clinical usefulness.

Discrimination Receiver operating curves (ROCs) were plotted to evaluate the performance of the nomogram in discriminating MVI positive from MVI negative in the training and validation cohorts. Discrimination was quantified with the area under the curve (AUC).

Calibration Calibration (i.e., agreement between the observed outcome frequencies and predicted probabilities) was plotted to explore the predictive accuracy of the nomogram [25].

Clinical usefulness Decision curve analysis (DCA) was conducted to determine the clinical usefulness of the nomogram by quantifying the net benefits at different threshold probabilities in the validation cohort [26].

Statistical analysis

Statistical analysis was conducted with SPSS 22.0 for Windows (Chicago, IL) and R software (version 3.4.1;

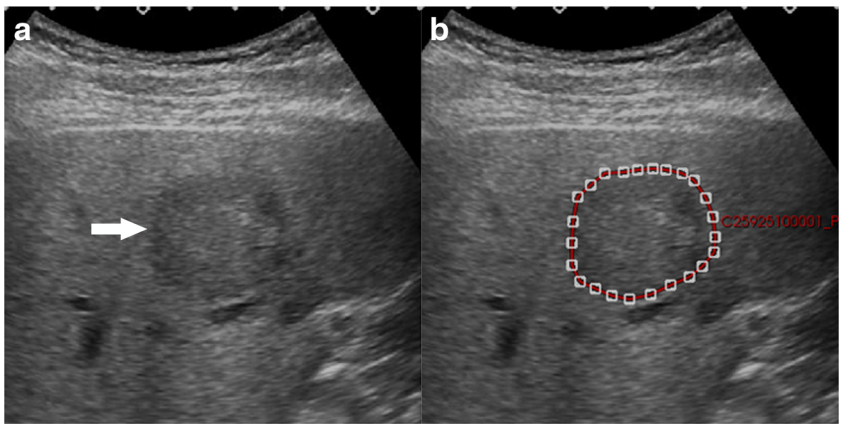

Fig. 2 Example of delineating region of interest (ROI) on grayscale US images. Arrows indicate solitary HCC lesion of a 54-year-old man (**a**). The tumor outline was delineated as the ROI (**b**)

Table 1 Clinic-pathologic characteristics and CEUS features in the training and validation cohorts

	Training cohort ($n = 341$)	Validation cohort ($n = 141$)	p value
Gender (male/female)	301/40	116/25	0.079
Age (years)			0.431
< 40	45 (13.2)	16 (11.3)	
40–50	98 (28.7)	32 (22.7)	
50–60	104 (30.5)	47 (33.3)	
> 60	94 (27.6)	46 (32.7)	
Family history of HCC (yes/no)	28/313	13/128	0.718
PLT (10^9/L)			0.797
< 100	49 (14.4)	19 (13.5)	
≥ 100	292 (85.6)	122 (86.5)	
ALT (U/L)			0.361
> 37	137 (40.2)	63 (44.7)	
≤ 37	204 (59.8)	78 (55.3)	
AST (U/L)			0.067
> 40	115 (33.7)	60 (42.6)	
≤ 40	226 (66.3)	81 (57.4)	
ALB (g/L)			0.857
< 35	53 (15.5)	21 (14.9)	
≥ 35	288 (84.5)	120 (85.1)	
Bilirubin (μmol/L)			0.569
> 22	51 (15.0)	24 (17.0)	
≤ 22	290 (85.0)	117 (83.0)	
PT (s)			0.839
< 14	21 (6.2)	8 (5.7)	
≤ 14	320 (93.8)	133 (94.3)	
AFP (μg/L)			0.080
< 20	127 (37.2)	62 (44.0)	
20–400	115 (33.7)	33 (23.4)	
> 400	99 (29.1)	46 (32.6)	
HBsAg (IU/mL)			0.575
≤ 0.05	51 (15.0)	17 (12.1)	
0.05–250	91 (26.7)	35 (24.8)	
> 250	199 (58.4)	89 (63.1)	
HBV-DNA (IU/mL)			0.686
< 100	113 (33.1)	45 (31.9)	
100–10^5	124 (36.4)	57 (40.4)	
> 10^5	104 (30.5)	39 (27.7)	
Child-Pugh (A/B)	332/9	133/8	0.109
Edmondson-Steiner classification			0.542
≤ II	210 (61.6)	91 (64.5)	
> II	131 (38.4)	50 (35.5)	
Cirrhosis (yes/no)	139/202	56/85	0.831
MVI (yes/no)	136/205	59/82	0.690
Tumor size			0.150
< 5 cm	228 (66.9)	81 (57.4)	
≥ 5 cm	113 (33.1)	60 (42.6)	
Echo			0.447
Mixed	159 (46.6)	74 (52.5)	

Table 1 (continued)

	Training cohort ($n = 341$)	Validation cohort ($n = 141$)	p value
Hypo	102 (29.9)	32 (22.7)	
Iso	59 (17.3)	26 (18.4)	
Hyper	21 (6.2)	9 (6.4)	
Shape			0.931
Regular	149 (43.7)	61 (43.3)	
Irregular	192 (56.3)	80 (56.7)	
Boundary			0.304
Well defined	143 (41.9)	52 (36.9)	
Poorly defined	198 (58.1)	89 (63.1)	
Margin			0.239
Smooth	113 (33.1)	39 (27.7)	
Non-smooth	228 (66.9)	102 (72.3)	
AP-enhanced level			0.087
Hyper	317 (93.0)	137 (97.2)	
Iso or hypo	24 (7.0)	4 (2.8)	
PVP-enhanced level			0.553
Hyper or iso	109 (32.0)	49 (34.8)	
Hypo	232 (68.0)	92 (62.2)	
DP-enhanced level			0.592
Hyper or iso	11 (3.2)	6 (4.3)	
Hypo	330 (96.8)	135 (95.7)	
AP-enhanced homogeneity			0.975
Homogeneous	194 (56.9)	80 (56.7)	
Heterogeneous	147 (43.1)	61 (43.3)	
AP-enhanced margin			0.809
Well defined	236 (69.2)	96 (68.1)	
Poorly defined	105 (30.8)	45 (31.9)	
Feeding artery (yes/no)	171/170	72/69	0.855
Radiomics score*	− 0.49 [− 0.69, − 0.24]	− 0.51 [− 0.69, − 0.19]	0.767

Unless otherwise indicated, data in parentheses are percentages. Mixed echo: hyperechogenicity in combination with a hypoechoic or anechoic portion. Feeding artery: hypertrophic artery directed towards the lesion and larger than the branches at the same depth during the arterial phase

AP arterial phase, *PVP* portal vein phase, *DP* delay phase, *hyper* hyper-enhancement, *iso* iso-enhancement, *hypo* hypo-enhancement

*Data in parentheses represents interquartile ranges

http://www.Rproject.org). Categorical variables were compared using the χ^2 test, and continuous variables were compared using the t test for variables with a normal distribution or the Mann-Whitney U test for variables with an abnormal or unknown distribution. The reported statistical significance levels were all two-sided, and p values of less than 0.05 were considered statistically significant. The inter-observer agreement of feature extraction was evaluated by using the interclass correlation coefficient (ICC). The ICC was graded as follows: poor (< 0.20), moderate (0.20 to 0.40), fair (0.40 to 0.60), good (0.60 to 0.80), or very good (0.80 to 1.00). R software was used to build and evaluate the prediction model. The "glmnet" package was used for LASSO regression. The "glm" function was used for the univariate and multivariate logistic regression analyses. The "Hmisc" package was used to plot the nomogram. The "pROC" package was used to plot the ROCs and measure the AUC. The "CalibrationCurves" package was used for the calibration curves. The "DecisionCurve" package was used to perform DCA.

Results

Clinic-pathological characteristics and CEUS features

The clinic-pathological characteristics and CEUS features of the training and validation cohorts are shown in Table 1.

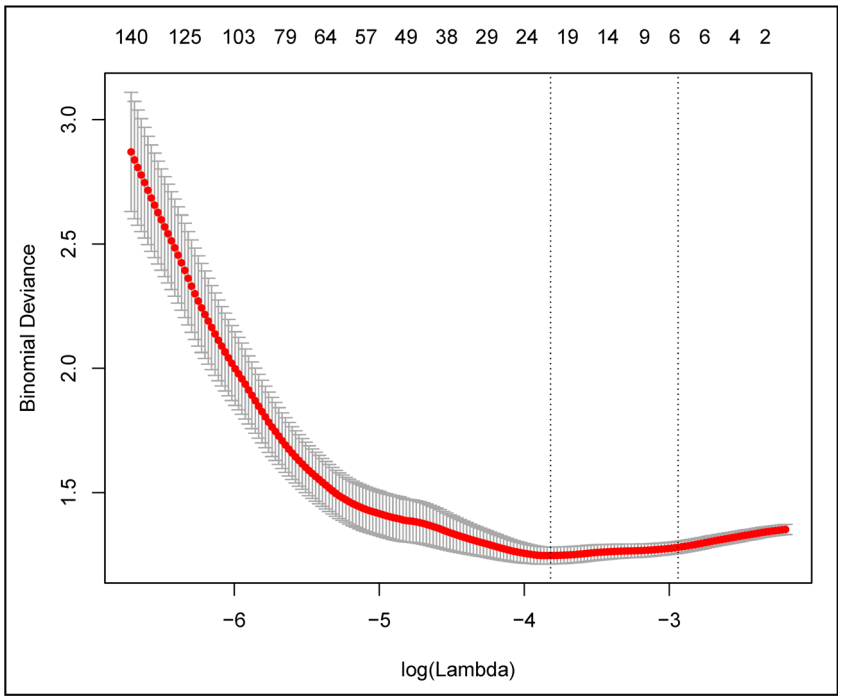

Fig. 3 Radiomics feature selection using the least absolute shrinkage and selection operator (LASSO) regression model in the training cohort. The 10-fold cross-validation (CV) process was repeated 50 times to generate the optimal penalization coefficient lambda (λ) in the LASSO model. The value of λ that gave the minimum average binomial deviance was used to select features. Dotted vertical lines were drawn at the optimal values using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria). A λ value of 0.053 with log (λ) − 2.937 was chosen (the 1-SE criteria) according to 10-fold CV, where optimal λ resulted in six nonzero coefficients

Positive MVI accounted for 39.9% (136/341) and 41.8% (59/141) of patients in the training and validation cohorts, respectively. There was no significant difference between the two cohorts in the presence of MVI ($p = 0.690$). In addition, there were no significant differences between the two cohorts in other clinic-pathological characteristics and CEUS features. These results justified the use of the training and validation cohorts.

Inter-observer reproducibility of feature extraction

The inter-observer reproducibility of feature extraction by the two radiologists was good, with ICCs ranging from 0.620 to 0.934. Therefore, all the outcomes were based on the measurements made by the first reader.

Radiomics score

Based on the training cohort, 1044 extracted features were reduced to six potential predictors using the LASSO regression model (Fig. 3). These six features were represented in the radiomics score formula (Supplement, Appendix A2). The radiomics score for each lesion in the training and validation cohorts was calculated using this formula. There was no significant difference between the two cohorts for the distribution of the radiomics score (Table 1, $p = 0.767$).

Development of MVI-predicting nomograms

Table 2 shows the results of the univariate and multivariate logistic regression analyses for MVI in the training cohort. AFP, tumor size, and radiomics score were the independent factors associated with MVI ($p < 0.001$). Thus, the radiomics nomogram was developed with AFP, tumor size, and radiomics score (Fig. 4a). The clinical nomogram was developed with AFP and tumor size (Fig. 4c).

Validation of MVI-predicting nomograms

Discrimination

Table 3 shows the performance of the radiomics score, clinical nomogram, and radiomics nomogram in discrimination between MVI-positive and MVI-negative patients. In the validation cohort, the AUCs for radiomics score, clinical nomogram, and radiomics nomogram were 0.689 (95% CI 0.598, 0.781), 0.634 (95% CI 0.543, 0.724), and 0.731 (95% CI 0.647, 0.815), respectively (in the training cohort, they were 0.712 [95% CI 0.655, 0.768], 0.674 [95% CI 0.617, 0.731], and 0.758 [95% CI 0.706, 0.810], respectively). The radiomics score and clinical nomogram showed similar discrimination in the training cohort ($p = 0.358$) and the validation cohort ($p = 0.398$), respectively. When combining the radiomics score with AFP and tumor size into the radiomics nomogram, the performance of the clinical nomogram was significantly improved from 0.674 to 0.758 in the training

Table 2 Results of the univariate and multivariate analyses based on the training cohort

Factors (reference)	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Gender (female)	1.870 (0.930, 4.050)	0.093	1.881 (0.824, 4.561)	0.145
Age, years				
< 40	Reference		Reference	
40–50	0.852 (0.419, 1.732)	0.657	1.050 (0.445, 2.496)	0.911
50–60	0.797 (0.394, 1.612)	0.527	1.329 (0.566, 3.165)	0.516
> 60	0.379 (0.179, 0.794)	0.010	0.535 (0.217, 1.310)	0.171
Family history of HCC (no)	0.825 (0.356, 1.813)	0.639	–	–
PLT (≥ 100 10^9/L)	1.272 (0.686, 2.338)	0.439	–	–
ALT (≤ 37 U/L)	2.075 (1.333, 3.245)	0.001	1.620 (0.851, 3.093)	0.141
AST (≤ 40 U/L)	2.039 (1.292, 3.230)	0.002	1.338 (0.681, 2.629)	0.397
ALB (< 35 g/L)	0.769 (0.426, 1.396)	0.383	–	–
Bilirubin (≤ 22 μmol/L)	1.868 (1.026, 3.420)	0.041	1.742 (0.863, 3.529)	0.121
PT (≤ 14 s)	0.585 (0.204, 1.479)	0.279	–	–
AFP (μg/L)				
< 20	Reference		Reference	
20–400	1.831 (1.065, 3.177)	0.030	1.592 (0.843, 3.023)	0.153
> 400	4.030 (2.312, 7.145)	0.000	4.021 (2.072, 7.971)	0.000*
HBsAg (IU/mL)				
≤ 0.05	Reference		–	–
0.05–250	0.873 (0.437, 1.749)	0.700	–	–
> 250	0.736 (0.396, 1.379)	0.335	–	–
HBV-DNA (IU/mL)				
< 100	Reference		–	–
100–10^5	1.072 (0.632, 1.821)	0.797	–	–
> 10^5	1.505 (0.875, 2.601)	0.140	–	–
Child-Pugh (B)	1.212 (0.295, 4.661)	0.777		
Edmondson-Steiner classification (≤ II)	1.651 (1.059, 2.580)	0.027	1.572 (0.925, 2.680)	0.095
Cirrhosis (no)	1.197 (0.770, 1.860)	0.423	–	–
Tumor size (≤ 5 cm)	2.028 (1.283, 3.218)	0.003	1.907 (1.108, 3.299)	0.020*
Echo				
Mixed	Reference		–	–
Hypo	0.995 (0.594, 1.656)	0.983	–	–
Iso	1.012 (0.594, 2.021)	0.756	–	–
Hyper	2.142 (0.857, 5.535)	0.105	–	–
Imaging signs				
Shape (irregular)	0.763 (0.490, 1.182)	0.227	–	–
Boundary (poorly defined)	0.816 (0.523, 1.267)	0.366	–	–
Margin (non-smooth)	0.797 (0.498, 1.265)	0.340	–	–
AP-enhanced level (iso or hypo)	0.924 (0.401, 2.203)	0.853	–	–
PVP-enhanced level (hyper or iso)	0.971 (0.611, 1.549)	0.900	–	–
DP-enhanced level (hyper or iso)	3.077 (0.778, 20.387)	0.155	–	–
AP homogeneity (heterogeneous)	0.934 (0.603, 1.448)	0.759	–	–
AP-enhanced margin (poorly defined)	0.886 (0.556, 1.418)	0.611	–	–
Feeding artery (no)	0.775 (0.501, 1.196)	0.251	–	–
Radiomics score	12.145 (5.761, 27.262)	0.000	11.079 (4.964, 26.315)	0.000*

Factors significant at the 0.10 level in the univariate analysis were included in the multivariate logistic regression analysis. In addition, *p* values below 0.05 (indicated with asterisk) were considered significant in the multivariate analysis

AP arterial phase, *PVP* portal vein phase, *DP* delay phase, *hyper* hyper-enhancement, *iso* iso-enhancement, *hypo* hypo-enhancement

cohort ($p < 0.001$). This significant improvement was demonstrated in the validation cohort (the AUC was improved from 0.634 to 0.731, $p = 0.015$), which demonstrated the incremental value of the radiomics score for MVI estimation. The ROCs of the nomograms and the radiomics score in both cohorts are plotted in Fig. 5. The ROC also showed outstanding discrimination performance of the radiomics nomogram.

Calibration and clinical usefulness

The calibration plots of the radiomics nomogram and clinical nomogram applied in the validation cohort are shown in Fig. 4b and Fig. 4d, respectively. Both the radiomics nomogram and clinical nomogram showed good agreement in detecting the presence of MVI between prediction and histopathologic confirmation based on surgical specimens.

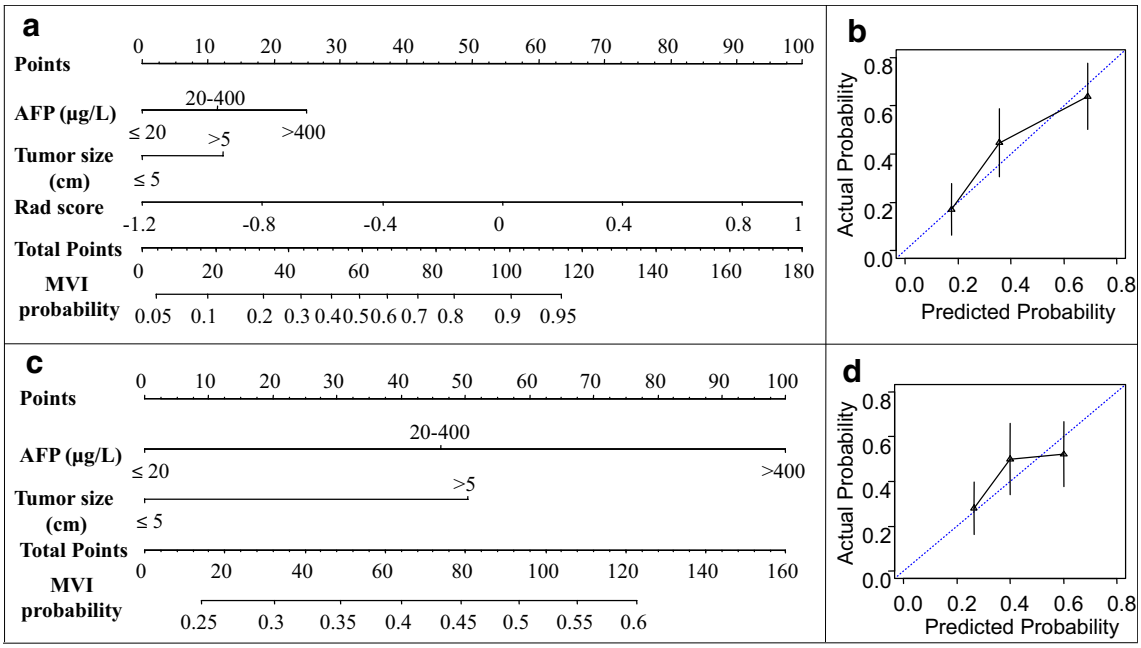

Fig. 4 Nomograms developed from the training cohort (**a**, **c**) and the assessment of the model calibration in the validation cohort (**b**, **d**). Both the nomograms showed good agreement on the presence of MVI between the predicted and actual probability in the validation cohort. **a** Radiomics nomogram incorporated with the AFP, tumor size, and radiomics score. (**b**) Calibration plot for the radiomics nomogram. (**c**) Clinical nomogram incorporated with the AFP and tumor size. (**d**) Calibration plot for the clinical nomogram

The clinical usefulness of the two nomograms was evaluated using DCA (Fig. 6). The DCA curves showed that if the threshold probability was > 20%, using the radiomics nomogram to predict MVI added more benefit for patients than using the clinical nomogram.

Discussion

In the present study, we developed a radiomics score for MVI prediction by using the radiomics technique with US images. The radiomics score was demonstrated to be an independent factor of MVI. When combining the radiomics score with clinical risk factors in the radiomics nomogram, the performance of the clinical nomogram was significantly improved, which demonstrated the incremental value of the radiomics score for individualized MVI prediction in HCC.

Our study demonstrated that an AFP level over 400 ng/mL and a tumor size over 5 cm were also significantly associated with MVI, which was similar to previous reports [27–29]. Prior studies reported that some image features that were observed on MRI or CT were potential predictors for MVI, such as non-

Table 3 AUC of the radiomics score and prediction nomograms

	Training cohort ($n = 341$)		Validation cohort ($n = 141$)	
	AUC (95% CI)	*p* value	AUC (95% CI)	*p* value
Rad-score	0.712 (0.655, 0.768)		0.689 (0.598, 0.781)	
Clin-nom	0.674 (0.617, 0.731)		0.634 (0.543, 0.724)	
Rad-nom	0.758 (0.706, 0.810)		0.731 (0.647, 0.815)	
Rad-score vs. Clin-nom		0.358		0.398
Rad-score vs. Rad-nom		0.027		0.225
Clin-nom vs. Rad-nom		*0.000*		*0.015*

Rad-score radiomics score, *clin-nom* clinical nomogram, combined with AFP and tumor size, *rad-nom* radiomics nomogram, combined with AFP, tumor size and radiomics score

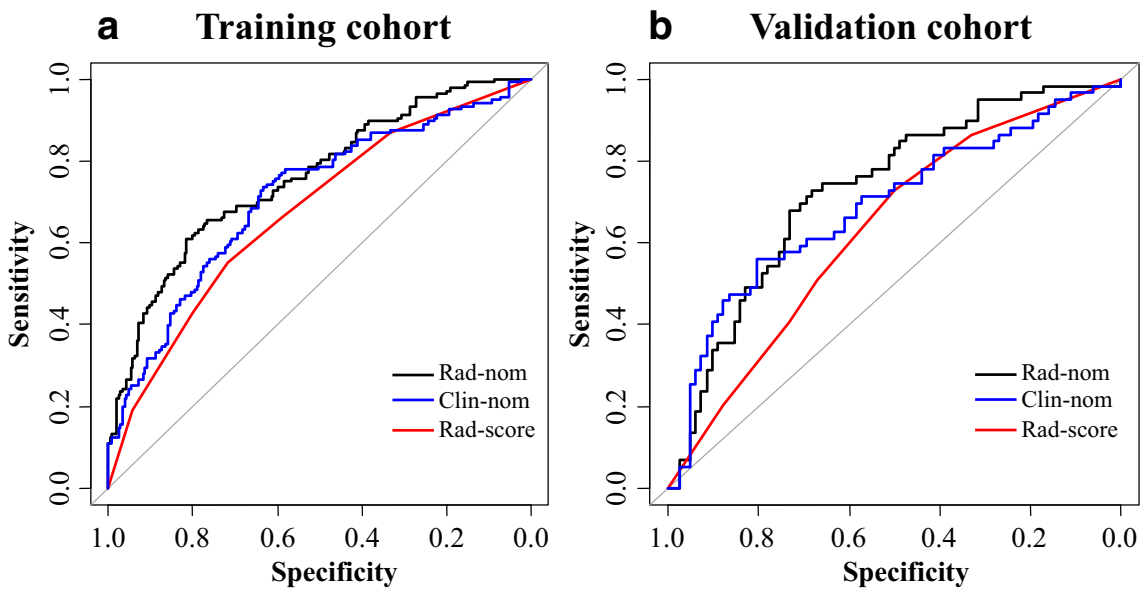

Fig. 5 ROC curves of the radiomics nomogram (black curves), clinical nomogram (red curves), and radiomics score (blue curves) derived from the training (**a**) and validation (**b**) cohorts. Rad-nom, radiomics nomogram; Clin-nom, clinical nomogram; Rad-score, radiomics score

smooth tumor margins and peritumoral enhancement [12–15]. However, some studies carried out negative results [13, 27, 30]. Our study is the first to investigate the image features of US/CEUS for preoperative assessment of MVI. Unfortunately, none of the qualitative CEUS features were independently associated with MVI. Non-smooth tumor margins were also evaluated on US but were not associated with MVI. Given that our study was retrospective, more studies, especially prospective studies, are required to explore the possible relationship between US/CEUS features and the presence of MVI.

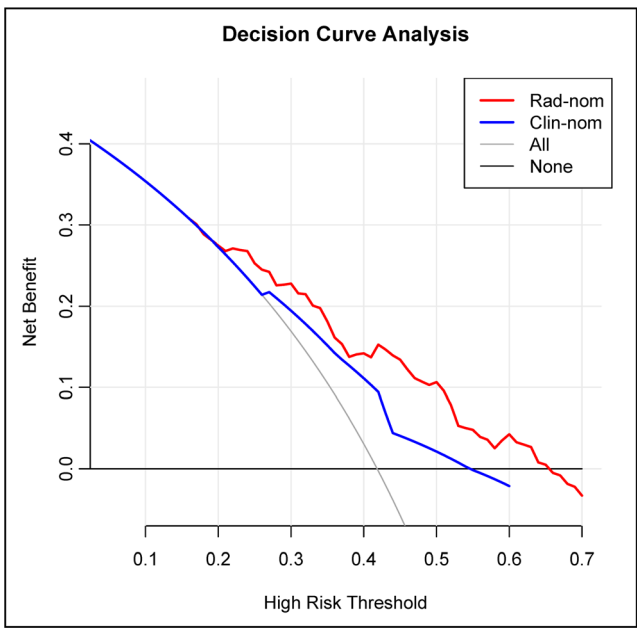

Fig. 6 Decision curve analysis (DCA) derived from the validation cohort. The *y*-axis measures the net benefit. The net benefit is determined by calculating the difference between the expected benefit and the expected harm associated with each proposed model [Net benefit = true positive rate − (false positive rate × weighting factor), weighting factor = threshold probability/(1-threshold probability)]. The gray line represents the assumption that all patients were MVI positive (the treat-all scheme). The black line represents the assumption that all patients were MVI negative (the treat-none scheme). If the threshold probability was > 20%, using the radiomics nomogram (red curve) to predict MVI added more benefit for patients than using the clinical nomogram (blue curve)

Computer-aided techniques have been applied to US imaging for disease diagnosis [31–33]. Radiomics is a computer-aided technique that translates medical imaging information into a series data through computerized algorithms [16, 17]. Previous studies have found that radiomics features were closely related to tumor microscopic structure and biological behavior [34–37]. Radiomics features reflect the texture information of the tumor, which is an important marker of intratumoral heterogeneity. Intratumoral heterogeneity may be the result of genomic heterogeneity, which is associated with a worse prognosis because tumors with more genomic heterogeneity are more likely to metastasize and develop resistance to treatment [38, 39]. Previous studies have suggested that genomic heterogeneity could be identified from medical images [40, 41]. However, the association between biologic processes and radiomics features is still complex [42]. When selecting biomarkers from thousands of radiomics features, it is difficult to clarify the correlation between a single radiomics feature with biological behavior in an intuitive way. The common approach is to build a multi-features parameter for outcome estimation using the radiomics technique [40, 41]. In our study, we extracted radiomics features of lesions from grayscale US images. The radiomics score was built based on the six selected radiomics features, which were indicative of the texture features of the tumor. The radiomics score was

independently associated with MVI, which supported that radiomics has the ability to reflect intratumoral heterogeneity.

A nomogram incorporating multiple risk factors has been used to predict medical outcomes and prognosis. A recent study incorporating the radiomics index into a nomogram along with clinical risk factors performed better in predicting disease-free survival in early-stage non-small cell lung cancer [43]. In the present study, AFP, tumor size, and radiomics score were independently associated with MVI, and the radiomics nomogram incorporating the above three variables achieved good performance for MVI prediction (AUC of 0.731 in validation cohort). When incorporating the radiomics score into the clinical nomogram, the AUC of the clinical nomogram was significantly improved from 0.634 to 0.731. In addition, the DCA showed that more patients would benefit from the radiomics nomogram than the clinical nomogram, which indicated that the radiomics score added incremental value to the clinical risk factors in terms of the clinical usefulness.

Our study has several limitations. First, as a retrospective study, bias was inevitable. In this retrospective study, the US examinations were performed by different radiologists with different US machines, which might lead to heterogeneity of the US images and further impact on the performance of the radiomics score. Although the radiomics score showed stable discrimination performance in the training (AUC = 0.712) and validation (AUC = 0.689) cohorts, both the AUCs were not satisfactory. The operator- and machine-dependence characteristic may be the crucial variable for the dissatisfactory AUCs. Therefore, a prospective longitudinal cohort study with strict control for operator and US machine as well as any other possible confounding variables is desired, which may improve the discrimination performance of the radiomics score. Second, our study was performed in a single institution. Although the radiomics nomogram was assessed in the validation cohort, additional validation in other centers is necessary to evaluate the reliability of our predictive nomogram. Third, as a common limitation of radiomics studies, application of the results requires specific techniques including the feature extracting software and statistical skills. To promote its application in clinical practice, radiomics technique needs to be developed as more simple and user-friendly.

Conclusion

In our study, we developed a radiomics score based on US images from patients with HCC that may be considered a potential biomarker for MVI prediction. The radiomics nomogram, which combined the radiomics score and the clinical risk factors, demonstrated improved performance compared with the clinical nomogram for MVI detection. Therefore, the radiomics nomogram is potentially useful for the preoperative prediction of MVI and allows more appropriate surgical planning before surgery.

Acknowledgements We acknowledge the following two pathologists for reviewing the specimen slices: Bing Liao and Li-li Chen (Department of Pathology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China).

Funding This study was supported by "National Nature Science Foundation of China" (No: 81701701), "Guangdong Natural Science Foundation" (No: 2017A030313661), "Training Project for Young Teacher of Sun Yat-sen University" (No: 16YKPY37), and "Guangdong Science and Technology Foundation" (No: 2017A020215195).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Wei Wang.

Conflict of interest The authors declare that they have no conflict of interest.

Statistics and biometry Not applicable.

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. Forner A, Llovet JM, Bruix J (2012) Hepatocellular carcinoma. Lancet 379:1245–1255
2. Fuks D, Dokmak S, Paradis V, Diouf M, Durand F, Belghiti J (2012) Benefit of initial resection of hepatocellular carcinoma followed by transplantation in case of recurrence: an intention-to-treat analysis. Hepatology 55:132–140
3. Zimmerman MA, Ghobrial RM, Tong MJ et al (2008) Recurrence of hepatocellular carcinoma following liver transplantation: a review of preoperative and postoperative prognostic indicators. Arch Surg 143:182–188 discussion 188
4. Bruix J, Gores GJ, Mazzaferro V (2014) Hepatocellular carcinoma: clinical frontiers and perspectives. Gut 63:844–855
5. Llovet JM, Schwartz M, Mazzaferro V (2005) Resection and liver transplantation for hepatocellular carcinoma. Semin Liver Dis 25: 181–200
6. 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
7. Lim KC, Chow PK, Allen JC et al (2011) Microvascular invasion is a better predictor of tumor recurrence and overall survival following surgical resection for hepatocellular carcinoma compared to the Milan criteria. Ann Surg 254:108–113
8. Pawlik TM, Gleisner AL, Anders RA, Assumpcao L, Maley W, Choti MA (2007) Preoperative assessment of hepatocellular carcinoma tumor grade using needle biopsy: implications for transplant eligibility. Ann Surg 245:435–442
9. Yao FY, Xiao L, Bass NM, Kerlan R, Ascher NL, Roberts JP (2007) Liver transplantation for hepatocellular carcinoma: validation of the

UCSF-expanded criteria based on preoperative imaging. Am J Transplant 7:2587–2596
10. Shindoh J, Andreou A, Aloia TA et al (2013) Microvascular invasion does not predict long-term survival in hepatocellular carcinoma up to 2 cm: reappraisal of the staging system for solitary tumors. Ann Surg Oncol 20:1223–1229
11. Rodriguez-Peralvarez M, Luong TV, Andreana L, Meyer T, Dhillon AP, Burroughs AK (2013) A systematic review of microvascular invasion in hepatocellular carcinoma: diagnostic and prognostic variability. Ann Surg Oncol 20:325–339
12. Chou CT, Chen RC, Lee CW, Ko CJ, Wu HK, Chen YL (2012) Prediction of microvascular invasion of hepatocellular carcinoma by pre-operative CT imaging. Br J Radiol 85:778–783
13. Chou CT, Chen RC, Lin WC, Ko CJ, Chen CB, Chen YL (2014) Prediction of microvascular invasion of hepatocellular carcinoma: preoperative CT and histopathologic correlation. AJR Am J Roentgenol 203:W253–W259
14. Renzulli M, Brocchi S, Cucchetti A et al (2016) Can current preoperative imaging be used to detect microvascular invasion of hepatocellular carcinoma? Radiology 279:432–442
15. Wu TH, Hatano E, Yamanaka K et al (2016) A non-smooth tumor margin on preoperative imaging predicts microvascular invasion of hepatocellular carcinoma. Surg Today. https://doi.org/10.1007/s00595-016-1320-x:1-7
16. Limkin EJ, Sun R, Dercle L et al (2017) Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol. https://doi.org/10.1093/annonc/mdx034
17. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577
18. Esteva A, Kuprel B, Novoa RA et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115–118
19. 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
20. Huang YQ, Liang CH, He L et al (2016) Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol. https://doi.org/10.1200/JCO.2015.65.9128
21. Zhang X, Li J, Shen F, Lau WY (2017) Significance of presence of microvascular invasion in specimens obtained after surgical treatment of hepatocellular carcinoma. J Gastroenterol Hepatol. https://doi.org/10.1111/jgh.13843
22. Claudon M, Dietrich CF, Choi BI et al (2013) Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver–update 2012: a WFUMB-EFSUMB initiative in cooperation with representatives of AFSUMB, AIUM, ASUM, FLAUS and ICUS. Ultraschall Med 34:11–29
23. Wang W, Chen LD, Lu MD et al (2013) Contrast-enhanced ultrasound features of histologically proven focal nodular hyperplasia: diagnostic performance compared with contrast-enhanced CT. Eur Radiol 23:2546–2554
24. Sauerbrei W, Royston P, Binder H (2007) Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Stat Med 26:5512–5528
25. Coutant C, Olivier C, Lambaudie E et al (2009) Comparison of models to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a prospective multicenter study. J Clin Oncol 27:2800–2808
26. Vickers AJ, Cronin AM, Elkin EB, Gonen M (2008) Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak 8:53
27. Lei Z, Li J, Wu D et al (2016) Nomogram for preoperative estimation of microvascular invasion risk in hepatitis B virus-related hepatocellular carcinoma within the Milan criteria. JAMA Surg 151:356–363
28. Zhao WC, Fan LF, Yang N, Zhang HB, Chen BD, Yang GS (2013) Preoperative predictors of microvascular invasion in multinodular hepatocellular carcinoma. Eur J Surg Oncol 39:858–864
29. Kim KA, Kim MJ, Jeon HM et al (2012) Prediction of microvascular invasion of hepatocellular carcinoma: usefulness of peritumoral hypointensity seen on gadoxetate disodium-enhanced hepatobiliary phase images. J Magn Reson Imaging 35:629–634
30. Kornberg A, Freesmeyer M, Barthel E et al (2009) 18F-FDG-uptake of hepatocellular carcinoma on PET predicts microvascular tumor invasion in liver transplant patients. Am J Transplant 9:592–600
31. Shan J, Alam SK, Garra B, Zhang Y, Ahmed T (2016) Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. Ultrasound Med Biol 42:980–988
32. Sugimoto K, Shiraishi J, Tanaka H et al (2016) Computer-aided diagnosis for estimating the malignancy grade of hepatocellular carcinoma using contrast-enhanced ultrasound: an ROC observer study. Liver Int 36:1026–1032
33. Zhang Q, Xiao Y, Dai W et al (2016) Deep learning based classification of breast tumors with shear-wave elastography. Ultrasonics 72:150–157
34. 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
35. Grossmann P, Stringfield O, El-Hachem N et al (2017) Defining the biological basis of radiomic phenotypes in lung cancer. Elife 6
36. 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
37. Guo Y, Hu Y, Qiao M et al (2017) Radiomics analysis on ultrasound for prediction of biologic behavior in breast invasive ductal carcinoma. Clin Breast Cancer. https://doi.org/10.1016/j.clbc.2017.08.002
38. Campbell PJ, Yachida S, Mudie LJ et al (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 467:1109–1113
39. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446
40. Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X (2007) Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. J Vasc Interv Radiol 18:821–831
41. Rutman AM, Kuo MD (2009) Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. Eur J Radiol 70:232–241
42. Tran B, Dancey JE, Kamel-Reid S et al (2012) Cancer genomics: technology, discovery, and translation. J Clin Oncol 30:647–660
43. Huang Y, Liu Z, He L et al (2016) Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. Radiology 281:947–957