BMC Medical Imaging

RESEARCH

Open Access

Ultrasomics differentiation of malignant and benign focal liver lesions based on contrast-enhanced ultrasound

Hang-Tong Hu[1†], Ming-De Li[1†], Jian-Chao Zhang[4], Si-Min Ruan[1], Shan-Shan Wu[3], Xin-Xin Lin[1], Hai-Yu Kang[3], Xiao-Yan Xie[1], Ming-De Lu[1,2], Ming Kuang[1,2], Er-Jiao Xu[3*] and Wei Wang[1*]

Abstract

Objectives To establish a nomogram for differentiating malignant and benign focal liver lesions (FLLs) using ultrasomics features derived from contrast-enhanced ultrasound (CEUS).

Methods 527 patients were retrospectively enrolled. On the training cohort, ultrasomics features were extracted from CEUS and b-mode ultrasound (BUS). Automatic feature selection and model development were performed using the Ultrasomics-Platform software, outputting the corresponding ultrasomics scores. A nomogram based on the ultrasomics scores from artery phase (AP), portal venous phase (PVP) and delayed phase (DP) of CEUS, and clinical factors were established. On the validation cohort, the diagnostic performance of the nomogram was assessed and compared with seniorexpert and resident radiologists.

Results In the training cohort, the AP, PVP and DP scores exhibited better differential performance than BUS score, with area under the curve (AUC) of 84.1-85.1% compared with the BUS (74.6%, $P<0.05$). In the validation cohort, the AUC of combined nomogram and expert was significantly higher than that of the resident (91.4% vs. 89.5% vs. 79.3%, $P<0.05$). The combined nomogram had a comparable sensitivity with the expert and resident (95.2% vs. 98.4% vs. 97.6%), while the expert had a higher specificity than the nomogram and the resident (80.6% vs. 72.2% vs. 61.1%, $P=0.205$).

Conclusions A CEUS ultrasomics based nomogram had an expert level performance in FLL characterization.

Keywords Focal liver lesion, Contrast-enhanced ultrasound, Ultrasomics, Nomogram

†Hang-Tong Hu and Ming-De Li contributed equally to this work.

*Correspondence:
Er-Jiao Xu
xuerjiao@mail.sysu.edu.cn
Wei Wang
wangw73@mail.sysu.edu.cn
[1]Department of Medical Ultrasonics, Institute of Diagnostic and Interventional Ultrasound, Ultrasomics Artificial Intelligence X-Lab, The First Affiliated Hospital of Sun Yat-Sen University, 58 Zhongshan Road 2, Guangzhou 510080, People's Republic of China
[2]Department of Hepatobiliary Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China
[3]Department of Medical Ultrasonics, The Eighth Affiliated Hospital of Sun Yat-sen University, No. 3025, Shennanzhong Road, Shenzhen 518033, PR China
[4]Puer People's Hospital, Puer, China

Introduction

Focal liver lesions (FLLs) are masses or areas of tissue that are considered to be abnormal parts of liver [1]. Accurate differential of the benign and malignant of FLLs was necessary to determine the most effective treatment. In most cases, FLLs are initially detected during conventional ultrasonography (US) examination [2, 3], while it can only provide limited information in definitive diagnosis [1, 4]. Contrast-enhanced ultrasound (CEUS) outperformed US by providing real-time perfusion information of FLL [5]. According to the study of Guang Y et al. [6]. , CEUS shows similar diagnostic value (Area under the receiver-operating characteristic curve, AUC: 94%, sensitive: 88%, specificity: 81%) for the characterizing of a defined FLL, compared with contrast-enhanced computed tomography (CECT) (AUC: 93%, sensitive: 90%, specificity: 77%) and contrast-enhanced magnetic resonance imaging (CEMRI) (AUC: 92%, sensitive: 86%, specificity: 81%).

However, the performance of CEUS depends on the clinicians' experience which is subjective and leads to great intra- and inter-observer variance especially in facing atypical imaging characteristics of lesions [7, 8]. Radiomics is a powerful tool for extracting clinically relevant information from medical images [9–13]. Compared with conventional qualitative visible features extracted by radiologists, radiomics can extract high-throughput quantitative features to reflect inter-tumoral heterogeneity and provide a more accurate diagnosis [14]. Liang et al. reported using a radiomics score model based on US images to discriminate malignancy from benign thyroid nodules. The model showed good discrimination performance with an AUC of 93.1%, which was higher than junior radiologists (AUC: 82.9%) [15]. Wang et al. developed a radiomics model to differentiate uterine sarcoma from atypical leiomyoma based on apparent diffusion coefficient maps of MRI. The radiomics model showed higher diagnostic efficacy than experienced radiologists (AUC: 83.0% vs. 75.2%) [16]. Guo et al. reported a radiomics signature based on MRI showing better performance than a radiology resident in differentiation ocular adnexal lymphoma from idiopathic orbital inflammation (AUC: 73.0% vs. 62.4%, $P<0.05$) [17].

Ultrasomics is the specified application of radiomics in ultrasonography, utilizing ultrasound specified features and selected computerized algorithm [18]. In this study, we intended to develop a nomogram based on ultrasomics for differentiating malignant from benign FLLs on multiphase CEUS and to compare the diagnostic efficacy between radiologists and the nomogram.

Materials and methods

Patients

This study was approved by the Institutional Committee on Ethics (ICE) for Clinical Research and Animal Trials of the First Affiliated Hospital of Sun Yat-sen University (NO. [2015]106). Informed consents of patients to participate in the study and for publication were waived for its retrospective nature. Between January 2014 and September 2015, we retrospectively analyzed 1525 patients referred to our hospital with FLLs. The clinic-pathological data and images were collected from the data system of our hospital. Figure 1 displayed the flow chart of patient recruitment and the inclusion and exclusion criteria. The inclusion criteria were: patients with FLLs underwent CEUS. Exclusion criteria were as follows: (1) treatments were performed before CEUS; (2) cases without a definitive diagnosis or follow up of less than 1 year; (3) cases with missing data of any needed CEUS phases or CEUS recorded for less than 3 min [19]; (4) lesions covered by acoustic shadow for more than 1/3 on the images.

Finally, 527 patients were included. The training cohort consisted of 364 patients obtained from January 2014 to May 2015 (263 men and 101 women, median age 55 years, age range 46–64 years). The validation cohort consisted of 163 patients obtained from June 2015 to September 2015 (132 men and 31 women, median age 56 years, age range 33–80 years). Baseline information and clinical data collection included: age, gender, size, history of hepatitis B or hepatitis C, history of malignancy, hepatocirrhosis and final diagnosis based on pathology or follow-up.

CEUS examination

The following ultrasound equipments were used: (1) Aixplorer Ultrasound system (SuperSonic Imagine, Aix-en-Provence, France) equipped with the SC6-1 convex probe with frequency range of 1.0 to 6.0 MHz. (2) Acuson Sequoia 512 (Siemens Medical Solutions, Mountain View, CA, United States) with a 4V1 vector transducer with frequency range of 1.0 to 4.0 MHz. (3) Aplio SSA-770 or Aplio 500 (Toshiba Medical Systems, Tokyo, Japan) with a 375BT convex transducer with frequency range, of 1.9 to 6.0 MHz. Patients were asked to lie in a supine position and keep relaxation during the whole examination. The entire liver was first scanned with the b-mode US (BUS) and the information of the targeted lesion was recorded (including the echo, largest diameter, location, shape and boundary). Then, the imaging mode was changed to CEUS, and a bolus of 2.4 mL SonoVue (Bracco, Milan, Italy) was injected intravenously. After that, a flush with 5 ml of 0.9% normal saline solution was injected immediately. The targeted lesion was observed for 3–6 min,

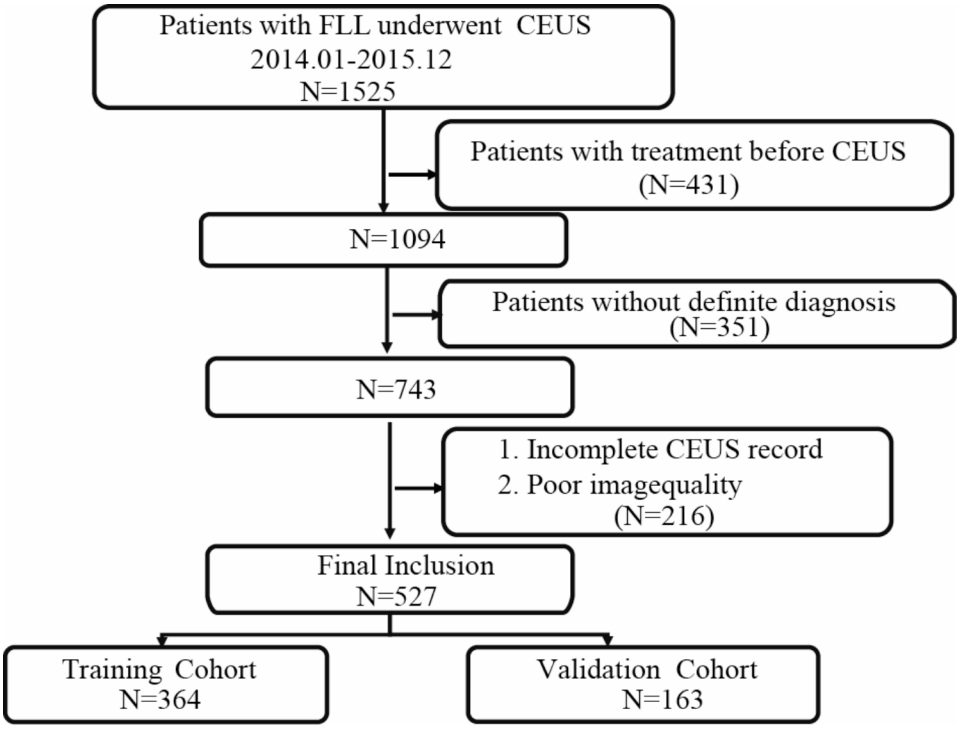

Fig. 1 Flow chart of enrolling the study population

and digital cine clips of artery phase (AP, 10–30 s after injection), portal venous phase (PVP, 31–120 s after injection) and delayed phase (DP, 121–360 s after injection) were stored on the hard disk [19].

Image preparation

First, the BUS image was selected. Then CEUS clips were converted into consecutive frames. And three representative frames of lesion's largest cross section corresponding to the AP, PVP and DP were selected for extracting ultrasomics features. A radiologist with 8 years' experience who was unaware of any clinical data of patients selected all images from study population.

Ultrasomics analysis

Tumor segmentation

The regions of interest (ROIs) were drawn on images of four phases (BUS, as well as AP, PVP and DP of CEUS) for each patient, and were delineated along the targeted lesion margin to cover the entire lesion. Two radiologists (W. Wang and L.D. Chen, with experience of 10 and 8 years respectively) independent delineated ROIs using ITK-SNAP soft (http://www.itksnap.org) in 50 randomly chosen patients, basing on which the inter-observer reproducibility was assessed using interclass correlation coefficient (ICC). And the rest ROIs of images were delineated by one radiologist (W. Wang). The ultrasomics features with ICC > 0.75 were used for further analysis.

Automatic ultrasomics signature construction

Feature extraction, feature selection, ultrasomics score development and evaluation were automatically performed using an in-house designed Ultrasomics-Platform (Version 2.1, Ultrasomics Artificial Intelligence X-lab, Guangzhou). A total of 5936 ultrasomics features of each phase were extracted, including First-order, grey level co-occurrence matrix (GLCM), grey level run-length matrix (GLRLM), gray level size zone matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM), Gray Level Dependence Matrix (GLDM), Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe) and wavelets transform (Supplementary Table 1). After feature normalization with z-score method, spearman correlation, and least absolute shrinkage selection operator (LASSO) were used to reduce the redundancy and select optimal features. According to the ultrasomics features from the four phases respectively, the Ultrasomics-Platform constructed and evaluated internally the ultrasomics scores automatically. Finally, the optimal performing models were output and used to calculate the ultrasomics scores. We constructed 4 single-phase ultrasomics scores (BUS score, AP score, PVP score, and DP score, respectively) from the image of each phase. We also developed a CEUS ultrasomics score which extracted from tri-phases of CEUS images.

Development and validation of the ultrasomics-based nomogram

In the training cohort, univariable and multivariable logistic regression analysis were used to screen out significant factors for FLL discrimination from clinical factors and the 4 single-phase ultrasomics scores. Factors with a *P* value less than 0.05 were included for nomogram construction. A nomogram combined the ultrasomics scores and clinical factors was developed. In the validation cohort, the performance of the nomogram was assessed by the differential efficiency, calibration power, and clinical usefulness.

Imaging analysis by the radiologists

To compare the diagnostic performance between the ultrasomics based nomogram and radiologists, CEUS clip of each patient in the validation cohort was independently reviewed by an expert radiologist (W.W. with experience of 10 years) and a resident radiologist (R.S.M. with experience of 2 years). Both radiologists only knew the patient's clinical information but were unware of the pathological results. The readers were asked to record the enhancement appearances of the lesion, including arterial phase enhancement and its pattern; presence, timing, and degree of washout. For high-risk patients, the lesions were categorized as CEUS LR-1 to LR-5 or LR-M according to American College of Radiology (ACR) CEUS LI-RADS version 2017 [20]. Then, we assigned LR-5, LR-4, LR-M lesions as malignant, and LR-1, LR-2, LR-3 lesions as benign. For non-high-risk patients, the lesions were assigned malignant or benign according to World Federation for Ultrasound in Medicine and Biology- European Federation of Societies for Ultrasound in Medicine and Biology (WFUMB-EFSUMB) criteria [19].

Statistical analysis

Statistical analysis was conducted with SPSS 22.0 for Windows (Chicago, IL) and R software (version 3.4.4, https://www.r-project.org/). Qualitatively, receiver operating characteristic (ROC) curve analysis was plotted to illustrate the discriminatory power of the ultrasomics score, nomogram and radiologists [21]. The calibration curve was plotted to descript the agreement of the nomogram outputs and the observed outcome [22]. The diagnostic efficacy of radiologists and nomogram were compared by sensitivity (SEN), specificity (SPE), positive predictive value (PPV), and negative predictive value (NPV). Estimated values of SEN, SPE of nomogram and radiologists were compared by using the McNemar test, the value of PPV and NPV were compared by Generalized Score Statistic test. Decision curve analysis (DCA) was conducted to compare the clinical usefulness of the nomogram and radiologists at different threshold probabilities [23].

Categorical variables were expressed with numbers or percentages and analyzed with χ^2 test. The continuous variables were expressed with mean ± SD, or median, and analyzed with *t* test for the normal distribution variables or Mann-Whitney U test for the abnormal or unknown distribution variables. A *P* values less than 0.05 was considered as statistically significant using two-sided testing.

Results

Clinic-pathological characteristics

The ratio of patients with malignant to benign FLLs was 77.2%/22.8% (280/84) and 77.9%/22.1% (127/36) in the training and validation cohort ($P=0.517$), respectively. In addition, there were no significant differences between the training cohort and validation cohort in other clinic-pathological characteristics such as hepatitis ($P=0.423$), cirrhosis ($P=0.767$), and size ($P=0.405$). The details of patient characteristics in the training cohort and validation cohort were presented in Table 1.

Ultrasomics score construction

In the training cohort, according to the feature selection with inter-observer ICC > 0.75, 449, 425, 473, and 501 features were selected for the next analysis from the BUS, AP, PVP, and DP, respectively. After optimal feature selection using spearman rank correlation and LASSO regression, 10 features from BUS images were ultimately left for BUS ultrasomics score construction. For CEUS images, 21, 23 and 20 features were left for the construction of AP, PVP and DP scores, respectively (Fig. 2). A detailed description of selected features was provided in Supplementary Tables 2–5.

Using logistic regression, the AP, PVP and DP scores exhibited better differential performance than BUS score, with AUC of 85.3%, 84.1%, 84.3% compared with the BUS (74.6%, $P<0.05$) in the training cohort. The CEUS ultrasomics score established by combining AP, PVP and DP features exhibited the higher AUC of 89.5%, compared with that of each single-phase score ($P_{AP}=0.07$, $P_{PVP}=0.17$, $P_{DP}=0.11$).

Nomogram construction and validation

In the training cohort, univariable and multivariable logistic regression results (Table 2) showed gender, age, AP score, PVP score, DP score, history of chronic hepatitis and history of malignancy were the significant predictive factors associated with classification of malignant and benign FLLs ($P<0.05$). The combined nomogram was developed with AP, PVP, DP scores, gender, age, history of chronic hepatitis and malignant

Table 1 Clinic-pathologic characteristics and focal liver lesions types in the training and validation cohorts

	Training cohort (n = 364)	Validation cohort (n = 163)	*P*
Gender (male/female)(n)	263/101	132/31	0.051
Age, n(%)			
< 40	57(15.6)	32(19.6)	0.123
40–60	188(51.6)	87(49.6)	0.393
> 60	119(32.6)	44(26.9)	0.157
Size	5.84 ± 3.89	5.56 ± 3.49	0.405
Viral hepatitis(with/without)(n)	257/107	121/42	0.423
Cirrhosis (yes/no)(n)	110/254	46/117	0.767
History of Malignant (yes/no)(n)	25/339	11/152	0.877
Malignant/Benign(n)	280/84	127/36	0.517
Malignant, n(%)			
HCC	239(65.6)	106(65.0)	0.921
ICC	13(3.5)	7(4.2)	0.688
Combined HCC and ICC	3(0.8)	2(1.2)	0.659
Hepatosarcoma	2(0.5)	0(0)	0.343
Neuroendocrine tumor	2(0.5)	1(0.6)	0.928
Metastasis	21(5.7)	11(6.7)	0.669
Benign lesion,n(%)			
Hemangioma	35(9.6)	17(10.4)	0.897
Focal nodular hyperplasia	12(3.2)	7(4.3)	0.615
Abscess	10(2.7)	3(1.8)	0.843
Nodule regenerative hyperplasia	9(2.4)	4(2.4)	0.986
Angioleiomyolipoma	7(1.9)	2(1.2)	0.566
Adenoma	6(1.6)	2(1.2)	0.712
Benign indeterminate	5(1.3)	1(0.6)	0.447
Radiomics score[median (IQR)]			
CEUS-Arterial phase	1.53[0.95,2.08]	1.47[1.01,2.15]	0.921
CEUS-Portal vein phase	1.44[1.01,1.58]	1.46[0.90,2.00]	0.964
CEUS-Delay phase	1.41[0.83,2.11]	1.49[0.81,2.16]	0.764
Baseline US	1.28[1.03,1.54]	1.27[0.99,1.53]	0.809

Unless otherwise indicated, data are the number of nodules, with percentages in parentheses. Abbreviations: CEUS: Contrast-enhanced ultrasound; HCC: Hepatocellular carcinoma; ICC: Intrahepatic cholangiocarcinoma; IQR: Interquartile range

history (Fig. 3). In both cohorts the combined nomogram exhibited best diagnostic performance, which significantly improved compared with 4 single-phase score and CEUS score in validation cohort (AUC 91.4% vs. 75.9%, 76.2%, 75.1%, 66.1%, 81.4%, $P<0.001$) (Table 3; Fig. 4).

In the training and validation cohort, the calibration curves indicated that the combined nomogram showed good agreement between the nomogram and final diagnosis.

Comparison of diagnostic performance between combined nomogram and radiologists

In the validation sets, the ROC analysis demonstrated that the AUC of the combined nomogram and expert were 91.4% and 89.5%, respectively, which was significantly (Delong test $P<0.05$) higher than that of the resident (AUC: 79.3%). (Table 4; Fig. 4). The sensitivity, specificity, PPV, and NPV of the combined nomogram were calculated according to the Youden cutoff to quantize the predictive ability of the predictive models in validation cohorts. Under the optimal cutoff (0.683), the sensitivity, specificity, PPV, and NPV of the combined nomogram achieved 95.3%, 72.2%, 92.36%, 81.3%, respectively (Table 5).

The expert yielded the highest sensitivity (98.4%), specificity (80.6%), NPV (94.7%), PPV (93.5%) in validation cohort. Among the expert, combined nomogram and resident, there were no significant differences in sensitivity, PPV and NPV in validation cohort. However, the specificity of the expert (80.6%) was significantly higher than the resident (61.1%, $P=0.01$). Although there was no significant difference ($P=0.21$) between combined nomogram (72.2%) and resident (61.1%) in specificity, a tendency towards strikingly higher specificity was observerd for combined nomogram compared to the resident (Table 5).

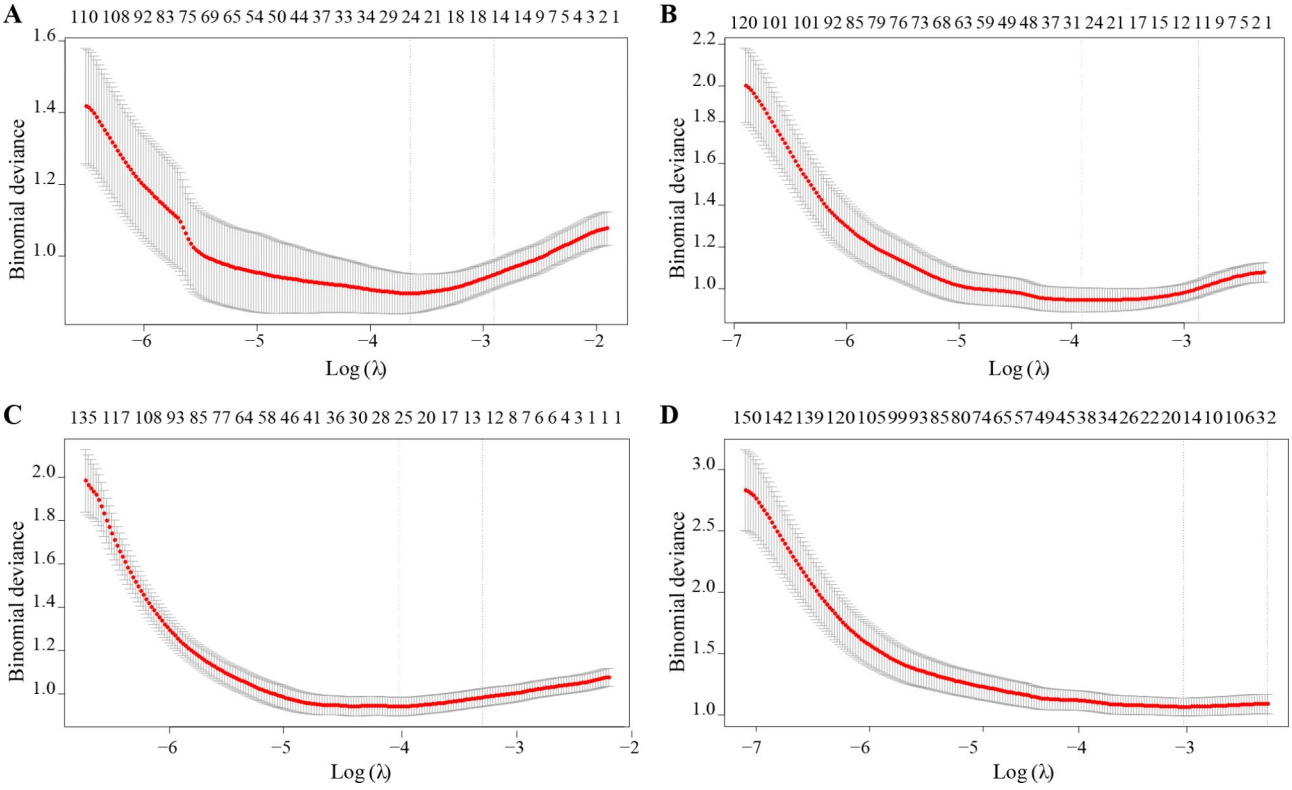

Fig. 2 Feature selection using the spearman and least absolute shrinkage and selection operator (LASSO) method. Finally, the classification ability of the ultrasomics scores was evaluated by the receiver-operating characteristics (ROC) curves obtained in both the training and validation cohort. Feature selection process using LASSO regression for BUS (A), AP (B), PVP (C) and DP (D) scores, respectively

Table 2 The results of the univariate and multivariate analyses based on the training set

Factors	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Gender	4.319(2.572,7.032)	< 0.01*	4.944(2.041, 12.583)	< 0.01*
Age	2.816(1.587,5.257)	< 0.01*	4.433(1.67, 13.081)	0.012*
Cirrhosis	5.260(2.577,12.267)	< 0.01*	1.305(0.467, 3.362)	0.616
History of hepatitis	16.967(9.498,36.431)	< 0.01*	30.348(12.438, 83.231)	< 0.01*
History of malignancy	7.657(1.581,137.945)	< 0.01*	262.536(12.042, 3.141 × 104)	0.011*
AP score	5.617(3.785,8.784)	< 0.01*	3.791(2.036, 7.587)	< 0.01*
PVP score	7.983(4.911,13.741)	< 0.01*	1.492(0.640, 3.586)	0.047*
DP score	5.949(3.886,9.663)	< 0.01*	2.658(1.191, 4.438)	< 0.01*
BUS score	11.314(5.489,24.835)	< 0.01*	1.001(0.841, 9.456)	0.089

The significant factors at the *P* > 0.10 level in the univariate analysis were selected to perform the multivariate logistic regression analysis. In the multivariate analysis, *P* values below 0.05 (indicated with asterisk) were considered significant

AP: arterial phase: PVP: portal vein phase: DP: delay phase: BUS: baseline ultrasound

In the high-risk and non-high-risk subgroups, the AUC of combined nomogram (89.8%, 88.2%) and expert (87.1%, 88.7%) was significantly higher than that of the resident (74.5%, 77.4%, $P < 0.001$) (Table 4). The high-risk and non-high-risk subgroups analyses indicated that expert still had the highest sensitivity (99.2%, 94.2%), specificity (75.0%, 83.3%), NPV (90.0%, 95.2%), PPV (97.3%, 80.0%). The specificity of expert (75.0%) was significantly higher than the combined nomogram (50.0%, $P = 0.04$) and resident (50.0%, $P = 0.04$) in the high-risk group. In the non-high-risk group, the specificity of expert (83.3%) and combined nomogram (83.3%) was significantly higher than the resident (66.3%, $P = 0.05$) (Table 5).

The DCA curves (Fig. 5) showed that the expert had the highest net benefit. If the threshold probability was setted as over 70%, using the combined nomogram added 0.012 net benefit (NB) than the resident for discriminating malignant from benign FLLs in validation [23]. When the threshold probability was setted

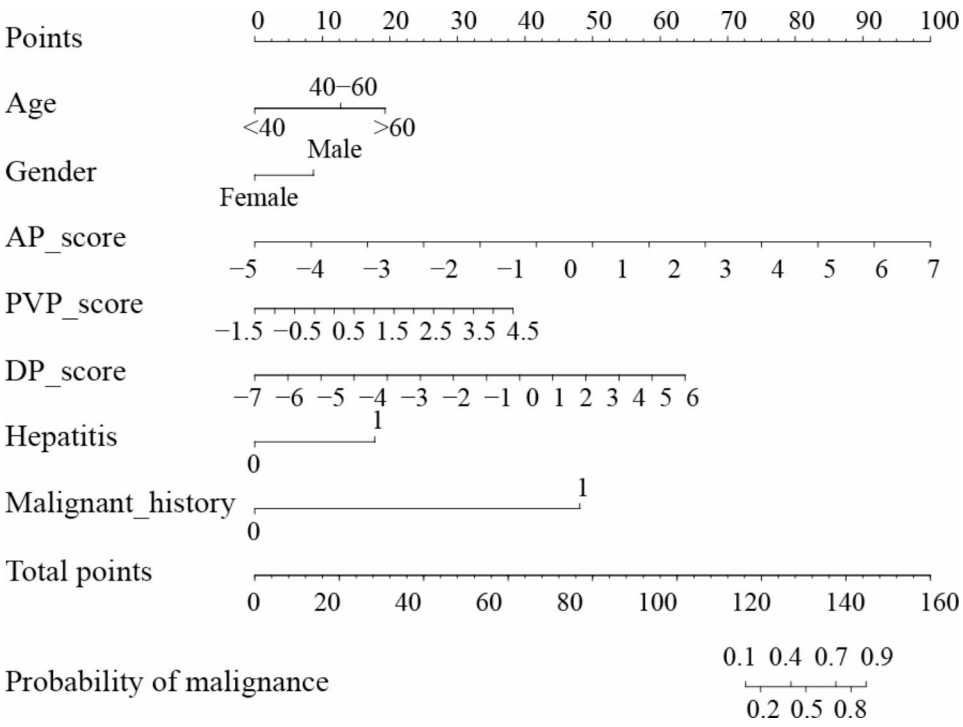

Fig. 3 Nomograms developed from the training cohort. The combined nomogram incorporated the artery phase (AP), portal venous phase (PVP) and delayed phase (DP) ultrasomics scores and independent clinical factor (CF)

Table 3 The diagnostic performance of the single phase ultrasomics score, tri-phases contrast-enhanced ultrasound score and combined nomogram based on training and validation cohort

	Training (*n* = 364)			Validation (*n* = 163)		
	AUC	**95%CI**	***P***	**AUC**	**95%CI**	***P***
AP_Score	0.853	0.810–0.906	< 0.001*	0.759	0.660–0.859	<0.001*
PVP_Score	0.841	0.799–0.891	< 0.001*	0.762	0.675–0.849	<0.001*
DP_Score	0.843	0.794–0.892	< 0.001*	0.751	0.650–0.851	<0.001*
BUS_Score	0.746	0.684–0.813	< 0.001*	0.661	0.563–0.760	<0.001*
CEUS Score	0.895	0.851–0.933	< 0.001*	0.814	0.725–0.904	<0.001*
Combined Nomogram	0.972	0.957–0.987	-	0.914	0.859–0.968	-

P value reference of the combined nomogram. AP: arterial phase; PVP: portal venous phase; DP: delay phase; BUS: baseline ultrasound; AUC: area under the curve; CEUS: contrast enhanced ultrasound

as over 19%, using the combined nomogram added net benefit (NB) than resident for discriminating malignant from benign FLLs in non-high-risk subgroup.

Discussion

In the present study, we employed CEUS and BUS images to investigate the diagnostic performance of ultrasomics features in discriminating malignancy from benign FLLs. we developed four single-phase ultrasomics scores based on CEUS/BUS images. AP,

Table 4 Compared of the area under the receiver-operating characteristics curve between combined nomogram, expert radiologist and resident radiologist in distinguishing malignant from benign lesions in validation cohort

	Validation cohort (*n* = 163)			High-Risk (*n* = 122)			Non-high-Risk (*n* = 41)		
	AUC	**95%CI**	***P***	**AUC**	**95%CI**	***P***	**AUC**	**95%CI**	***P***
Combined nomogram	0.914	0.859–0.968	-	0.898	0.815–0.982	-	0.882	0.782–0.982	-
Expert	0.895	0.829–0.961	0.606	0.871	0.742–0.999	0.594	0.887	0.792–0.983	0.941
Resident	0.793	0.712–0.876	0.009*	0.745	0.597–0.894	0.034*	0.774	0.650–0.899	0.036*
Expert vs. Resident			0.031*			0.056			0.022*

P values were combined nomogram vs. expert and resident, or expert vs. resident

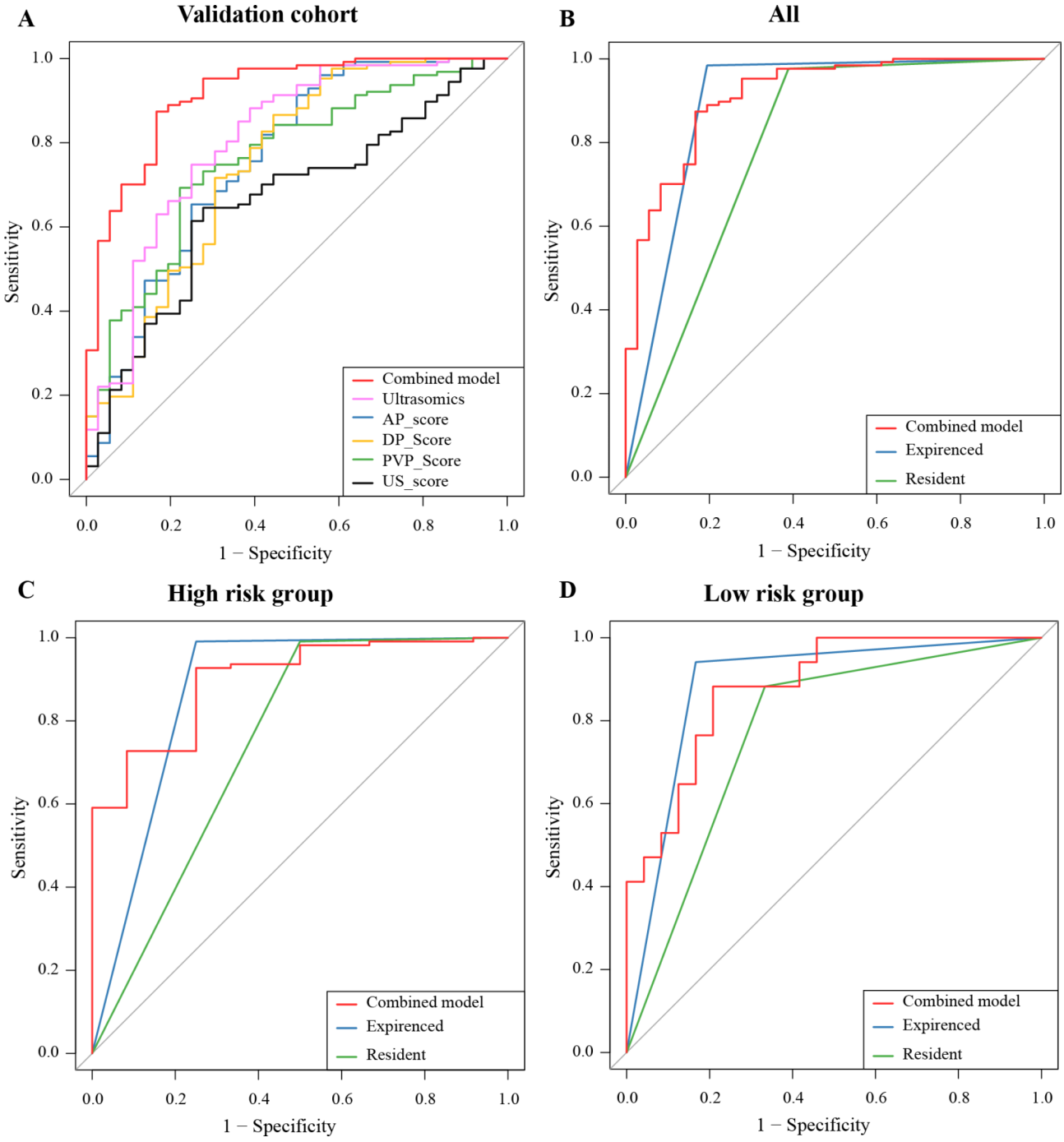

Fig. 4 Receiver-operating characteristics (ROC) curves of the combined nomogram (red curves), contrast-enhanced ultrasound (CEUS) ultrasomics score (orchid curves), artery phase (AP) score (blue curves), portal venous phase (PVP) score (green curves), delayed phase (DP) score (goldenrod curves), and baseline ultrasound (BUS) score (black curves) derived from the validation cohorts (**A**). ROC curves of the combined nomogram (red curves), expert radiologist (blue curves) and resident radiologist (green curves) derived from the total validation cohort (**B**), high-risk subgroup (**C**) and non-high-risk subgroup (**D**) in validation cohort

PVP, DP ultrasomics scores were demonstrated to be independent factors for differentiating malignant and benign FLLs. The combined nomogram integrated ultrasomics scores (AP, PVP, DP) and clinical factors significantly improved the diagnostic performance compared with CEUS scorein training and validation cohort. Our results also showed that the diagnostic performance of the combined nomogram was comparable with the expert radiologist, which was significantly better than the resident radiologist.

Table 5 Comparison of the diagnostic performance of the combined nomogram, expert radiologist and resident radiologist in distinguishing malignant from benign lesions in validation cohort

	SEN	95%CI	*P*	SPE	95%CI	*P*	PPV	95%CI	*P*	NPV	95%CI	*P*
Validation cohort ($n=163$)												
Combined nomogram	95.27%	91.50–98.9%		72.22%	57.60–86.80%		92.36%	87.81–96.90%		81.25%	67.70–94.70%	
Expert	98.42%	96.10–100%	0.157	80.55%	67.20–93.10%	0.256	94.69%	90.21–98.32%	0.221	93.54%	85.00-100%	0.135
Resident	97.63%	94.30–100%	0.256	61.11%	45.20–77.10%	0.205	89.82%	84.81–94.83%	0.243	88.13%	75.2–100%	0.393
Expert vs. Resident			0.563			0.008*			0.071			0.365
High-Risk ($n=122$)												
Combined nomogram	98.13%	95.60–100%		50.00%	21.70–78.20%		94.73%	90.62–98.81%		75.00%	45.10–100%	
Expert	99.15%	97.30–100%	0.563	75.00%	50.10–99.20%	0.035*	97.31%	94.22–100%	0.401	90.00%	71.40–100%	0.081
Resident	99.12%	97.30–100%	0.317	50.00%	21.70–78.20%	1.000	94.82%	90.73–98.81%	0.968	85.71%	59.10–100%	0.343
Expert vs. Resident			1.000			0.035*			0.083			0.784
Non-high-Risk ($n=41$)												
Combined nomogram	76.42%	56.10–96.60%		83.33%	68.40–98.20%		76.41%	56.32–96.63%		83.33%	68.40–98.20%	
Expert	94.15%	82.90–100%	0.179	83.33%	68.40–98.20%	1	80.00%	62.42–97.55%	0.181	95.23%	86.10–100%	0.696
Resident	88.23%	72.90–100%	0.414	66.32%	47.30–85.50%	0.045*	65.22%	45.71–84.62%	0.282	88.88%	74.30–100%	0.581
Expert vs. Resident			0.317			0.045*			0.033*			0.242

P values were combined nomogram vs. expert and resident, or expert vs. resident. SEN: sensitive; SPE: specificity; PPV: Positive predictive value; NPV: Negative predictive value

Widely applied in Asian countries, CEUS was considered an ideal imaging modality for the differentiation of benign and malignant FLLs [24, 25]. Compared with BUS, CEUS exhibited higher accuracy in the differentiating malignant from benign FLLs [5]. Similar results were presented in our study, the diagnostic efficacy of ultrasomics scores based on AP, PVP, DP contrast images were superior to that based on BUS image. It meant that the contrast-enhanced images may contain more information reflecting the heterogeneity of benign and malignant FLLs. In the daily process of diagnosis, radiologists usually integrated the performance of three-phase contrast imaging to make a diagnosis. Therefore, we had established a CEUS score by combining three contrast phases scores and validated in validation cohort. The results showed that the diagnostic efficiency of the CEUS scorelightly improved than single-phase score.

Previous studies have shown that patients with older age, men, a history of chronic hepatitis, and a history of malignant tumors have a higher risk of developing malignant liver tumors than general patients [26, 27], and similar results were obtained in our study. A diagnosis based on a multivariable model was considered a more reliable way, as the way that doctors naturally integrated patient clinical information and image features to make an optimal diagnosis. In present study, patient's age, gender, history of chronic hepatitis and history of malignancy were independent clinical factors associated with FLLs discrimination. We integrated above clinical risk factors and three ultrasomics scores (AP, PVP, DP) into a combined nomogram, the nomogram had a significantly higher classification performance than CEUS score(AUC: 91.4% vs. 81.4%, $P<0.001$). The result showed that clinical factors can provide complementary information to ultrasomics score.

Compared with the features based on visual analysis, a lot of image information such as the intensity features, texture features and wavelet decomposition forms can be captured by using computer analysis [28, 29]. Many studies reported that the image information can reflect genomic and phenotypic information of the tumors [9, 30]. Radiomics facilitates the integration of multiparametric ultrasound (including b-mode ultrasound, color Doppler, elastography, CEUS, etc.), enabling the extraction of quantitative features from medical images and videos, and facilitating a comprehensive analysis of image data [31]. We compared the diagnostic efficacy of the combined nomogram with an expert radiologist and a resident radiologist. According to results, combined nomogram and the expert radiologist had higher diagnostic efficiency than the resident in validation cohort and non-high-risk subgroup.

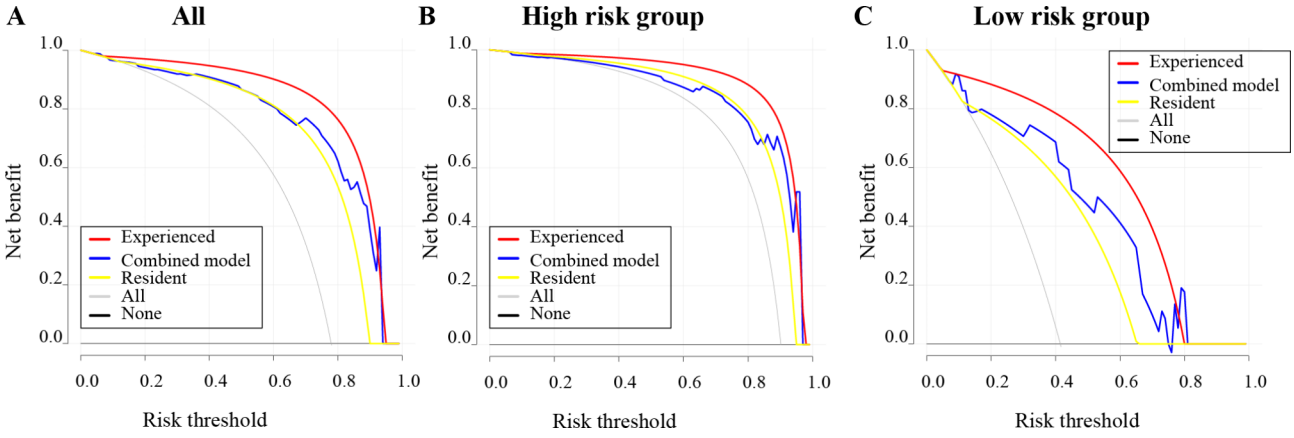

Fig. 5 Decision curve analysis for the combined nomogram (blue curve), expert radiologist (red curves) and resident radiologist (yellow curves) in the total validation cohort (**A**), high-risk subgroup (**B**) and non-high-risk subgroup (**C**) in validation cohort. The y-axis represents the net benefit, and the x-axis represents the threshold probability

In non-high-risk subgroup, combined nomogram had remarkable higher specificity than resident radiologist ($P=0.05$), but the sensitive was rather low ($P=0.41$). It illustrated the resident radiologist tended to misdiagnose the benign FLL in non-high-risk patients which may increase the burden on patients. In the high-risk group, the combined nomogram had a lower specificity, partly because of the fewer benign patients in the high-risk group. The decision curves analysis demonstrated that the combined nomogram added clinical benefits to the resident with higher diagnostic efficiency.

The present study has several limitations. First, as a retrospective study, it was inevitable to raise selection bias. Although we used random grouping and ten-fold cross-validation to reduce the risk, a prospective study is warranted. Second, reproducibility was a common deficiency in current radiomics researches, and the generalizability of reported findings was limited. Although we calculated ICC values of the ultrasomics features and selected features with ICC values greater than 0.75 to develop our models, different US machines and parameter settings may affect image features. The reliability and stability of our nomogram needed further validation with multi-center datasets. Thirdly, reproducibility poses a significant challenge in ultrasomics, particularly due to the operator-dependent nature of ultrasound. Operator experience not only influences ROI delineation but also impacts the selection of representative frames from CEUS videos and the acquisition of ultrasound images during CEUS examination. In this study, all ROIs were manually drawn by 2 radiologists; although good inter-observer agreement was reached, automatic segmentation would provide a more objective assessment and also save time [32]. And future studies on bridging data heterogeneity in ultrasound is important. Specifically, we acknowledged that by computing ultrasomics scores separately for each CEUS phase, we may have overlooked the crucial sequential diagnostic information provided by CEUS. This oversight could potentially limit the comprehensive understanding of lesion characteristics. We recognize the importance of future studies that focus on extracting CEUS's sequential diagnostic information, as it would be invaluable for improving FLL differentiation.

In conclusion, the nomogram based on contrast-enhanced ultrasomics and clinical factors had a good performance of classifying malignancy and benign FLLs. A CEUS ultrasomics based nomogram had an expert level performance in FLL characterization. Ultrasomics could be a useful and practicable tool for FLL diagnosis, especially for less experienced sonographers. Further validation in future studies would be needed.

Abbreviations

AP	Artery phase
AUC	Area under the curve
BUS	B-mode ultrasound
CECT	Contrast-enhanced computed tomography
CEMRI	Contrast-enhanced magnetic resonance imaging
CEUS	Contrast-enhanced ultrasound
CoLIAGe	Co-occurrence of Local Anisotropic Gradient Orientations
DP	Delayed phase
FLL	Focal liver lesion
GLCM	Grey level co-occurrence matrix
GLDM	Gray Level Dependence Matrix
GLRLM	Grey level run-length matrix
GLSZM	Gray level size zone matrix
ICC	Interclass correlation coefficient
LASSO	Least absolute shrinkage selection operator
NB	Net benefit
NGTDM	Neighbouring Gray Tone Difference Matrix
NPV	Negative predictive value
PPV	Positive predictive value
PVP	Portal venous phase
ROC	Receiver operating characteristic
ROI	Region of interest
SEN	Sensitivity

SPE	Specificity
US	Ultrasonography

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12880-024-01426-x.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
Guarantors of integrity of entire study, W.W.; study concepts/study design or data acquisition or data analysis/interpretation, E.J.X., W.W.; manuscript drafting for important intellectual content, H.T.H., M.D.L.(Ming-De Li); approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, X.Y.X, M.D.L.(Ming-De Lu), M.K.; literature research, J.C.Z, S.S.W.; clinical studies, S.M.R, W.W.; statistical analysis, H.T.H., M.D.L.(Ming-De Li); article submission, H.Y.K., X.X.L., and manuscript editing, H.T.H., M.D.L.(Ming-De Li), E.J.X., W.W.

Funding
This work was supported by the National Natural Science Foundation of China (No: 82102141, 82171960, 81971630), Natural Science Foundation of Guangdong Province (NO: 2021B1515020054), and Guangzhou Basic and Applied Basic Research Project (No: 2023A04J2231).

Data availability
The datasets generated and analyzed in this study are available upon reasonable request by the corresponding author.

Declarations

Ethics approval and consent to participate
This study was approved by the ICE for Clinical Research and Animal Trials of the First Affiliated Hospital of Sun Yat-sen University (NO. [2015]106). While for the retrospective nature of this study, informed consents for participating in this study was waived by the Institutional Review Board of the First Affiliated Hospital of Sun Yat-sen University. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 29 November 2022 / Accepted: 11 September 2024
Published online: 16 September 2024

References

1. Marrero JA, Ahn J, Rajender Reddy K. and G. Americal College of, ACG clinical guideline: the diagnosis and management of focal liver lesions. Am J Gastroenterol. 2014;109(9):1328–47. quiz 1348.
2. Segura Grau A, Valero López I, Díaz N, Rodríguez, Segura JM, Cabral. Liver ultrasound: focal lesions and diffuse diseases. Semergen. 2016;42(5):307–14.
3. Rungsinaporn K, Phaisakamas T. Frequency of abnormalities detected by upper abdominal ultrasound. J Med Assoc Thai. 2008;91(7):1072–5.
4. D'Onofrio M, Crosara S, De Robertis R, Canestrini S, Mucelli RP. Contrast-enhanced ultrasound of focal liver lesions. AJR Am J Roentgenol. 2015;205(1):W56–66.
5. Quaia E, Stacul F, Gaiani S, Ricci P, Passariello R, Curzio D, Pozzi Mucelli R. Comparison of diagnostic performance of unenhanced vs SonoVue - enhanced ultrasonography in focal liver lesions characterization. The experience of three Italian centers, Radiol Med. 2004;108(1–2):71–81.
6. Guang Y, Xie L, Ding H, Cai A, Huang Y. Diagnosis value of focal liver lesions with SonoVue®-enhanced ultrasound compared with contrast-enhanced computed tomography and contrast-enhanced MRI: a meta-analysis. J Cancer Res Clin Oncol. 2011;137(11):1595–605.
7. Li W, Wang W, Liu GJ, Chen LD, Wang Z, Huang Y, Liu JY, Xie XY, Lu MD. Differentiation of atypical hepatocellular carcinoma from focal nodular hyperplasia: diagnostic performance of contrast-enhanced US and microflow imaging. Radiology. 2015;275(3):870–9.
8. Quaia E, Alaimo V, Baratella E, Pizzolato R, Cester G, Medeot A, Cova MA. Effect of observer experience in the differentiation between benign and malignant liver tumors after ultrasound contrast agent injection. J Ultrasound Med. 2010;29(1):25–36.
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77.
10. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. Radiology. 2021;298(3):505–516.
11. Zhong L, Shi L, Zhou L, Liu X, Gu L, Bai W. Development of a nomogram-based model combining intra- and peritumoral ultrasound radiomics with clinical features for differentiating benign from malignant in breast imaging reporting and data system category 3–5 nodules. Quant Imaging Med Surg. 2023;13(10):6899–910.
12. Hu Y, Li A, Zhao CK, Ye XH, Peng XJ, Wang PP, Shu H, Yao QY, Liu W, Liu YY, Lv WZ, Xu HX. A multiparametric clinic-ultrasomics nomogram for predicting extremity soft-tissue tumor malignancy: a combined retrospective and prospective bicentric study. Radiol Med. 2023;128(6):784–797.
13. Jin J, Jiang Y, Zhao YL, Huang PT. Radiomics-based machine learning to predict the recurrence of hepatocellular carcinoma: a systematic review and meta-analysis. Acad Radiol. 2024;31(2):467–479.
14. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuze S, Schernberg A, Paragios N, Deutsch E, Ferte C. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol. 2017;28(6):1191–206.
15. Liang J, Huang X, Hu H, Liu Y, Zhou Q, Cao Q, Wang W, Liu B, Zheng Y, Li X, Xie X, Lu M, Peng S, Liu L, Xiao H. Predicting malignancy in thyroid nodules: radiomics score versus 2017 American College of Radiology Thyroid Imaging, Reporting and Data System. Thyroid. 2018;28(8):1024–1033.
16. Xie H, Hu J, Zhang X, Ma S, Liu Y, Wang X. Preliminary utilization of radiomics in differentiating uterine sarcoma from atypical leiomyoma: Comparison on diagnostic efficacy of MRI features and radiomic features. Eur J Radiol. 2019;115:39–45.
17. Guo J, Liu Z, Shen C, Li Z, Yan F, Tian J, Xian J. MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation. Eur Radiol. 2018;28(9):3872–3881.
18. Li W, Huang Y, Zhuang BW, Liu GJ, Hu HT, Li X, Liang JY, Wang Z, Huang XW, Zhang CQ, Ruan SM, Xie XY, Kuang M, Lu MD, Chen LD, Wang W. Multiparametric ultrasomics of significant liver fibrosis: A machine learning-based analysis. Eur Radiol. 2019;29(3):1496–1506.
19. Claudon M, Dietrich CF, Choi BI, Cosgrove DO, Kudo M, Nolsoe CP, Piscaglia F, Wilson SR, Barr RG, Chammas MC, Chaubal NG, Chen MH, Clevert DA, Correas JM, Ding H, Forsberg F, Fowlkes JB, Gibson RN, Goldberg BB, Lassau N, Leen EL, Mattrey RF, Moriyasu F, Solbiati L, Weskott HP, Xu HX. 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. 2013;34(1):11–29.
20. Kono Y, Lyshchik A, Cosgrove D, Dietrich CF, Jang HJ, Kim TK, Piscaglia F, Willmann JK, Wilson SR, Santillan C, Kambadakone A, Mitchell D, Vezeridis A, Sirlin CB. Contrast Enhanced Ultrasound (CEUS) Liver imaging reporting and data system (LI-RADS(R)): the official version by the American College of Radiology (ACR). Ultraschall Med. 2017;38(1):85–86.
21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45.
22. Coutant C, Olivier C, Lambaudie E, Fondrinier E, Marchal F, Guillemin F, Seince N, Thomas V, Leveque J, Barranger E, Darai E, Uzan S, Houvenaeghel G, Rouzier R. 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. 2009;27(17):2800–8.
23. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565–74.

24. Jo PC, Jang HJ, Burns PN, Burak KW, Kim TK, Wilson SR. Integration of contrast-enhanced us into a multimodality approach to imaging of nodules in a cirrhotic liver: how I do it. Radiology. 2017;282(2):317–331.
25. Sporea I, Badea R, Popescu A, Sparchez Z, Sirli RL, Danila M, Sandulescu L, Bota S, Calescu DP, Nedelcu D, Brisc C, Ciobaca L, Gheorghe L, Socaciu M, Martie A, Ioanitescu S, Tamas A, Streba CT, Iordache M, Simionov I, Jinga M, Anghel A, Cijevschi Prelipcean C, Mihai C, Stanciu SM, Stoicescu D, Dumitru E, Pietrareanu C, Bartos D, Manzat Saplacan R, Parvulescu I, Vadan R, Smira G, Tuta L, Saftoiu A. Contrast-enhanced ultrasound (CEUS) for the evaluation of focal liver lesions - a prospective multicenter study of its usefulness in clinical practice. Ultraschall Med. 2014;35(3):259–66.
26. Chen WQ, Sun KX, Zheng RS, Zeng HM, Zhang SW, Xia CF, Yang ZX, Li H, Zou XN, He J. Cancer incidence and mortality in China, 2014. Chinese Journal of Cancer Research. 2018;30(1):1–12.
27. Tang A, Hallouch O, Chernyak V, Kamaya A, Sirlin CB. Epidemiology of hepatocellular carcinoma: target population for surveillance and diagnosis. Abdom Radiol (NY). 2018;43(1):13–25.
28. Wu ML, Li L, Wang JH, Zhang YY, Guo Q, Li X, Zhang XN. Contrast-enhanced US for characterization of focal liver lesions: a comprehensive meta-analysis. Eur Radiol. 2018;28(5):2077–2088.
29. Schellhaas B, Waldner MJ, Gortz RS, Vitali F, Kielisch C, Pfeifer L, Strobel D, Janka R, Neurath MF, Wildner D. Diagnostic accuracy and interobserver variability of dynamic vascular pattern (DVP) in primary liver malignancies - a simple semiquantitative tool for the analysis of contrast enhancement patterns. Clin Hemorheol Microcirc. 2017;66(4):317–31.
30. Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. Eur J Radiol. 2009;70(2):232–41.
31. Zhang H, Meng Z, Ru J, Meng Y, Wang K. Application and prospects of AI-based radiomics in ultrasound diagnosis. Vis Comput Ind Biomed Art. 2023;6(1):20.
32. Gross M, Arora S, Huber S, Kücükkaya AS, Onofrey JA. LiverHccSeg: a publicly available multiphasic MRI dataset with liver and HCC tumor segmentations and inter-rater agreement analysis. Data Brief. 2023;51:109662.

Publisher's note