European Journal of Radiology 120 (2019) 108685

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

The interreader agreement and validation of contrast-enhanced ultrasound liver imaging reporting and data system

JiaWu Li, WenWu Ling, Shuang Chen, Lin Ma, Lulu Yang, Qiang Lu, Yan Luo*

Department of Ultrasound, West China Hospital of Sichuan University, Chengdu 610041, Sichuan Province, China

ARTICLE INFO

Keywords:
Contrast-enhanced ultrasound
Liver imaging reporting and data system
Liver
Hepatocellular carcinoma

ABSTRACT

Purpose: This study explored the interreader agreement and diagnostic performance of contrast-enhanced ultrasound liver imaging reporting and data system (CEUS LI-RADS).

Methods: Between January 2014 and December 2017, 1366 patients at risk for hepatocellular carcinoma (HCC) who underwent CEUS were included in this retrospective study. Four ultrasound physicians rated the HCC likelihood of focal liver lesions (FLLs) using CEUS LI-RADS v2017. Interreader agreement on CEUS LI-RADS categories and major features (arterial phase hyperenhancement (APHE), washout appearance) were assessed using weighted kappa statistics (κ). Diagnostic performance was described by sensitivity, specificity, PPV and NPV, +LR, -LR.

Results: The interreader agreement (κ) for CEUS LI-RADS categories, APHE, and washout appearance ranged from 0.61 to 0.73, 0.65 to 0.83, and 0.58 to 0.71, respectively. Interreader agreement for LI-RADS categories and APHE were almost substantial between FLLs < 2 cm, ≥ 2 cm, < 5 cm; interreader agreement for major features were fair to substantial for FLLs ≥ 5 cm. The accuracy, PPV and + LR for HCC and malignancy in FLLs < 2 cm, ≥ 2 cm, < 5 cm, ≥ 5 cm were high, with values of 84.7% to 91.9%, 90.2% to 94.2%, and 2.2 to 8.0, respectively. CEUS LI-RADS had the highest specificity for HCC (90.2%) and malignancy (90.9%) diagnosis for FLLs < 2 cm and < 5 cm, respectively; specificity was lowest for HCC (54.7%) and malignancy (68.3%) diagnosis for FLLs ≥ 5 cm.

Conclusions: CEUS LI-RADS is a good standardized categorization system for high-risk patients, and the combination of two or three LR-M features may improve the true-negative classification of HCC diagnosis.

1. Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignant tumor in the world [1] and the third leading cause of cancer death in China [2]. Cirrhosis, chronic hepatitis B or C and alcohol abuse are the main causes of liver cancer [3–5]. For patients at risk for HCC, many guidelines have recommended imaging examinations and the assessment of serum tumor markers at an interval of 1–6 months [6–9].

Contrast-enhanced imaging modalities play a pivotal role in monitoring and diagnosing HCC. Moreover, HCC is a unique solid organ tumor that can be definitely diagnosed solely by imaging methods with no need for pathology confirmation in patients with cirrhosis and chronic liver disease. However, there are some differences in the diagnostic criteria among different guidelines on HCC [6–9]. As recommended by the American Association for the Study of Liver Diseases (AASLD), for patients with high risk factors, if arterial phase hyperenhancement combined with washout presents in a nodule ≥ 1 cm, this nodule should be considered HCC [7]. However, the Asian Pacific Association for the Study of the Liver (APASL) suggests that HCC can be diagnosed as long as a nodule shows arterial phase hyperenhancement and late phase washout regardless of the nodule size in high-risk patients [9]. Therefore, for the same hepatic nodule, different medical institutions and radiologists could make different diagnoses, which may further affect clinical management.

Liver imaging reporting and data system (LI-RADS), developed by the American College of Radiology (ACR), is a comprehensive system used to standardize interpretations and reporting imaging examinations in patients at risk for HCC. With the wide use of contrast-enhanced ultrasound (CEUS) in clinical practice [10–12], ACR also launched LI-RADS on CEUS [13]. However, few studies [14–16] have investigated the interreader agreement and diagnostic performance of CEUS LI-RADS because it has been developed for only a short time. Additionally, in these studies, the results were inconsistent. Therefore, further identification in a large population is necessary. In this study, we

* Corresponding author.
E-mail address: luo_yan777@163.com (Y. Luo).

https://doi.org/10.1016/j.ejrad.2019.108685
Received 8 July 2019; Received in revised form 4 September 2019; Accepted 18 September 2019

investigated the interreader agreement and diagnostic performance of CEUS LI-RADS in 1366 high-risk patients.

2. Methods

2.1. Patient selection

This retrospective study was approved by the institutional ethics committee, and informed consent was waived. All patients underwent contrast enhanced ultrasound (CEUS) of the liver in our Ultrasound Department between January 2014 and December 2017 and were consecutively enrolled. For patients with more than one hepatic lesion, the most visible and accessible lesion was chosen for study. The inclusion criteria were as follows: (1) Focal liver lesions (FLLs) were solid and detected by conventional ultrasound with a CEUS examination also available; (2) patients with cirrhosis or chronic hepatitis; and (3) hepatic lesions were pathologically proven. The exclusion criteria were clinical intervention (such as radiofrequency ablation or hepatic arterial chemoembolization and partial liver resection) before CEUS. One board-certified ultrasound physician (15 years of experience in CEUS) reviewed the liver CEUS examinations. A total of 6861 patients underwent liver CEUS examinations during the period. Patients without any underlying liver disease (n = 3172); those who underwent transarterial chemoembolization (n = 252), radiofrequency ablation (n = 583), or partial liver resection (n = 423) before CEUS examinations; or those whose liver lesion nature was not confirmed by pathology (n = 1065) were excluded. Fig. 1 shows the flow chart patient inclusion in the study.

2.2. Contrast-enhanced ultrasonography (CEUS) examination

An iU22 ultrasound system (Royal Philips, the Netherlands) equipped with a C5-1 (1–5 MHz) transducer was used for the examinations. The CEUS examination was performed with a real-time, low-mechanical index (MI: 0.05–0.08) and reverse pulse imaging technique. In total, 2.4 mL of the SonoVue ultrasound contrast agent (Bracco, Milan, Italy) was injected into the medial cubital vein and immediately flushed with 5 mL of 0.9% sodium chloride solution. A timer was immediately started after the contrast agent was injected, and a dynamic digital video of the target lesion and surrounding liver parenchyma was recorded continuously within the first minute for analysis. After 60 s, intermittent scanning was performed to prevent the microbubbles from being destroyed too quickly; scanning continued until the microbubbles were completely cleared, and typical contrast-enhanced images of the lesions were stored for each scan.

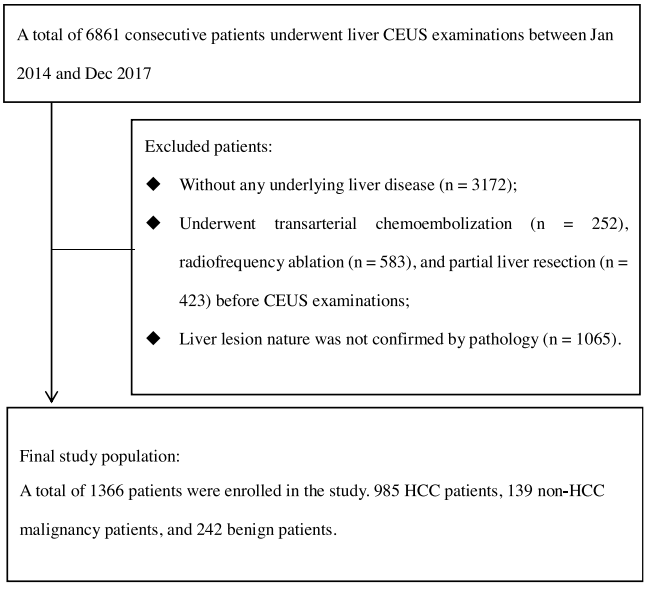

Fig. 1. The flow chart for inclusion of patients in the study.

2.3. Imaging analysis

Four ultrasound physicians independently classified all liver nodules according to CEUS LI-RADS v2017 [12]. Four readers were numbered as reader 1, reader 2, reader 3, and reader 4, who had 4, 5, 5, and 3 years of experience with CEUS, respectively. Four readers were trained together on CEUS LI-RADS for an hour, and two examples for each LI-RADS category were provided, which were not part of the study. All cases were numbered, with each having a unique number. Each reader was provided with the dynamic digital video within the first minute, typical contrast-enhanced images at rest in the portal phase and late phase, conventional ultrasound images and the diameter of the target lesion. All readers were unaware of any clinical information or other imaging results. The four readers assigned the LI-RADS categories and major enhancement features (arterial hyperenhancement, washout appearance (late in onset (ε 60 s) and mild washout)) independently. The final CEUS LI-RADS categorization for diagnostic performance evaluation was the consensus of the 4 readers.

2.4. Reference standard

All FLLs were confirmed by pathology. Due to uncertain imaging diagnosis or prior to radiofrequency ablation, ultrasound-guided puncture biopsy was performed for 198 FLLs; surgical resection was conducted for the other 1168 FLLs.

2.5. Statistical analysis

SAS 9.4 statistical software was used to perform data analysis. The weighted kappa test (k) with 95% confidence intervals (CIs) was used to evaluate the interreader agreement of category and major enhancement features of CEUS LI-RADS. A crossover table was used to show the consistency of different readers. The accuracy, sensitivity, specificity, positive and negative predictive value (PPV and NPV), and positive and negative likelihood ratios (+ LR and -LR) of CEUS LI-RADS for HCC and malignancy were calculated by comparison with pathology. Kappa results were stratified qualitatively by score (slight agreement, 0.01–0.20; fair agreement, 0.21–0.40; moderate agreement, 0.41–0.60; substantial agreement, 0.61–0.80; and almost perfect agreement, 0.81–0.99 [17]).

3. Results

3.1. Patient and tumor characteristics

The characteristics of the study patients are shown in Table 1. A total of 1366 patients were enrolled in the study, including 1097 males and 269 females. The age ranged from 18 to 90 years, with an average age (SD) of 52.3 (12.0) years. The average tumor diameter (SD) was 4.7 (3.5) cm, ranging from 0.5 to 20 cm. All patients had a history of chronic liver disease, and the etiologies were as follows: hepatitis B virus (n = 1300), hepatitis C virus (n = 38), alcohol-related liver disease (n = 2), autoimmune hepatitis (n = 4), nonalcoholic steatohepatitis (n = 11), and schistosomiasis infection (n = 11). Among the 1366 patients, 512 (37.5%, 512/1366) had cirrhosis (489 caused by HBV, eighteen by HCV, one by autoimmune hepatitis, and four by chronic schistosomiasis infection), and the other 854 (62.5%, 854/1366) did not have cirrhosis.

The final diagnoses for all FLLs were 985 (72.1%, 985/1366) HCCs, 139 (10.2%, 139/1366) non-HCC malignancies, and 242 (17.7%, 242/1366) benign lesions. The mean sizes (SD, range) of HCCs, benign lesions and non-HCC malignancies were 5.2 (3.6, 0.8 ~ 20) cm, 5.3 (3.3, 0.7–14) cm, and 4.8 (3.9, 0.8–16) cm, respectively. One hundred thirty-

Table 1
Characteristics of the Study Subjects (n = 1366).

Variable	Count
Age, mean (SD), y	52.3 (12.0)
Sex	
Male	1097 (80.3%)
Female	269 (19.7%)
Tumor size	
Mean (SD)	4.7 (3.5) cm
Range	(0.5-20) cm
Etiology	
Hepatitis B virus	1300 (95.2%)
Hepatitis C virus	38 (2.8%)
Alcoholic abuse	2 (0.1%)
Autoimmune hepatitis	4 (0.3%)
Non-alcoholic steatohepatitis	11 (0.7%)
Schistosomiasis infection	11 (0.8%)
Pathology	
HCC	985 (72.1%)
Non-HCC malignancies	139 (10.2%)
Benign lesions	242 (17.7%)

SD, standard deviation. cm, centimeter. HCC, hepatocellular carcinoma.

nine non-HCC lesions included 59 intrahepatic cholangiocarcinoma (ICC), 62 liver metastases, 14 combined hepatocellular-cholangiocarcinoma, two hepatic sarcomas, and two liver lymphomas. Two hundred forty-two benign lesions included 198 hepatic hemangiomas, nine hyperplastic nodules, eighteen focal nodular hyperplasia (FNH), twelve hepatic hydatidosis, and five hepatic adenomas.

3.2. CEUS LI-RADS categories

The frequency distribution of LI-RADS for 1366 FLLs classified by reader 1 to reader 4 is presented in Table 2. The final CEUS LI-RADS categories, which were based on the consensus of reader 1 to reader 4, are displayed in Table 3.

Among the 1366 FLLs, none of the LI-RADS category 1 and 2 lesions, 9 (11.5%) of 78 category 3 lesions, 81 (72.3%) of 112 category 4 lesions, 723 (93.3%) of 775 category 5 lesions (Fig. 2), and 172 (59.7%) of 288 category M lesions were HCCs (Fig. 3). All LI-RADS category 1 and category 2 lesions, 67 (85.9%) of 78 category 3 lesions, 29 (25.9%) of 112 category 4 lesions, 12 (1.5%) of 775 category 5 lesions, and 21 (7.3%) of 288 category M lesions were benign. None of the LI-RADS category 1 and 2 lesions, 2 (2.6%) of 78 category 3 lesions, 2 (1.8%) of 112 category 4 lesions, 40 (5.2%) of 775 category 5 lesions, and 95 (33.0%) of 288 category M lesions were malignancy.

Further analysis of CEUS patterns of HCCs and non-HCC malignancies in the LR-M category was performed. Among 172 LR-M HCCs, 79.7% (137/172) only had one of the three features of LR-M criteria, which was mainly characterized by early washout (80.3%, 110/137). However, among the 95 LR-M non-HCC malignancies, 81.1% (77/95) lesions had at least two of three features of LR-M criteria, mainly arterial phase hyperenhancement (APHE) accompanied by early (< 60 s) washout and marked washout (61.0%, 47/77) and rim APHE accompanied by early (< 60 s) washout and marked washout (35.1%, 27/77). The detail CEUS characteristics for HCC and non-HCC malignancies within the LR-M category are shown in Table 4.

Table 2
Frequency distributions of CEUS LI-RADS for 1366 focal liver lesions classified by reader 1 to reader 4.

LI-RADS category	Reader 1	Reader 2	Reader 3	Reader 4
LR-1	100 (7.3)	112 (8.2)	82 (6.0)	104 (7.6)
LR-2	9 (0.7)	7 (0.5)	7 (0.5)	8 (0.6)
LR-3	83 (6.1)	71 (5.2)	90 (6.6)	83 (6.1)
LR-4	146 (10.7)	68 (5.0)	154 (11.3)	117 (8.6)
LR-5	799 (58.5)	763 (55.9)	748 (54.7)	771 (56.4)
LR-M	229 (16.7)	345 (25.2)	285 (20.9)	283 (20.7)

The table presented the frequency distribution of CEUS LI-RADS categories assigned by four readers. Numbers in parentheses are percentages of each category. CEUS, contrast enhanced ultrasound. LI-RADS, Liver Imaging Reporting and Data System.

3.3. Interreader agreement on CEUS LI-RADS categories and major features

The interreader agreement for CEUS LI-RADS and major enhancement features that were classified by the four readers are shown in Table 5. The interreader agreement on FLL categories was substantial (κ coefficient from 0.61 to 0.73). The interreader agreement for major features was substantial to almost perfect for "arterial hyperenhancement" (κ coefficient from 0.65 to 0.83) and moderate to substantial for "washout appearance" (κ coefficient from 0.58 to 0.71).

3.4. In-depth analysis of interreader agreement

The interreader agreement for patients with cirrhosis (n = 512) was evaluated separately. The interreader agreement for FLL categories was substantial (κ coefficient from 0.60 to 0.70). The interreader agreement was substantial for "arterial hyperenhancement" (κ coefficient from 0.64 to 0.79) and moderate to substantial for "washout appearance" (κ coefficient from 0.47 to 0.77).

Subanalysis revealed substantial to almost perfect interreader agreement for FLLs with a diameter < 2 cm (κ coefficient from 0.70 to 0.83) and moderate to substantial interreader agreement for FLLs ≥ 2 cm (κ coefficient from 0.55 to 0.67).

Interreader agreement of major features for FLLs with a diameter < 2 cm and ≥ 2 cm was similar for "arterial hyperenhancement" (κ coefficient from 0.67 to 0.80, and 0.61 to 0.79, respectively) and "washout appearance" (κ coefficient from 0.53 to 0.78, and 0.50 to 0.71, respectively).

Moreover, subanalysis showed substantial interreader agreement for FLLs with a diameter < 5 cm (κ coefficient from 0.71 to 0.76) and moderate to substantial interreader agreement for FLLs ≥ 5 cm (κ coefficient from 0.45 to 0.63).

Interreader agreement for "arterial hyperenhancement" was higher for FLLs < 5 cm than for FLLs ≥ 5 cm (κ coefficient from 0.68 to 0.84, and 0.40 to 0.73, respectively) as well as for "washout appearance" for the same size categories (κ coefficient from 0.56 to 0.75, and 0.30 to 0.65, respectively). Fig. 4 depicts an example of a case that led to potential variability in assigning major LI-RADS features.

3.5. Diagnostic performance of CEUS LI-RADS for HCC and malignancy

For all 1366 FLLs, the proportions of HCC in LR-3, LR-4, and LR-5 were 11.5% (9/78), 72.3% (81/112), and 93.3% (723/775), respectively. As illustrated in Table 6, when LR-M (because it is not specific for HCC) was excluded and LR-5 was used as the criteria for a definite determination of HCC, the accuracy, sensitivity, specificity, PPV, NPV, +LR and -LR of CEUS LI-RADS for HCC were 86.8% (95%CI: 84.7%, 88.7%), 89.1% (95%CI: 86.7,91.1), 80.4% (95%CI: 75.0%, 84.9%), 93.4% (95%CI: 91.3%, 95.0%), 70.3% (95%CI: 64.8%, 75.3%), 4.5 (3.5,5.8), and 0.14 (0.11,0.17), respectively.

When considering LR-5 and M as definite malignancy, the sensitivity, specificity, PPV and NPV of CEUS LI-RADS for malignancy were 90.7% (95%CI: 89.0%, 92.1%), 91.6% (95%CI: 89.8%, 93.2%), 86.4% (95%CI:81.2%, 90.3%), 96.9% (95%CI: 95.6%, 97.8%), 69.0% (95%CI: 63.4%, 74.1%), 6.7(4.9,9.2), and 0.10(0.08, 0.12), respectively.

3.6. In-depth analysis of diagnostic performance of CEUS LI-RADS for HCC and malignancy for FLLs of different sizes

For FLLs with a diameter < 2 cm, ≥ 2 cm, < 5 cm, ≥ 5 cm, the

Table 3
The final frequency distributions of CEUS LI-RADS.

Group	Reference standard	LR-1	LR-2	LR-3	LR-4	LR-5	LR-M
Overall	HCC	0	0	9 (11)	81 (72)	723 (93)	172 (60)
(n = 1366)	Benign Lesions	106 (100)	7 (100)	67 (86)	29 (26)	12 (2)	21 (7)
	Non-HCC malignancies	0	0	2 (3)	2 (2)	40 (5)	95 (33)
	Total	106	7	78	112	775	288
Diameter < 2 cm	HCC	0	0	1 (2)	21 (60)	77 (87)	17 (61)
(n = 250)	Benign Lesions	34 (100)	3 (100)	60 (98)	13 (37)	11 (12)	2 (7)
	Non-HCC malignancies	0	0	0	1 (3)	1 (1)	9 (32)
	Total	34	3	61	35	89	28
Diameter ≥ 2 cm	HCC	0	0	8 (47)	60 (78)	646 (94)	155 (60)
(n = 1116)	Benign Lesions	72 (100)	4 (100)	7 (41)	16 (21)	1 (0)	19 (7)
	Non-HCC malignancies	0	0	2 (12)	1 (1)	39 (6)	86 (33)
	Total	72	4	17	77	686	260
Diameter < 5 cm	HCC	0	0	9 (12)	77 (74)	433 (94)	89 (57)
(n = 886)	Benign Lesions	81 (100)	7 (100)	67 (86)	26 (25)	5 (1)	15 (10)
	Non-HCC malignancies	0	0	2 (2)	1 (1)	23 (5)	51 (33)
	Total	81	7	78	104	461	155
Diameter ≥ 5 cm	HCC	0	0	0	4 (50)	290 (92)	83 (62)
(n = 480)	Benign Lesions	25 (100)	0	0	3 (38)	7 (2)	6 (5)
	Non-HCC malignancies	0	0	0	1 (12)	17 (6)	44 (33)
	Total	25	0	0	8	314	133

This table showed the distribution of CEUS LI-RADS categories of focal liver lesions with different size. CEUS, contrast enhanced ultrasound. LI-RADS Liver Imaging Reporting and Data System. HCC, Hepatocellular carcinoma. Note. Data are numbers of patients, data in parentheses are percentages.

proportions of HCC in LR-3, LR-4, and LR-5 were 2%, 60%, 87% vs 8%, 78%, 94% vs 12%, 74%, 94% vs 0%, 50%, 92%, respectively.

As indicated in Table 6, when considering the LR-5 category as HCC (LR-M was excluded due to the lack of specificity for HCC), the accuracy, PPV and + LR of CEUS LI-RADS for HCC in FLLs with a diameter < 2 cm, ≥2 cm, < 5 cm, and ≥5 cm were high, with corresponding values of 84.7% (95%CI: 79.3%, 88.9%), 90.2% (95%CI: 83.2%, 94.6%) and 8.0 (95%CI: 4.6, 13.8) vs 87.4% (95%CI: 85.0%, 89.5%), 94.2% (95%CI: 92.1%, 95.8%), and 3.2 (95%CI: 2.5,4.2) vs 84.4% (95%CI: 81.6%, 86.9%), 93.9% (95%CI: 91.2%, 95.9%), and 6.3(4.5,8.9) vs 91.9% (95%CI: 88.6%, 94.4%), 92.4% (95%CI: 88.7%, 94.9%), and 2.2 (1.6,2.9).

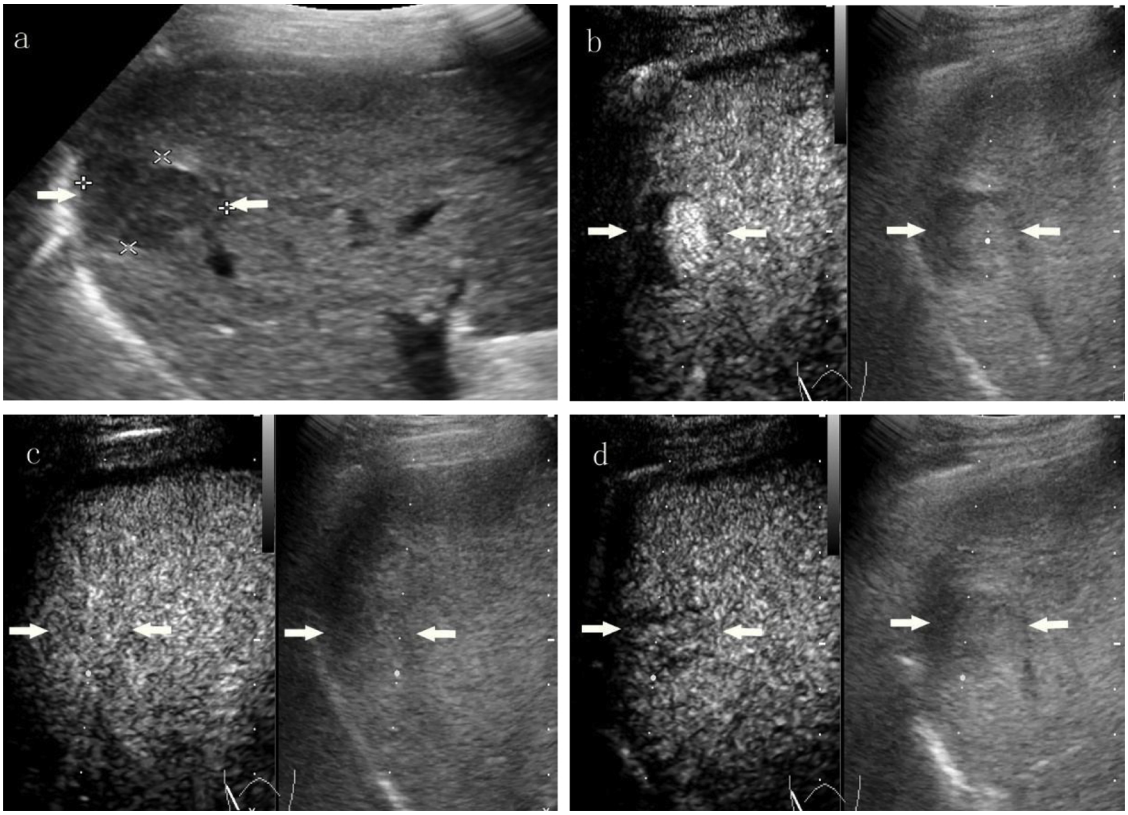

Fig. 2. The arrow highlights a 2.7 cm hypoechoic tumor in liver segment VII of a 60-year-old male patient with cirrhosis (a). The tumor presented with partial hyperenhancement in the arterial phase (b), isoenhancement in the portal phase (c) and mild washout in late phase (d). The lesion was classified as LR-5 and proved to be highly differentiated hepatocellular carcinoma by pathology.

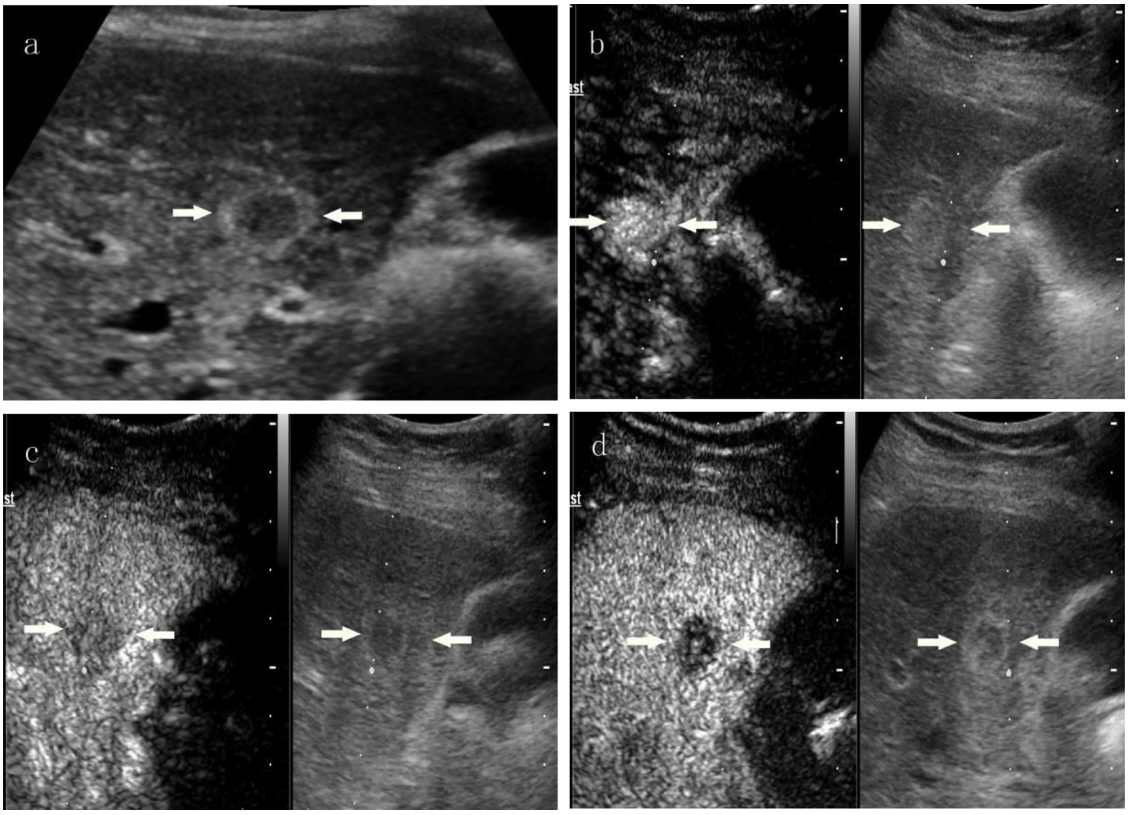

Fig. 3. The arrow highlights a 1.9 cm hypoechoic tumor in liver segment V of a 45-year-old female patient with cirrhosis (a). The tumor presented with hyper-enhancement in the arterial phase (b), washout started after 49 s (c) and mild washout in late phase (d). The lesion was classified as LR-M and proved to be moderately differentiated hepatocellular carcinoma by pathology.

Table 4
The detail CEUS characteristics of HCC and non-HCC malignancies within LR-M category.

LR-M	rim APHE & early washout & marked washout	rim APHE & early washout & mild washout	rim APHE & late washout & mild washout	APHE & early washout & marked washout	APHE & early washout & mild washout	APHE & late washout & marked washout
HCC	2	7	5	26	110	22
ICC	8	1	1	27	5	4
cHCC-CC	0	1	0	2	2	0
Metastasis	19	1	1	16	4	1
Sarcoma	0	0	0	2	0	0

CEUS, contrast enhanced ultrasound. APHE, arterial phase hyperenhancement. HCC, Hepatocellular carcinoma. ICC, intrahepatic cholangiocarcinoma. cHCC-CC, combined hepatocellular cholangiocarcinoma.

Table 5
Interreader Agreement (κ) for CEUS LI-RADS categories and major enhancement features classified by four readers.

Reader	Variable	*k* coefficients		
		Reader2	Reader3	Reader4
Reader1	LI-RADS categories	0.65 (0.63,0.67)	0.61 (0.59,0.63)	0.73 (0.71,0.75)
	Arterial hyper-enhancement	0.67 (0.63,0.70)	0.65 (0.63,0.67)	0.73 (0.71,0.75)
	Washout appearance	0.67 (0.65,0.69)	0.64 (0.62,0.66)	0.69 (0.67,0.71)
Reader2	LI-RADS categories		0.67 (0.65,0.69)	0.73 (0.71,0.75)
	Arterial hyper-enhancement		0.70 (0.68,0.72)	0.76 (0.74,0.78)
	Washout appearance		0.71 (0.68,0.74)	0.64 (0.62,0.66)
Reader3	LI-RADS categories			0.69 (0.67,0.71)
	Arterial hyper-enhancement			0.83 (0.80,0.86)
	Washout appearance			0.58 (0.56,0.60)

Interreader agreement was calculated by Weighted kappa. Arterial hyper-enhancement: whole or in part, not rim or peripheral discontinuous globular enhancement. Washout appearance: late in onset (≥ 60 s) and mild washout. CEUS, contrast enhanced ultrasound. LI-RADS, Liver Imaging Reporting and Data System. Note. Data in parentheses are 95% confidence intervals.

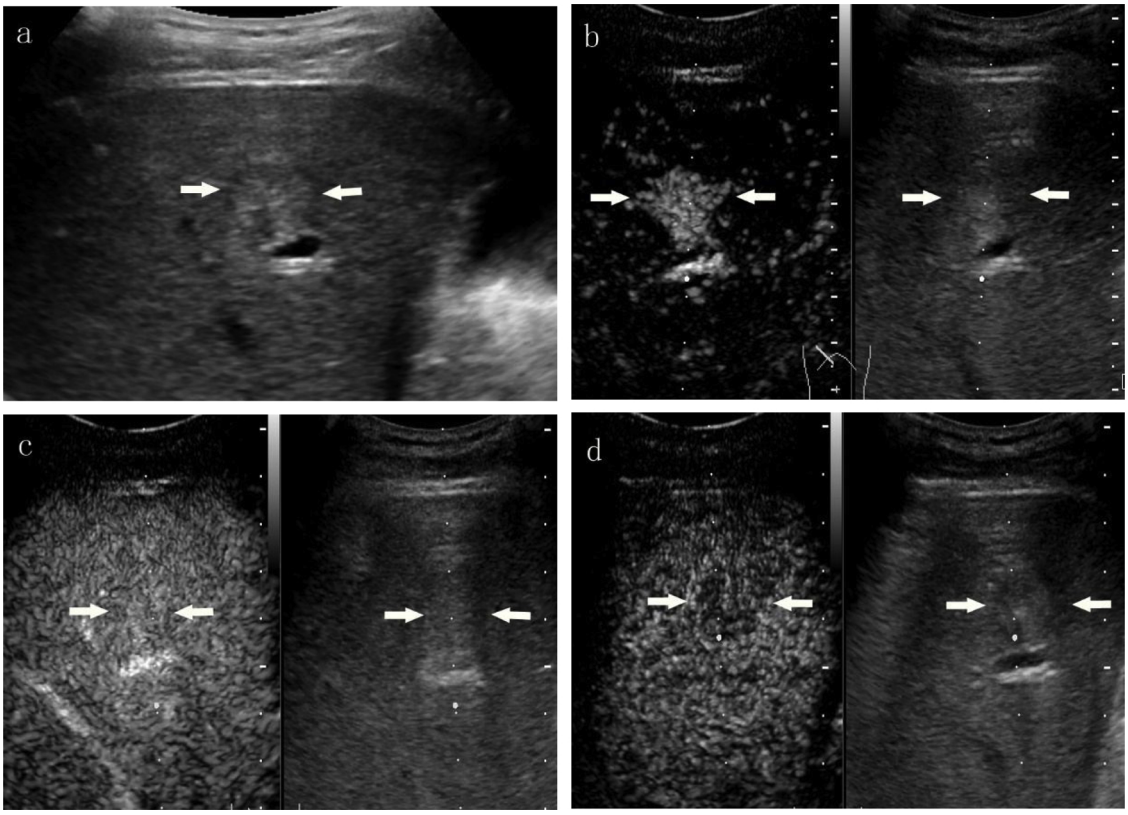

Fig. 4. It illustrates an example of a case which led to the potential variability in assigning major LI-RADS features. The arrow highlights a 2.4 cm hypoechoic tumor in liver segment V of a 63-year-old male patient with chronic hepatitis B (a). The assigned Liver Imaging Reporting and Data System (LI-RADS) categories were LR-5 (three readers) and LR-4 (one reader), demonstrating substantial interreader agreement for LI-RADS. For this particular observation, three readers agreed on two major features (arterial hyperenhancement, late (≥60 s) and mild washout), and only one reader interpreted hyperenhancement in arterial phase, isoenhancement in portal and late phase. The figure shows a hyperenhancement at arterial phase (b) and isoenhancement in the portal phase (c), and mild washout in late phase (d). The lesion was proved to be moderately differentiated hepatocellular carcinoma by pathology.

CEUS LI-RADS also exhibited the highest specificity (90.2%, 95%CI: 83.2%, 94.6%) for HCC diagnosis for FLLs with a diameter < 2 cm and the lowest specificity (54.7%, 95%CI: 40.6%, 68.2%) for FLLs ≥5 cm.

For malignancy diagnosis, the accuracy, PPV and + LR of CEUS LI-RADS in FLLs with a diameter < 2 cm, ≥2 cm, < 5 cm, and ≥5 cm were high, with corresponding values of 85.6% (95%CI: 80.7%, 89.5%), 88.9 (95%CI: 81.4%, 93.7%), and 7.7 (4.6, 13.0) vs 91.9% (95%CI: 90.1%, 93.3%), 97.9% (95%CI: 96.7%, 98.7%), and 5.5 (3.7, 8.2) vs 87.7% (95%CI: 85.4%, 89.7%), 96.8% (95%CI: 94.9%, 98.0%), 8.7 (5.8, 13.3) vs 96.3% (95%CI: 94.1%, 97.7%), 97.1% (95%CI: 94.9%, 98.4%), and 3.1 (2.0, 4.9).

CEUS LI-RADS demonstrated the lowest specificity (68.3%, 95%CI: 51.8%, 81.4%) for malignancy diagnosis for FLLs with a diameter ≥5 cm and the highest specificity for a diameter < 5 cm (90.9%, 95%CI: 84.8%, 93.7%).

4. Discussion

Our study showed that the consistency of the four readers on categories of CEUS LI-RADS was substantial (κ, 0.61-0.73), and there was no obvious difference between every two readers, which may be related to each reader having at least 3 years of experience with CEUS in the study. We further analyzed the consistency among four readers for cases of cirrhosis and focal liver lesions with different sizes. The results showed that CEUS LI-RADS had similar interreader agreement in categories and major features between patients with cirrhosis and those with risk factors for HCC. In addition, interreader agreement for LI-RADS categories was almost substantial for FLLs < 2 cm, ≥2 cm, and < 5 cm, and interreader agreement for major features was moderate to substantial for FLLs ≥5 cm. The results indicated that CEUS LI-RADS could reduce the differences in the diagnosis of HCC among ultrasound physicians, which is also consistent with the purpose of CEUS LI-RADS. Ling et al [14] reported that the interreader agreement for CEUS LI-RADS on small hepatic nodules (≤2 cm) between two readers who had more than 5 years of experience in abdominal ultrasound imaging was substantial (κ, 0.690), which is consistent with our results. Furthermore, Terzi et al [18] reported an interobserver reproducibility for 120 patients of 0.755, which is in line with our results. However, Schellhaas [15] indicated that the interobserver agreement for the CEUS LI-RADS category in 50 patients was fair (κ, 0.309). They postulated that there were two reasons for this result. One reason was that the perception of arterial phase hyperenhancement in CEUS was different from that in MRI. The other reason was that readers were more familiar with MRI LI-RADS than CEUS LI-RADS, since CEUS LI-RADS was only recently developed.

Because CEUS is an imaging modality associated with certain observer-experience and subjectivity, we further explored the interreader agreement on the major features of CEUS LI-RADS. Substantial to perfect interreader agreement for arterial hyperenhancement (k, 0.65 to 0.83), and moderate to substantial interreader agreement for washout appearance (k, 0.58 to 0.71) was observed in the study. In addition, interreader agreement for arterial phase hyperenhancement was substantial and moderate to substantial for washout appearance for FLLs < 2 cm, ≥ 2 cm, and < 5 cm. Nonetheless, interreader agreement for major features was fair to substantial for FLLs ≥5 cm, which suggests considerable inconsistency for FLLs ≥5 cm. The results of studies on the interreader agreement for arterial hyperenhancement and wash appearance varied from fair to substantial [16,19–22]. Researchers

Table 6
Diagnostic performance of CEUS LI-RADS for HCC and malignanc.

		No. of Nodules				Diagnostic Performance (%)						
		TP	FP	FN	TN	Accuracy	Sensitivity	Specificity	PPV	NPV	+ LR	- LR
Overall	**Diagnostic Criteria for HCC**											
(n = 1366)	LR-5	723	52	90	213	86.8(84.7,88.7)	89.1(86.7,91.1)	80.4(75.0,84.9)	93.4(91.3,95.0)	70.3(64.8,75.3)	4.5(3.5,5.8)	0.14(0.11,0.17)
	Diagnostic Criteria for malignancy											
	LR-5 and M	1030	33	94	209	90.7(89.0,92.1)	91.6(89.8,93.2)	86.4(81.2,90.3)	96.9(95.6,97.8)	69.0(63.4,74.1)	6.7(4.9,9.2)	0.10(0.08, 0.12)
Diameter < 2 cm	**Diagnostic Criteria for HCC**											
(n = 250)	LR-5	77	12	22	111	84.7(79.3,88.9)	77.8(68.1,85.3)	90.2(83.2,94.6)	86.5(77.2,92.5)	83.5(75.8,89.1)	8.0(4.6,13.8)	0.25(0.17,0.36)
	Diagnostic Criteria for malignancy											
	LR-5 and M	104	13	23	110	85.6(80.7,89.5)	81.9(73.9,87.9)	89.4(82.3,94.0)	88.9(81.4,93.7)	82.7(75.0,88.5)	7.7(4.6,13.0)	0.20(0.14,0.29)
Diameter ≥ 2 cm	**Diagnostic Criteria for HCC**											
(n = 1116)	LR-5	646	40	68	102	87.4(85.0,89.5)	90.5(88.0,92.5)	71.8(63.6,78.9)	94.2(92.1,95.8)	60.0(52.2,67.3)	3.2(2.5,4.2)	0.13(0.11,0.17)
	Diagnostic Criteria for malignancy											
	LR-5 and M	926	20	71	99	91.9(90.1,93.3)	92.9(91.1,94.5)	83.2(75.0,89.2)	97.9(96.7,98.7)	58.2(50.4,65.7)	5.5(3.7,8.2)	0.09(0.07,0.11)
Diameter < 5 cm	**Diagnostic Criteria for HCC**											
(n = 886)	LR-5	433	28	86	184	84.4(81.6,86.9)	83.4(79.9,86.5)	86.8(81.3,90.9)	93.9(91.2,95.9)	68.1(62.2,73.6)	6.3(4.5,8.9)	0.19(0.16,0.23)
	Diagnostic Criteria for malignancy											
	LR-5 and M	596	20	89	181	87.7(85.4,89.7)	87.8(84.2,89.4)	90.9(84.8,93.7)	96.8(94.9,98.0)	67.0(61.0,72.5)	8.7(5.8,13.3)	0.14(0.12,0.18)
Diameter ≥ 5 cm	**Diagnostic Criteria for HCC**											
(n = 480)	LR-5	290	24	4	29	91.9(88.6,94.4)	98.6(96.3,99.6)	54.7(40.6,68.2)	92.4(88.7,94.9)	87.9(70.9,96.0)	2.2(1.6,2.9)	0.02(0.01,0.07)
	Diagnostic Criteria for malignancy											
	LR-5 and M	434	13	5	28	96.3(94.1,97.7)	98.9(97.2,99.6)	68.3(51.8,81.4)	97.1(94.9,98.4)	84.4(67.3,94.3)	3.1(2.0,4.9)	0.02(0.01,0.04)

CEUS, contrast enhanced ultrasound. LI-RADS, Liver Imaging Reporting and Data System. TP, true-positive. FP, false-positive. FN, false-negative. TN, true-negative. PPV, positive predictive value. NPV, negative predictive value. Note. Data in parentheses are 95% confidence intervals.

attribute this discrepancy to the latent subjective cognition of readers in recognition of enhancement features as well as the clinical experience of readers. Our results showed that interreader agreement was slightly better for arterial hyperenhancement than washout appearance, which may be related to the readers' cognitive difference in washout appearance.

None of LR-1 and 2 were HCCs or malignancies in the study, and the proportion of HCC lesions increased gradually from LR-3, LR-4 to LR-5, at 11.5% (9/78), 72.3% (81/112), and 93.3% (723/775), respectively. The study of van der Po et al l [23] consisted of a systematic review of 17 studies on LI-RADS, with 2760 patients and 3556 FLLs, and the authors found 94% LR-5 lesions and 74% LR-4 lesions were HCCs, which is almost consistent with our results. However, 38% LR-3 lesions were HCCs in their study, which was different from our results and may be due to the number of LR-3 lesions was relatively smaller in our study. In addition, 59.7% (172/288) of LR-M lesions were HCCs in the study, which is in line with the findings of some other studies [18,23–25]. Lee et al. [24] showed that at least one LR-M feature was observed in 51.5% of patients (34/66) with HCC in their study. In a study by Kim et al [25], 76.9% (10/13) of LR-M lesions were HCCs. In this study, we found that among the 172 cases of HCC in LR-M, 137 (79.7%, 137/172) cases had only one of three features for LR-M, which led to many HCCs classified as LR-M. However, for the other 95 cases of non-HCC malignancies in LR-M, 77 (81.1%, 77/95) cases had at least two of three features for LR-M. Therefore, the combination of two or three LR-M features may help improve the true-negative classification for HCC diagnosis.

This study showed that CEUS LI-RADS has a good diagnostic performance for HCC when LR-M lesions are excluded. Especially for LR-5 lesions, the diagnostic accuracy was up to 93.3% (723/775), and LR-5 criteria showed high specificity for HCC diagnosis, which is comparable to the findings of a previous study [5,23,26]. We further analyzed the diagnostic performance of CEUS LI-RADS for FLLs with different sizes, and the results showed that the accuracy and + LR of CEUS LI-RADS in the diagnosis of HCC and malignancy were high regardless of the FLL size. Moreover, CEUS LI-RADS showed high specificity in the diagnosis of HCC and malignancy for FLLs with a diameter < 2 cm (90.2% and 89.4%, respectively) and < 5 cm (86.8% and 90.9%, respectively). This is similar to the results of Ronot et al [27], who reported a specificity of LI-RADS for HCC on MRI and CT of 89.9% and 88.3%, respectively, when FLLs were smaller than 3 cm. However, the specificity of CEUS LI-RADS for HCC and malignancy was low (54.7% and 68.3%, respectively) for FLLs ≥ 5 cm, and the reason needs further exploration.

However, there were serval limitations in our study. First, this was a retrospective and single-center study, and a prospective study is needed to verify the results. Second, the four readers have similar CEUS experience (from 3 to 5 years), which may have caused the high interreader agreement of CEUS LI-RADS in this study, and interreader agreement among readers with different CEUS experiences should be further explored. Third, there were relatively few patients with LR-2 lesions, which was due to these lesions usually being followed-up by imaging modalities. To strengthen the reference standard, only patients with pathological confirmation were included, which also led to patient selection bias in the study. Finally, the readers did not rate the quality of the acquisitions, which is also a limitation of the study.

In conclusion, CEUS LI-RADS has good consistency among different ultrasound physicians and is a good standardized categorization system for high-risk patients. In addition, the combination of two or three LR-M features may help improve the true-negative classification of HCC diagnosis.

Funding

This research was supported by Post-Doctor Research Project, West China Hospital, Sichuan University (NO. 2018HXBH073), and National Natural Science Foundation of China (No. 81701702).

Acknowledgements

The authors wish to thank Cai-Rong Zhu, PhD (West China School of Public Health, Sichuan University) for her statistical suggestions.

References

[1] L.A. Torre, F. Bray, R.L. Siegel, J. Ferlay, J. Lortet-Tieulent, A. Jemal, Global cancer statistics, 2012, CA Cancer J. Clin. 65 (2) (2015) 87–108.
[2] W. Chen, R. Zheng, P.D. Baade, S. Zhang, H. Zeng, F. Bray, et al., Cancer statistics in China, 2015, CA Cancer J. Clin. 66 (2) (2016) 115–132.
[3] J. Fehér, G. Lengyel, Hepatocellular carcinoma: occurrence, risk factors, biomarkers, Orv. Hetil. 151 (23) (2010) 933–940.
[4] R.H. Luo, Z.X. Zhao, X.Y. Zhou, Z.L. Gao, J.L. Yao, Risk factors for primary liver carcinoma in Chinese population, World J. Gastroenterol. 11 (28) (2005) 4431–4434.
[5] C. Cassinotto, C. Aubé, A. Dohan, Diagnosis of hepatocellular carcinoma: an update on international guidelines, Diagn. Interv. Imaging 98 (5) (2017) 379–391.
[6] European Association for the Study of the Liver, EASL clinical practice guidelines: management of hepatocellular carcinoma, J. Hepatol. 69 (1) (2018) 182–236.
[7] J.A. Marrero, L.M. Kulik, C.B. Sirlin, A.X. Zhu, R.S. Finn, M.M. Abecassis, et al., Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the american association for the study of liver diseases, Hepatology 68 (2) (2018) 723–750.
[8] J. Zhou, H.C. Sun, Z. Wang, W.M. Cong, J.H. Wang, M.S. Zeng, et al., Guidelines for diagnosis and treatment of primary liver cancer in china (2017 edition), Liver Cancer 7 (2018) 235–260.
[9] M. Omata, A.L. Cheng, N. Kokudo, M. Kudo, J.M. Lee, J. Jia, et al., Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update, Hepatol. Int. 11 (4) (2017) 317–370.
[10] M. D'Onofrio, S. Crosara, R. De Robertis, S. Canestrini, R.P. Mucelli, Contrast-enhanced ultrasound of focal liver lesions, AJR 205 (1) (2015) W56–66.
[11] R.F. Hanna, V.Z. Miloushev, A. Tang, L.A. Finklestone, S.Z. Brejt, R.S. Sandhu, et al., Comparative 13-year meta-analysis of the sensitivity and positive predictive value of ultrasound, CT, and MRI for detecting hepatocellular carcinoma, Abdom. Radiol. (NY) 41 (1) (2016) 71–90.
[12] M. Claudon, C.F. Dietrich, B.I. Choi, et al., 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, Ultrasound Med. Biol. 39 (2) (2013) 187–210 Ultraschall Med 2013;34(1):11-29.
[13] American College of Radiology. Liver Imaging Reporting and Data System, (2017) https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/CEUS-LI-RADS-v2017.
[14] W. Ling, M. Wang, X. Ma, et al., The preliminary application of liver imaging reporting and data system (LI-RADS) with contrast-enhanced ultrasound (CEUS) on small hepatic nodules ($\leq$ 2cm), J. Cancer 9 (16) (2018) 2946–2952.
[15] B. Schellhaas, M. Hammon, D. Strobel, et al., Interobserver and intermodality agreement of standardized algorithms for non-invasive diagnosis of hepatocellular carcinoma in high-risk patients: CEUS-LI-RADS versus MRI-LI-RADS, Eur. Radiol. 28 (10) (2018) 4254–4264.
[16] B. Schellhaas, L. Pfeifer, C. Kielisch, R.S. Goertz, M.F. Neurat, D. Strobel, Interobserver agreement for contrast-enhanced ultrasound (CEUS)-based standardized algorithms for the diagnosis of hepatocellular carcinoma in high-risk patients, Ultraschall Med. 39 (6) (2018) 667–674.
[17] A.J. Viera, J.M. Garrett, Understanding interobserver agreement: the Kappa statistic, Fam. Med. 37 (5) (2005) 360–363.
[18] E. Terzi, M. Iavarone, M. Pompili, L. Veronese, G. Cabibbo, M. Fraquelli, et al., Contrast ultrasound LI-RADS LR-5 identifies hepatocellular carcinoma in cirrhosis in a multicenter restropective study of 1,006 nodules, J. Hepatol. 68 (3) (2018) 485–492.
[19] M.S. Davenport, S. Khalatbari, P.S. Liu, et al., Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging, Radiology 272 (1) (2014) 132–142.
[20] B.K. Barth, O.F. Donati, M.A. Fischer, et al., Reliability, validity, and reader acceptance of LI-RADS-an in-depth analysis, Acad. Radiol. 23 (9) (2016) 1145–1153.
[21] M.R. Bashir, R. Huang, N. Mayes, et al., Concordance of hypervascular liver nodule characterization between the organ procurement and transplant network and liver imaging reporting and data system classifications, J. Magn. Reson. Imaging 42 (2) (2015) 305–314.
[22] Y.D. Zhang, F.P. Zhu, X. Xu, et al., Classifying CT/MR findings in patients with suspicion of hepatocellular carcinoma: comparison of liver imaging reporting and data system and criteria-free Likert scale reporting models, J. Magn. Reson. Imaging 43 (2) (2016) 373–383.
[23] C.B. van der Pol, C.S. Lim, C.B. Sirlin, T.A. McGrath, J.P. Salameh, M.R. Bashir, et al., Accuracy of the liver imaging reporting and data system in computed tomography and magnetic resonance image analysis of hepatocellular carcinoma or overall malignancy-a systematic review, Gastroenterology 156 (4) (2019) 976–986.
[24] H.S. Lee, M.J. Kim, C. An, How to utilize LR-M features of the LI-RADS to improve the diagnosis of combined hepatocellular-cholangiocarcinoma on gadoxetate-enhanced MRI? Eur. Radiol. 29 (5) (2019) 2408–2416.
[25] Y.Y. Kim, C. An, S. Kim, M.J. Kim, Diagnostic accuracy of prospective application of the Liver Imaging Reporting and Data System (LI-RADS) in gadoxetate-enhanced MRI, Eur. Radiol. 28 (5) (2018) 2038–2046.
[26] B. Schellhaas, D. Wildner, L. Pfeifer, R.S. Goertz, A. Hagel, M.F. Neurath, et al., LI-RADS-CEUS-Proposal for a contrast-enhanced ultrasound algorithm for the diagnosis of hepatocellular carcinoma in high-risk populations, Ultraschall Med. 37 (6) (2016) 627–634.
[27] M. Ronot, O. Fouque, M. Esvan, J. Lebigot, C. Aube, V. Vilgrain, Comparison of the accuracy of AASLD and LI-RADS criteria for the non-invasive diagnosis of HCC smaller than 3 cm, J. Hepatol. 68 (4) (2018) 715–723.