



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

Barbara Schellhaas¹ · Matthias Hammon² · Deike Strobel¹ · Lukas Pfeifer¹ · Christian Kielisch¹ · Ruediger S. Goertz¹ · Alexander Cavallaro² · Rolf Janka² · Markus F. Neurath¹ · Michael Uder² · Hannes Seuss² 

Received: 2 October 2017 / Revised: 26 January 2018 / Accepted: 7 February 2018 / Published online: 19 April 2018

© European Society of Radiology 2018

Abstract

Objectives We compared the interobserver agreement for the recently introduced contrast-enhanced ultrasound (CEUS)-based algorithm CEUS-LI-RADS (Liver Imaging Reporting and Data System) versus the well-established magnetic resonance imaging (MRI)-LI-RADS for non-invasive diagnosis of hepatocellular carcinoma (HCC) in high-risk patients.

Methods Focal liver lesions in 50 high-risk patients (mean age 66.2 ± 11.8 years; 39 male) were assessed retrospectively with CEUS and MRI. Two independent observers reviewed CEUS and MRI examinations, separately, classifying observations according to CEUS-LI-RADSv.2016 and MRI-LI-RADSv.2014. Interobserver agreement was assessed with Cohen's kappa.

Results Forty-three lesions were HCCs; two were intrahepatic cholangiocarcinomas; five were benign lesions. Arterial phase hyperenhancement was perceived less frequently with CEUS than with MRI (37/50 / 38/50 lesions = 74%/78% [CEUS; observer 1/observer 2] versus 46/50 / 44/50 lesions = 92%/88% [MRI; observer 1/observer 2]). Washout appearance was observed in 34/50 / 20/50 lesions = 68%/40% with CEUS and 31/50 / 31/50 lesions = 62%/62% with MRI. Interobserver agreement was moderate for arterial hyperenhancement ($\kappa = 0.511/0.565$ [CEUS/MRI]) and “washout” ($\kappa = 0.490/0.582$ [CEUS/MRI]), fair for CEUS-LI-RADS category ($\kappa = 0.309$) and substantial for MRI-LI-RADS category ($\kappa = 0.609$). Intermodality agreement was fair for arterial hyperenhancement ($\kappa = 0.329$), slight to fair for “washout” ($\kappa = 0.202$) and LI-RADS category ($\kappa = 0.218$)

Conclusion Interobserver agreement is substantial for MRI-LI-RADS and only fair for CEUS-LI-RADS. This is mostly because interobserver agreement in the perception of washout appearance is better in MRI than in CEUS. Further refinement of the LI-RADS algorithms and increasing education and practice may be necessary to improve the concordance between CEUS and MRI for the final LI-RADS categorization.

Key Points

- CEUS-LI-RADS and MRI-LIRADS enable standardized non-invasive diagnosis of HCC in high-risk patients.
- With CEUS, interobserver agreement is better for arterial hyperenhancement than for “washout”.
- Interobserver agreement for major features is moderate for both CEUS and MRI.
- Interobserver agreement for LI-RADS category is substantial for MRI, and fair for CEUS.
- Interobserver-agreement for CEUS-LI-RADS will presumably improve with ongoing use of the algorithm.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-018-5379-1>) contains supplementary material, which is available to authorized users.

✉ Hannes Seuss
hannes.seuss@uk-erlangen.de

¹ Department of Internal Medicine 1, Friedrich-Alexander-Universität Erlangen-Nürnberg, University Hospital Erlangen, Ulmenweg 18, 91054 Erlangen, Bayern, Germany

² Department of Radiology, Friedrich-Alexander-Universität Erlangen-Nürnberg, University Hospital Erlangen, Maximiliansplatz 3, 91054 Erlangen, Bayern, Germany

Keywords Carcinoma, hepatocellular · Magnetic resonance imaging · Ultrasonography · Diagnostic techniques and procedures · Liver cirrhosis

Abbreviations

ACR	American College of Radiology
bh	Breath-hold
BMI	Body mass index
CE-CT	Contrast-enhanced computed tomography
CE-MRI	Contrast-enhanced magnetic resonance imaging
CEUS	Contrast-enhanced Ultrasound
cor	Coronal
CT	Computed tomography
DWI	Diffusion-weighted image
fs	Fat saturation
GRE	Gradient echo
HASTE	Half-Fourier acquisition single-shot turbo spin echo
HCC	Hepatocellular carcinoma
ICC	Intrahepatic cholangiocellular carcinoma
LI-RADS	Liver Imaging Reporting and Data System
LR	LI-RADS category
MRI	Magnetic resonance imaging
T	Tesla
T1w	T1-weighted
T2w	T2-weighted
TE	Echo time
TR	Repetition time
tra	Transversal
TSE	Turbo spin echo
VIBE	Volumetric-interpolated breath-hold examination

Introduction

Hepatocellular carcinoma (HCC) can be diagnosed non-invasively in high-risk patients owing to typical features on contrast-enhanced imaging. Hallmarks of HCC include arterial phase hyperenhancement followed by a gradual “washout” of the contrast agent during portal venous and late phase. However, this characteristic contrast enhancement pattern is not found in all cases [1–8].

Attempts to improve standardization in the interpretation, documentation and reporting of contrast-enhanced imaging led to the development of LI-RADS (Liver Imaging Reporting and Data System) by the American College of Radiology (ACR) in 2011 [9, 10]. With LI-RADS, focal liver lesions in high-risk patients are categorized according to “major features” (lesion diameter, arterial phase hyperenhancement, washout appearance, capsule appearance, threshold growth) and optionally using ancillary features. The probability of a lesion being an HCC is expressed by

assigning a LI-RADS category between LR-1 (definitely benign) and LR-5 (definitely HCC). LI-RADS is only defined for observations in contrast-enhanced computed tomography (CE-CT) and magnetic resonance imaging (CE-MRI).

German National Guidelines regard contrast-enhanced ultrasound (CEUS) as an imaging modality equivalent of CE-MRI and CE-CT in the non-invasive diagnosis of HCC in high-risk patients [11]. CEUS provides a unique real-time assessment of contrast enhancement patterns using contrast agents that remain strictly intravascular, allowing for very sensitive assessment of tumour vascularity. Several multicentre studies and meta-analyses have demonstrated the excellent diagnostic accuracy of CEUS in the differential diagnosis of focal liver lesions [12–21].

Standardized CEUS-based diagnostic algorithms in HCC such as ACR-CEUS-LI-RADS have not been developed until very recently [22–25].

To date, there are no studies evaluating these CEUS-based algorithms in a clinical setting. Only a few studies have assessed inter-reader agreement with LI-RADS in CT and MRI [26–32]. Thus, for the first time, this pilot study aimed to directly compare the diagnostic accuracy and interobserver agreement between ACR-CEUS-LI-RADS and MRI-LI-RADS.

Materials and methods

Study design

Figure 1 illustrates the design of this retrospective study. The risk population for HCC was defined according to national guidelines as patients with cirrhosis of any origin, chronic hepatitis B infection, chronic hepatitis C infection with advanced fibrotic changes, and histologically proven non-alcoholic steatohepatitis [11]. Inclusion criteria were at least one focal liver lesion visible on conventional ultrasound and availability of both CEUS and CE-MRI of the liver. All patients had both CEUS and CE-MRI within at least 3 months after initial detection of the lesion. Patients with prior systemic or local ablative treatment for HCC were excluded. All patients provided written informed consent for de-identified data evaluation. The study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the local ethics committee.

A systematic, retrospective search of the interdisciplinary liver cancer board was performed to identify high-risk patients. Patients’ electronic records were assessed to select those with CEUS examinations of the liver. Additionally, patients were identified when presenting for conventional liver ultrasound or CEUS, including patients undergoing HCC

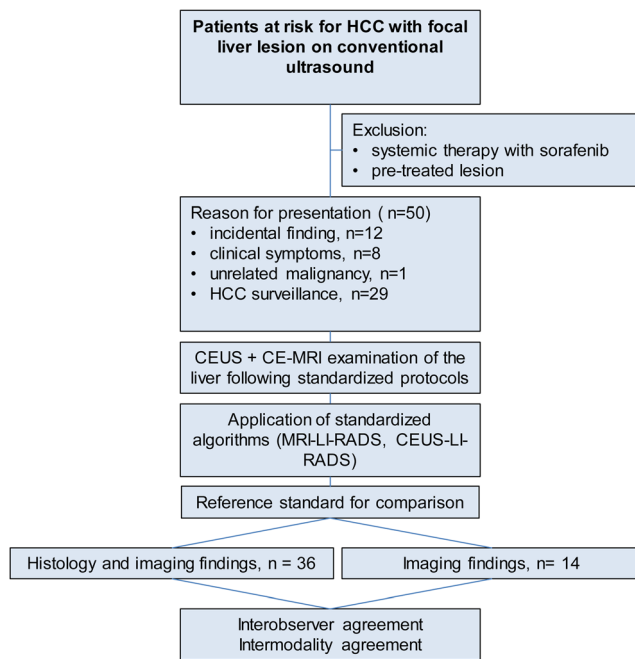


Fig. 1 Study design and patient selection

surveillance, symptomatic patients and patients with incidental focal liver lesions.

CEUS

CEUS examinations were performed according to European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines for the characterization of focal liver lesions following a standardized protocol with low mechanical index and intravenous bolus injection of 1.5 mL SonoVue® (Bracco Imaging, Konstanz, Germany) followed by a saline flush [12]. Video clips of the examinations were recorded over a time period of 3–5 min, beginning prior to the first arrival of detectable microbubbles and continuing until the beginning of clearance of microbubbles from parenchymal tissue in the late phase. Contrast enhancement patterns during arterial, portal venous and late phase were assessed. Vascular phases were defined according to EFSUMB guidelines (arterial phase, beginning within 20 s after injection of the contrast agent, duration until 30–45 s after injection, depending on the cardiocirculatory situation; portal venous phase, beginning 30–45 s after injection, ending about 120 s after injection; late phase, beginning 2 min after injection and lasting until the clearance of microbubbles from liver tissue) [12].

CEUS was performed using three different ultrasound devices (Siemens Acuson S2000, GE Logiq E9, Toshiba Aplio 500) by three physicians with 5–10 years of experience in liver ultrasound. All patients underwent conventional B-mode of the liver prior to CEUS. In cases with more than one lesion, all lesions were recorded. Only the lesion most accessible to ultrasound examination was chosen for CEUS

and assessment by two independent observers (not involved in the CEUS examination) using the standardized diagnostic algorithm CEUS LI-RADS.

MRI

MRI was performed following standardized protocols with two different 1.5-T MR scanners: one Magnetom Aera and one Magnetom Avanto (both Siemens Healthineers, Erlangen, Germany). A dedicated HCC protocol was used for each scanner. All images had a slice thickness of 5 mm. The spacing was 6 mm for two-dimensional sequences and 5 mm for three-dimensional acquisition. After the pre-contrast sequences, weight-adapted gadobutrol 1.0 mmol/mL (Gadovist®, Bayer Pharma AG, Berlin, Germany) was injected intravenously as a non-liver-specific contrast agent. T1w sequences were acquired in the arterial, venous and several post-CE phases. The matrix was roughly 260×320 pixels for all sequences except the diffusion-weighted images, for which it was 130×160 pixels.

The default protocols are presented in detail in Table 1.

The target lesion assessed with CEUS-LI-RADS was identified on MRI scans and evaluated by two observers using the standardized diagnostic algorithm MRI-LI-RADS. A lesion was measured on the axial T1w sequence it was best visible.

Standardized algorithms

CEUS-LI-RADS

CEUS examinations were evaluated according to ACR-CEUS-LI-RADSv.2016 (Supplemental Fig. 1) developed by the American College of Radiology. Prior to the study, observers received theoretical training (2 h) to become familiarized with the use of the CEUS-LI-RADS algorithm. The algorithm and its features were explained and example CEUS clips of all categories were reviewed. In a subsequent practical training phase, classification according to CEUS-LI-RADS was taught using five example lesions.

Two observers with 2 and 5 years of experience, respectively, in CEUS examinations of the liver blinded to patients' clinical data and final diagnosis (except for knowledge of a high-risk constellation) independently reviewed the CEUS examinations and assigned a CEUS-LI-RADS category to the target observation. Observers assessed uptake of contrast agent in the target observation relative to the surrounding parenchyma. They decided on hyper-, iso- or hypoenhancement of the target observation in the arterial, portal venous and late phase and, if available, very late phase (> 240 s). In case of “washout”, observers were asked to distinguish between early “washout” (starting < 60 s) and late “washout” (≥ 60 s).

Table 1 Magnetic resonance imaging protocols

Sequence	Contrast	TR (ms)	TE (ms)	TE (ms)	TE (ms)
Scanner (1 / 2)		1	2	1	2
T2w HASTE tra bh		900	638	122	119
T2w (HASTE / TSE) tra bh fs		1000	5217	90	102
T2w HASTE bh cor		1000	1060	122	119
DWI/ADC tra b50 to b800	None	3294	2000	72	56
T1w (VIBE / VIBE Dixon) GRE tra in		7.96	6.65	4.76	4.77
T1w (VIBE / VIBE Dixon) GRE tra opposed		7.96	6.65	2.38	2.39
T1w (VIBE / VIBE Dixon) GRE fs tra bh		5.42	6.65	2.39	2.39
	Arterial	5.42	6.65	2.39	2.39
	Venous	5.42	6.65	2.39	2.39
		5.42	6.65	2.39	2.39
T1w (VIBE / VIBE Dixon) GRE fs cor bh		5.6	6.65	2.56	2.39
T1w FLASH fs tra bh	Delayed	130	169	7.15	7.15

Magnetic resonance sequences with their respective contrast phase, relaxation time (TR) and echo time (TE). Differences between the sequences of the two scanners are put in parentheses. The protocol for scanner 1 (Magentom Avanto, Siemens Healthineers, Erlangen, Germany) consisted of a T2-weighted (T2w) half-Fourier acquisition single-shot turbo spin echo (HASTE) with transversal (tra) image orientation, acquired during one single breath-hold (bh), followed by a T2w HASTE tra bh with fat saturation (fs), a T2w HASTE bh in coronal (cor) and a diffusion weighted image (DWI) sequence in tra orientation with *b* values ranging from 50 to 800. The apparent diffusion coefficient was calculated from the DWI. These sequences were followed by a T1-weighted (T1w) volumetric-interpolated breath-hold examination (VIBE) gradient echo (GRE) sequence in transversal orientation in an opposed-phase (opp) and an in-phase (in) manner. Subsequently, a series of T1w VIBE GRE fs tra bh sequences was acquired without contrast and in the arterial, venous and post-contrast enhanced (CE) phase. Further post-CE phases included a T1w VIBE GRE fs cor bh as well as a high resolution T1w fast low angle shot (FLASH) fs tra bh sequence. In the protocol for scanner 2 (Magnetom Aera, Siemens Healthineers) the T2w HASTE fs tra bh was replaced by a T2w turbo spin echo (TSE) fs tra bh sequence. The T1w VIBE GRE sequences were replaced by T1w VIBE DIXON sequences

MRI-LI-RADS

MRI examinations were evaluated according to LI-RADSv.2014 (Supplemental Fig. 2). One certified radiologist with 7 years and one resident with 4 years of experience in hepatobiliary imaging blinded to patients' clinical data and final diagnosis (except for knowledge of a high-risk constellation, location and size of the target observation) independently reviewed the MRI examinations and assigned a LI-RADS category to the target observation.

MRI observers did not receive a particular training session; MRI-LI-RADS has been implemented for routine diagnosis in suspected HCC lesions in our department for 2 years.

Ancillary features and tie-breaking rules were not used. As MRI scans were available for one examination per patient, the major feature of "threshold growth" could not be assessed.

Reference standard

Final diagnosis was based on histology or, if histological findings were not available, on characteristic findings upon contrast-enhanced imaging and, in cases of benign lesions, constant appearance and lack of interval growth during follow-up imaging. Histological findings were

obtained via ultrasound-guided biopsy in 30 cases (core biopsy in 29 cases, fine needle biopsy in one case), CT-guided biopsy in two cases, and surgical resection in four cases. For biopsy, a mean of 2.7 separate biopsies (range, 1–6) was taken at the decision of the examiner. Mean length of total tissue samples available for one patient was 38 mm (range, 4–97 mm). All histological diagnoses were made by two expert pathologists in consensus.

Statistical analysis

Quantitative variables are expressed as a mean \pm standard deviation. Categorical variables are expressed as frequencies. Groups were compared using Fisher's exact test. Cohen's κ statistics were used for the evaluation of interobserver agreement. Results were interpreted as follows: $\kappa = 0.81$ – 1.00 , (almost) perfect agreement; $\kappa = 0.61$ – 0.80 , substantial agreement; $\kappa = 0.41$ – 0.60 , moderate agreement; $\kappa = 0.21$ – 0.40 , fair agreement; $\kappa \leq 0.20$, slight agreement. SPSS-21 (IBM Corporation, Armonk, NY, USA) and Excel 2010 (Microsoft Corporation, Redmond, Washington, USA) were used for statistical analyzes. Differences were considered statistically significant for $p < 0.05$.

Table 2 Patient characteristics ($n = 50$)

Age [years] (mean \pm SD/ range)	66.2 \pm 11.8 (53–86)	
Male/female (n , %)	39/11	(78.0%/22%)
Clinical or histological findings of cirrhosis	45	(90.0%)
Fibrosis (Ishak 2–3)	4	(8.0%)
Child–Pugh stage ($n = 45$)	A	35 (77.8%)
	B	8 (17.8%)
	C	2 (4.4%)

SD standard deviation, n number

Results

Patient and tumour characteristics

Patient and tumour characteristics are shown in Tables 2 and 3. Of 50 lesions, 43 were HCCs (86%); two were intrahepatic cholangiocellular carcinomas (ICCs); five were benign lesions. Histological findings were available in 36/50 lesions (72%; 32 HCCs, two ICCs, two regenerate/dysplastic nodules).

Table 3 Tumour characteristics

Lesion characteristics ($n = 50$)		
Size of tumour [mm] (mean \pm SD / range)	40.9 \pm 28.3 (14–150)	
≥ 20 mm	45	(90.0%)
Tumour number (n /%)	Single	27 (54.0%)
	2–3	12 (24.0%)
	> 3	11 (22.0%)
Tumour entity	HCC	43 (86.0%)
	Non-HCC	7 (14.0%)
HCC characteristics ($n = 43$)		
Size of HCC [mm] (mean \pm SD / range)	42.9 \pm 29.2 (14–150)	
≥ 20 mm	40	(93.0%)
HCCs within Milan criteria	26	(60.5%)
BCLC stages	A	26 (60.5%)
	B	11 (24.4%)
	C	5 (11.6%)
	D	1 (2.3%)
Histological grading (n /%)	Available	32 (74.4%)
	G1	8 (25.0%)
	G2	20 (62.5%)
	G3	4 (12.5%)
Diffuse infiltration	4	(9.5%)

Final diagnosis was based on histological findings, if available, or on findings upon contrast-enhanced imaging according to German national HCC guidelines and, in cases of benign lesions, constant findings over at least 6 months

HCC hepatocellular carcinoma, BCLC Barcelona Clinic Liver Cancer, SD standard deviation, n number, mm millimetres

Of the benign lesions, three were regenerate/dysplastic nodules; one was a cyst; and one was a focal fat sparing. Three HCCs and two benign lesions were less than 20 mm in diameter. Mean lesion size on conventional ultrasound was 30 \pm 16.6 mm (range, 16–69 mm) for non-HCC lesions, versus 42.9 \pm 29.2 mm (range, 14–150 mm) for HCC lesions. In 31 cases, lesion size was larger on ultrasound than on MRI; in 19 cases, lesions were measured with greater size on MRI. However, there were no cases where different measurements upon ultrasound or MRI would have led to a discrepancy in size category as < 20 mm or \geq 20 mm. Thus, slight differences in size measurements between imaging modalities did not affect LI-RADS categorization.

Concordant/discordant findings for CEUS-LI-RADS versus MRI-LI-RADS

Perception of major features differed between CEUS and MRI, although statistical significance was reached only for the perception of arterial phase hyperenhancement, but not washout appearance. Arterial phase hyperenhancement was observed in 76% of lesions ($n = 38/50$) with CEUS and 90% of lesions ($n = 45/50$) with MRI ($p = 0.038$; mean values from both observers); intermodality agreement between CEUS and MRI for the perception of arterial phase hyperenhancement was fair ($\kappa = 0.329$). “Washout” was seen in 54% of lesions ($n = 27/50$) in CEUS and 62% of lesions ($n = 31/50$) in MRI ($p = 0.420$); intermodality agreement for washout appearance was slight to fair ($\kappa = 0.202$).

With CEUS-LI-RADS, a considerable proportion of the 32 histologically proven HCCs were classified as LR-4 (8/32 versus 16/32 for observer 1 versus observer 2) and LR-3 (3/32 versus 1/32). With MRI-LI-RADS, 4/32 (12.5%) (observer 1) versus 6/32 (18.8%) (observer 2) histologically proven HCCs were categorized as LR-3, LR-4 or LR-M. In detail, one HCC was categorized as LR-3, two as LR-4 and one as LR-M by observer 1; one HCC was categorized as LR-3, four as LR-4 and one as LR-M by observer 2.

Intermodality agreement for CEUS and MRI for LI-RADS category was slight to fair ($\kappa = 0.218$). A direct comparison of LI-RADS classification with CEUS and MRI is presented in Table 4.

Examples of LI-RADS categories in CEUS and MRI are shown in Figs. 2, 3 and 4.

Interobserver agreement

CEUS-LI-RADS

There was no discordance between the two observers for the distinction between early “washout” (< 60 s) and late “washout” (\geq 60 s). Early “washout” (< 60 s) was not perceived in any of the lesions in the study collective by either observer. The two observers perceived arterial phase

Table 4 Comparison of CEUS-LI-RADS and MRI-LI-RADS category (combined values from both observers of every modality)

		CEUS						Total
		LR-1	LR-2	LR-3	LR-4	LR-5/5V	LR-M	
MRI	LR-1	4	0	0	0	0	0	4
	LR-2	0	0	0	0	0	0	0
	LR-3	0	0	5	2	1	0	8
	LR-4	0	0	4	7	15	0	26
	LR-5/5V	0	0	15	57	86	0	158
	LR-M	0	0	2	2	0	0	4
	Total	4	0	26	68	102	0	200

Intermodality agreement between MRI and CEUS. For every lesion the assessments of the two MRI observers and the two CEUS observers were compared. Therefore 2×2 comparisons were performed for each patient ($n = 2 \times 2 \times 50 = 200$). In the lower left (dark grey), CEUS under-categorizes ($n = 80 / 40\%$) and in the upper right (light grey), it over-categorizes ($n = 18 / 9\%$) findings compared to MRI. In 102 (51%) comparisons observations of both modalities agree upon the LI-RADS category

CEUS contrast-enhanced ultrasound, MRI magnetic resonance imaging

hyperenhancement in 74%/76% of cases (37/50 versus 38/50 cases; $p = 0.640$) and “washout” in 68%/40% (34/50 versus 20/50 cases; $p = 0.005$).

Interobserver agreement according to Cohen’s kappa was moderate for arterial phase hyperenhancement ($\kappa = 0.511$) and washout appearance ($\kappa = 0.490$), and fair for the CEUS-LI-RADS category ($\kappa = 0.309$) (Table 5).

MRI-LI-RADS

The two observers perceived capsule appearance in 47/50 versus 40/50 cases (94%/80%; $p = 0.037$), arterial phase hyperenhancement in 46/50 versus 44/50 cases (92%/88%; $p = 0.505$) and “washout” in 31/50 versus 31/50 cases (62%/62%; $p = 1.000$).

Interobserver agreement for MRI-LI-RADS was moderate for capsule appearance ($\kappa = 0.449$), arterial phase hyperenhancement ($\kappa = 0.565$) and washout appearance ($\kappa = 0.582$), and substantial for the LI-RADS category ($\kappa = 0.609$, Table 5).

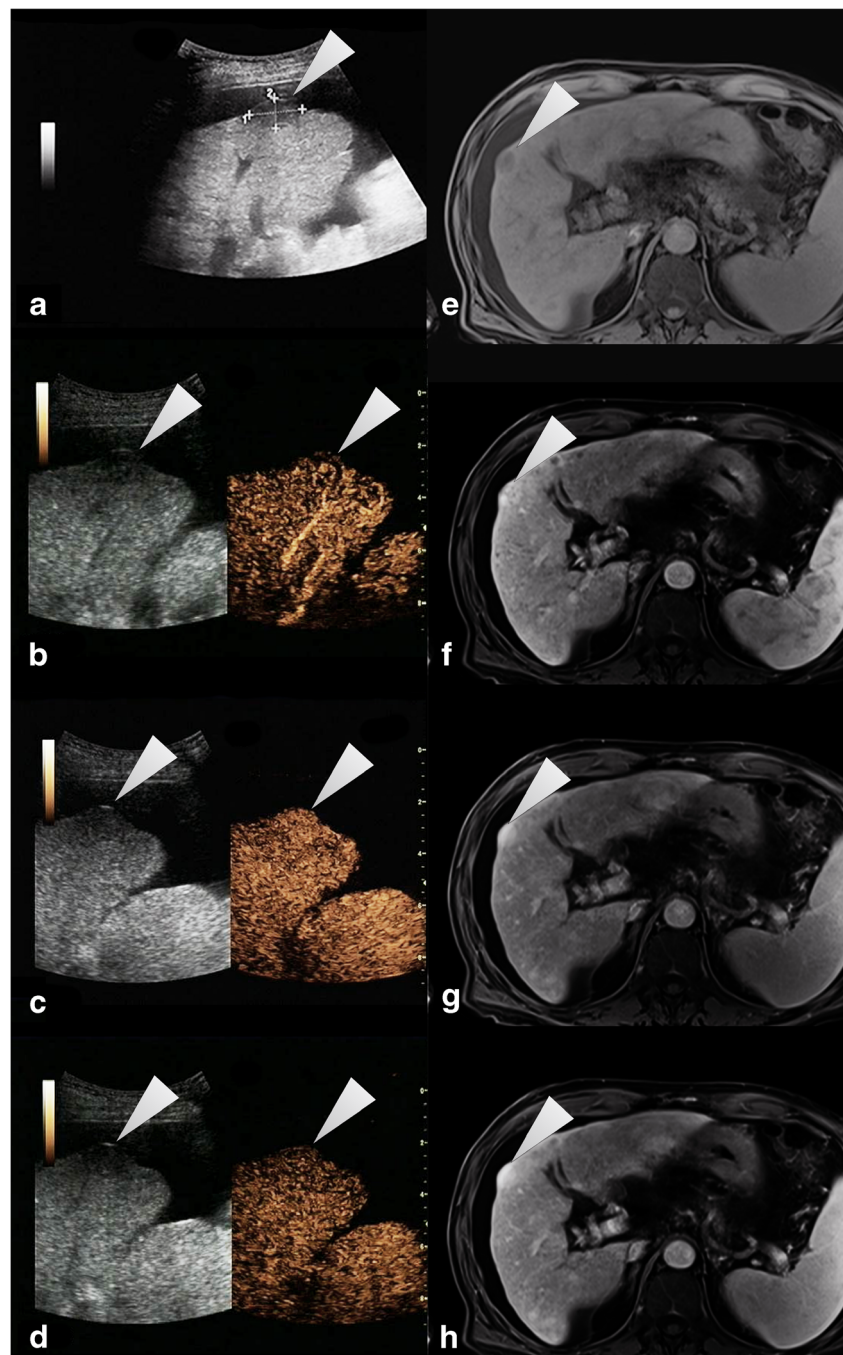
Discussion

Our study is the first one to assess interobserver agreement for MRI-LI-RADS and CEUS-LI-RADS in direct comparison. We found moderate interobserver agreement for MRI-LI-RADS for all three major features (arterial phase hyperenhancement, washout appearance, capsule appearance) and substantial agreement for MRI-LI-RADS category. With CEUS-LI-RADS we found moderate interobserver agreement for arterial phase hyperenhancement ($\kappa = 0.511$) and washout appearance ($\kappa = 0.490$) and only fair agreement concerning CEUS-

LI-RADS category ($\kappa = 0.309$). Importantly, intermodality agreement between CEUS and MRI was only fair for arterial phase hyperenhancement ($\kappa = 0.329$), slight to fair for “washout” ($\kappa = 0.202$) and slight to fair for LI-RADS category ($\kappa = 0.218$). The fact that interobserver agreement for the final category was substantial for MRI and only fair for CEUS although interobserver agreement for major features was moderate for both modalities might be due to the fact that “washout” plays a more important role in MRI than in CEUS.

The results of this study demonstrate considerable discrepancy between MRI and CEUS in terms of major feature assessment and the final LI-RADS category assignment. We found that this is mostly due to the fact that perception of arterial phase hyperenhancement differs between CEUS and MRI. With the LR-5 category definition in CEUS-LI-RADS in its current version, a relevant proportion of HCCs is categorized as LR-4 or LR-3. With CEUS-LI-RADS, a lesion cannot possibly be categorized as definite HCC if it is either lacking arterial phase hyperenhancement or contrast “washout”. This results in the fact that a substantial proportion of HCCs is categorized as LR-4. These findings raise the question of whether “washout” should be a necessary prerequisite for the categorization of a lesion as HCC in CEUS. Contrast agents in CEUS differ from those used in MRI in that they remain strictly intravascular. Therefore, “washout” in CEUS cannot be equated with “washout” in MRI. There is evidence from the literature that “washout” should not be mandatory for the non-invasive diagnosis of HCC with CEUS in cirrhotic patients [8, 33]. In our study detection of “washout” differed significantly between observers with CEUS, but not with MRI, indicating the limitation of “washout” as a major criterion in CEUS

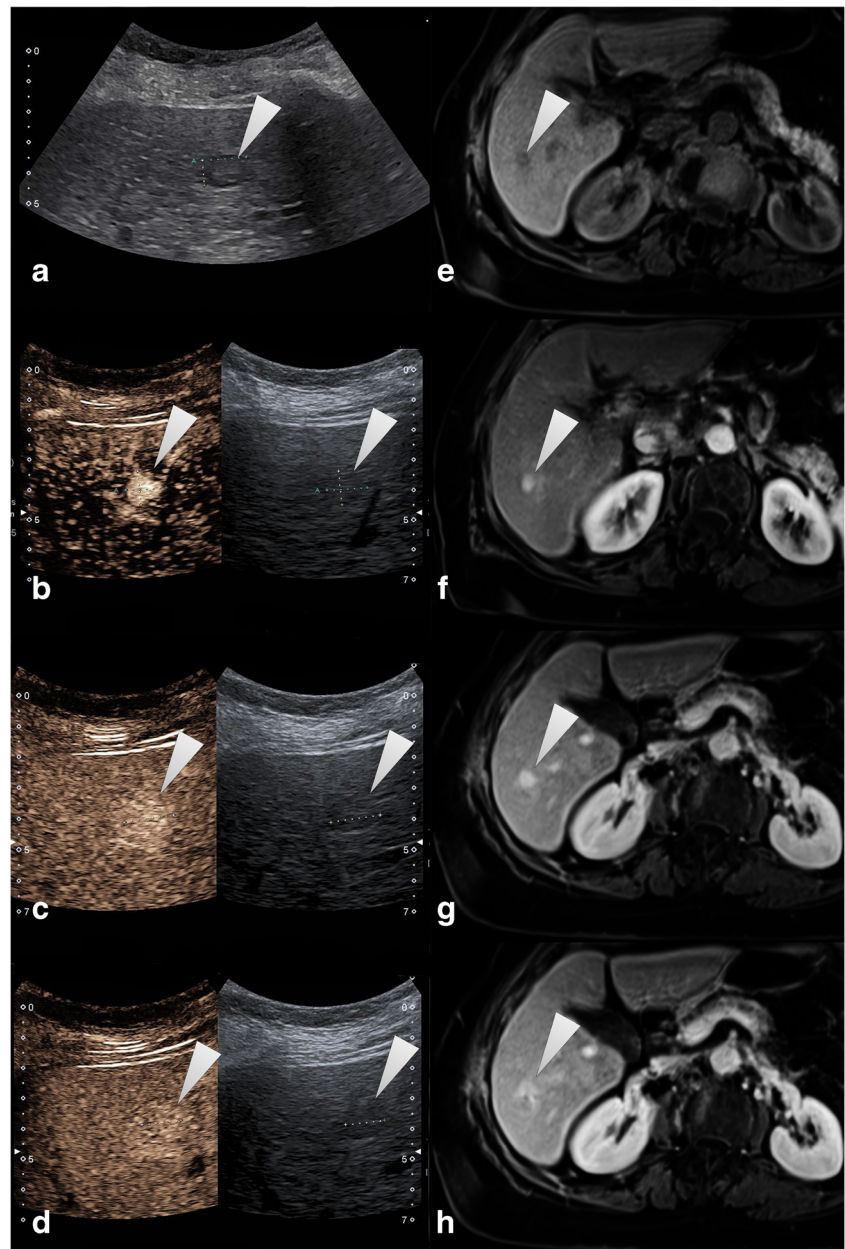
Fig. 2 Exemplary LI-RADS-3 lesion in both CEUS and MRI. Atypical HCC (18 mm; white arrowhead). a B-mode ultrasound shows a hypochoic lesion. CEUS: b arterial isoenhancement, c portal-venous isoenhancement and d no “washout” in the delayed phase. T1w VIBE fat-sat: e non-contrast phase, f arterial phase with weak hyperenhancement, g portal-venous phase with weak hyperenhancement and h no measureable washout in the delayed phase



Another reason for the higher agreement of MRI-LI-RADS in comparison to CEUS-LI-RADS in our study can be seen in the issue of learning curves. CEUS-LI-RADS is a very recent development, whereas MRI-LI-RADS has been widely adopted since its first release in 2011. In our department, MRI-LI-RADS has been

routinely used for several years by the MRI observers, whereas the CEUS observers needed a special training session prior to the study to become familiar with the use of CEUS-LI-RADS. Correspondingly, Quaia et al. showed that interobserver agreement for the assessment of hyper- or hypoenhancement of focal liver lesions in

Fig. 3 Exemplary LI-RADS-4 lesion in CEUS. For MRI tie-breaking rules have to be applied between LR-4 and LR-5. In this study no follow-up examination or ultrasound was available, therefore the lesion was classified LR-4. With threshold growth (LR-5g) or “washout” and visibility in ultrasound (LR-5us) it would have been upgraded. Small HCC (14 mm; white arrowhead). a Hypoechoic lesion in B-mode ultrasound. CEUS: b homogeneous arterial hyperenhancement, c sustained portal-venous hyperenhancement and d no “washout” in the delayed phase. T1w VIBE fat-sat: e hypointense lesion in the non-contrast phase, f arterial hyperenhancement, g portal-venous hyperenhancement and h “washout” in the delayed phase



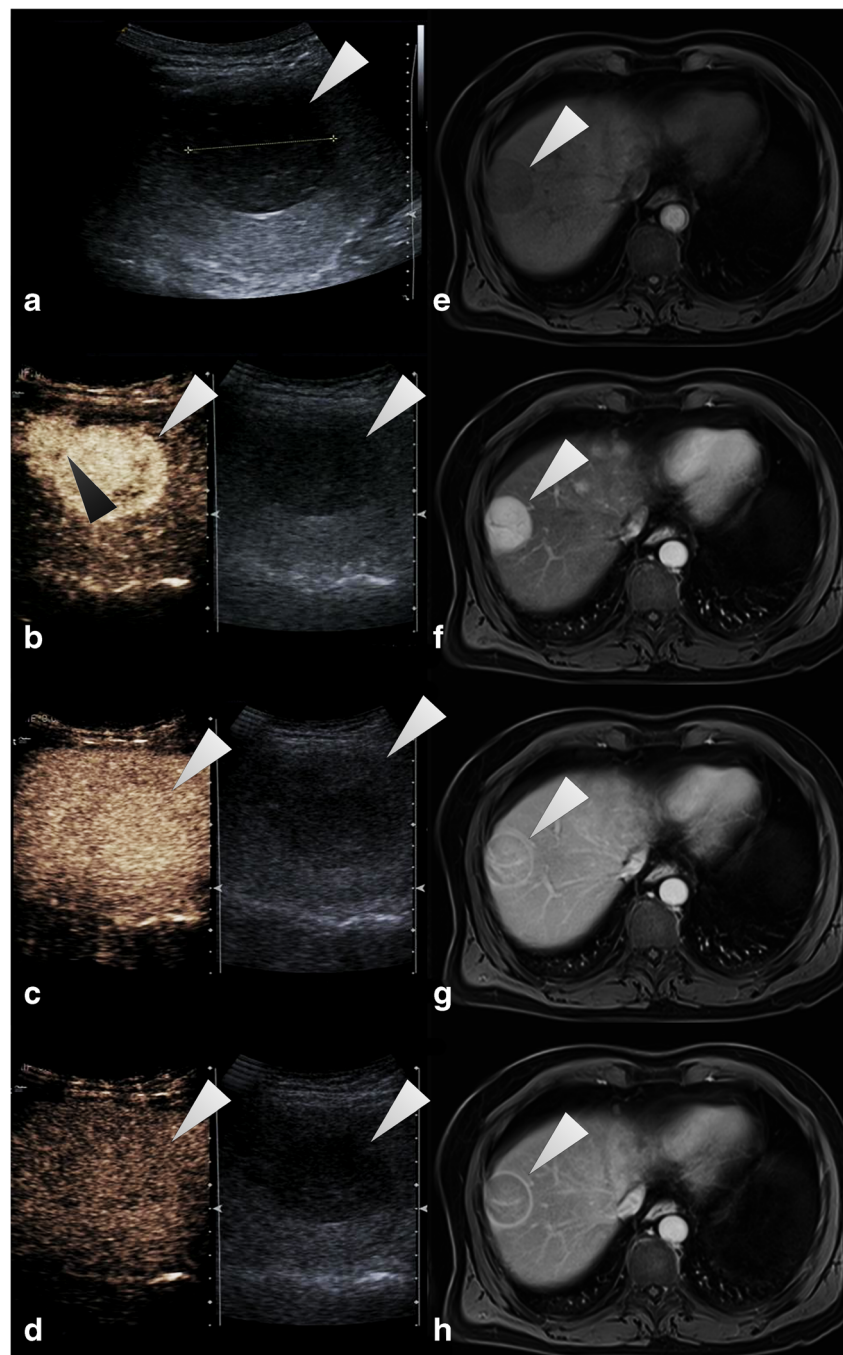
CEUS was moderate ($\kappa = 0.47\text{--}0.63$), but was better in experienced readers [34]. It might thus be expected that interobserver agreement will improve along with increasing application of the algorithm as has been shown for other “RADS” algorithms such as Breast Imaging Reporting and Data System [30, 35].

To date, there are no studies assessing interobserver agreement for standardized CEUS-based algorithms, and only a few studies addressing this issue for LI-RADS in

MRI or CT. Thus, our results are not directly comparable to the literature. Studies assessing interobserver agreement for MRI-LI-RADS found values for Cohen’s kappa between 0.35 and 0.44 [30, 32], with strongest interobserver agreement for arterial phase hyperenhancement [30]. For LI-RADS-CT, κ values between 0.56 and 0.69 are reported [36–38].

However, our results suggest best intermodality agreement between CEUS and MRI for arterial phase hyperenhancement.

Fig. 4 Exemplary LI-RADS-5 lesion. Large HCC (46 mm; white arrowhead). a Hypoechoic lesion in B-mode ultrasound. CEUS: b arterial hyperenhancement with an adjacent, now clearly visible lesion (black arrowhead), c slight “washout” in the portal-venous phase and d clear “washout” in the delayed phase. T1w VIBE fat-sat: e hypointense lesion in the non-contrast phase, f strong arterial hyperenhancement, g “washout” and hyperenhancing capsula in the portal-venous and h delayed phase



This emphasizes the point that arterial phase hyperenhancement should be regarded as the CEUS key imaging feature of HCC in high-risk patients, whereas the diagnostic value of “washout” in CEUS deserves further investigation.

Our study has some limitations. These are the relatively small sample size, the retrospective nature and the

single-centre design. However, the work was intended as a pilot study, which (for the first time) evaluates interobserver and intermodality agreement of CEUS-LI-RADS and MRI-LI-RADS in direct comparison. Another major limitation is the design with different observers for CEUS and MRI. The agreement might have

Table 5 Interobserver agreement for major features and lesion categories

	Cohen's κ	Proportion of agreement	
CEUS			
Arterial phase hyperenhancement	0.511	84%	(42/50)
Washout appearance	0.490	80%	(40/50)
CEUS-LI-RADS category	0.309	68%	(34/50)
MRI			
Capsule appearance	0.449	82%	(41/50)
Arterial phase hyperenhancement	0.565	92%	(46/50)
Washout appearance	0.582	80%	(40/50)
MRI-LI-RADS category	0.609	86%	(43/50)

Interobserver agreement according to Cohen's kappa and percentage agreement for major features and categories in CEUS-LI-RADS and MRI-LI-RADS. A Cohen's kappa value of 0.201–0.400 is considered as fair, 0.401–0.600 as moderate and 0.601–0.800 as substantial agreement

CEUS contrast-enhanced ultrasound, MRI magnetic resonance imaging, LI-RADS Liver Imaging Reporting and Data System

been stronger if the same observers had interpreted both CEUS and MRI.

In conclusion, the interobserver agreement for major features was moderate for both CEUS and MRI. The interobserver agreement for the final LI-RADS category was substantial for MRI and only fair for CEUS. Furthermore, intermodality agreement for the final LI-RADS category between MRI and CEUS was only slight to fair. Further refinement of the LI-RADS algorithms and increasing education and practice may be necessary to improve the concordance between CEUS and MRI for the final LI-RADS categorization.

Acknowledgements We thank the editors of *European Radiology* and those who reviewed this article.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Dr. Hannes Seuss.

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

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was waived by the institutional review board.

Ethical approval Institutional review board approval was obtained.

Methodology

- retrospective
- diagnostic study
- performed at one institution

References

1. Westwood M, Joore M, Grutters J et al (2013) Contrast-enhanced ultrasound using SonoVue(R) (sulphur hexafluoride microbubbles) compared with contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging for the characterisation of focal liver lesions and detection of liver metastases: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 17:1–243
2. 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. *Ultrasound Med Biol* 39:187–210
3. Bota S, Piscaglia F, Marinelli S, Pecorelli A, Terzi E, Bolondi L (2012) Comparison of international guidelines for noninvasive diagnosis of hepatocellular carcinoma. *Liver Cancer* 1:190–200
4. Bruix J, Sherman M (2011) Management of hepatocellular carcinoma: an update. *Hepatology* 53:1020–1022
5. Kokudo N, Hasegawa K, Akahane M et al (2015) Evidence-based clinical practice guidelines for hepatocellular carcinoma: The Japan Society of Hepatology 2013 update (3rd JSH-HCC guidelines). *Hepatol Res* 45. <https://doi.org/10.1111/hepr.12464>
6. Friedrich-Rust M, Klopffleisch T, Nierhoff J et al (2013) Contrast-enhanced ultrasound for the differentiation of benign and malignant focal liver lesions: a meta-analysis. *Liver Int* 33:739–755
7. Shah S, Shukla A, Paunipagar B (2014) Radiological features of hepatocellular carcinoma. *J Clin Exp Hepatol* 4:S63–S66
8. Leoni S, Piscaglia F, Granito A et al (2013) Characterization of primary and recurrent nodules in liver cirrhosis using contrast-enhanced ultrasound: which vascular criteria should be adopted? *Ultraschall Med* 34:280–287
9. Mitchell DG, Bruix J, Sherman M, Sirlin CB (2015) LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. *Hepatology* 61:1056–1065
10. Elsayes KM, Kielar AZ, Agrons MM et al (2017) Liver Imaging Reporting and Data System: an expert consensus statement. *J Hepatocell Carcinoma* 4:29–39
11. AWMF-Leitlinie S3 - Leitlinie Deutschland: "Hepatozelluläres Karzinom", Registernummer 032/053OL Stand: 01.05.2013, gültig bis 30.04.2018; http://www.awmf.org/uploads/tx_szleitlinien/032-053OL1_S3_Hepatozellul%C3%A4res_Karzinom_Diagnostik_Therapie_2013-verlaengert.pdf accessed: May 19 2017. Available via http://www.awmf.org/uploads/tx_szleitlinien/032-053OL1_S3_Hepatozellul%C3%A4res_Karzinom_Diagnostik_Therapie_2013-verlaengert.pdf. Accessed May 19 2017
12. 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
13. Strobel D, Bernatik T, Blank W et al (2011) Diagnostic accuracy of CEUS in the differential diagnosis of small (\leq 20 mm) and subcentimetric (\leq 10 mm) focal liver lesions in comparison with histology. Results of the DEGUM multicenter trial. *Ultraschall Med* 32:593–597

14. Strobel D, Seitz K, Blank W et al (2009) Tumor-specific vascularization pattern of liver metastasis, hepatocellular carcinoma, hemangioma and focal nodular hyperplasia in the differential diagnosis of 1,349 liver lesions in contrast-enhanced ultrasound (CEUS). *Ultraschall Med* 30:376–382
15. Seitz K, Greis C, Schuler A et al (2011) Frequency of tumor entities among liver tumors of unclear etiology initially detected by sonography in the noncirrhotic or cirrhotic livers of 1349 patients. Results of the DEGUM multicenter study. *Ultraschall Med* 32:598–603
16. Darnell A, Forner A, Rimola J et al (2015) Liver Imaging Reporting and Data System with MR imaging: evaluation in nodules 20 mm or smaller detected in cirrhosis at screening US. *Radiology* 275:698–707
17. Seitz K, Piscaglia F (2013) Ultrasound: the only "one stop shop" for modern management of liver disease. *Ultraschall Med* 34:500–503
18. Wildner D, Bernatik T, Greis C, Seitz K, Neurath MF, Strobel D (2015) CEUS in hepatocellular carcinoma and intrahepatic cholangiocellular carcinoma in 320 patients - early or late washout matters: a subanalysis of the DEGUM multicenter trial. *Ultraschall Med* 36:132–139
19. Wildner D, Pfeifer L, Goertz RS et al (2014) Dynamic contrast-enhanced ultrasound (DCE-US) for the characterization of hepatocellular carcinoma and cholangiocellular carcinoma. *Ultraschall Med* 35:522–527
20. McEvoy SH, McCarthy CJ, Lavelle LP et al (2013) Hepatocellular carcinoma: illustrated guide to systematic radiologic diagnosis and staging according to guidelines of the American Association for the Study of Liver Diseases. *Radiographics* 33:1653–1668
21. Seitz K, Bernatik T, Strobel D et al (2010) Contrast-enhanced ultrasound (CEUS) for the characterization of focal liver lesions in clinical practice (DEGUM Multicenter Trial): CEUS vs. MRI—a prospective comparison in 269 patients. *Ultraschall Med* 31:492–499
22. Schellhaas B, Wildner D, Pfeifer L et al (2016) LI-RADS-CEUS - proposal for a contrast-enhanced ultrasound algorithm for the diagnosis of hepatocellular carcinoma in high-risk populations. *Ultraschall Med* 37:627–634
23. Karlas T (2017) LI-RADS-CEUS for the classification of HCC risk in liver lesions. *Z Gastroenterol* 55:507–508
24. Piscaglia F, Wilson SR, Lyshchik A et al (2017) American College of Radiology Contrast Enhanced Ultrasound Liver Imaging Reporting and Data System (CEUS LI-RADS) for the diagnosis of hepatocellular carcinoma: a pictorial essay. *Ultraschall Med*. <https://doi.org/10.1055/s-0042-124661>
25. ACR American College of Radiology: Liver Imaging Reporting and Data System <https://www.acr.org/Quality-Safety/Resources/LIRADS>. Accessed 12 Sep 2017
26. Zhang YD, Zhu FP, Xu X et al (2016) Liver Imaging Reporting and Data System: substantial discordance between CT and MR for imaging classification of hepatic nodules. *Acad Radiol* 23:344–352
27. Corwin MT, Fananapazir G, Jin M, Lamba R, Bashir MR (2016) Differences in liver imaging and reporting data system categorization between MRI and CT. *AJR Am J Roentgenol* 206:307–312
28. Chernyak V, Flusberg M, Law A, Kobi M, Paroder V, Rozenblit AM (2017) Liver Imaging Reporting and Data System: discordance between computed tomography and gadoxetate-enhanced magnetic resonance imaging for detection of hepatocellular carcinoma major features. *J Comput Assist Tomogr*. <https://doi.org/10.1097/rct.0000000000000642>
29. Hope TA, Aslam R, Weinstein S et al (2017) Change in Liver Imaging Reporting and Data System characterization of focal liver lesions using gadoxetate disodium magnetic resonance imaging compared with contrast-enhanced computed tomography. *J Comput Assist Tomogr* 41:376–381
30. Barth BK, Donati OF, Fischer MA et al (2016) Reliability, validity, and reader acceptance of LI-RADS-An in-depth analysis. *Acad Radiol* 23:1145–1153
31. Bashir MR, Huang R, Mayes N et al (2015) 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* 42:305–314
32. Davenport MS, Khalatbari S, Liu PS et al (2014) Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging. *Radiology* 272:132–142
33. Giorgio A, Montesarchio L, Gatti P et al (2016) Contrast-enhanced ultrasound: a simple and effective tool in defining a rapid diagnostic work-up for small nodules detected in cirrhotic patients during surveillance. *J Gastrointest Liver Dis* 25:205–211
34. Quaia E, Alaimo V, Baratella E et al (2010) Effect of observer experience in the differentiation between benign and malignant liver tumors after ultrasound contrast agent injection. *J Ultrasound Med* 29:25–36
35. Berg WA, D'Orsi CJ, Jackson VP et al (2002) Does training in the Breast Imaging Reporting and Data System (BI-RADS) improve biopsy recommendations or feature analysis agreement with experienced breast imagers at mammography? *Radiology* 224:871–880
36. Zhang YD, Zhu FP, Xu X et al (2016) 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* 43:373–383
37. Ehman EC, Behr SC, Umetsu SE et al (2016) Rate of observation and inter-observer agreement for LI-RADS major features at CT and MRI in 184 pathology proven hepatocellular carcinomas. *Abdom Radiol (NY)* 41:963–969
38. McHugh ML (2012) Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 22:276–282