

Interobserver Agreement for Contrast-Enhanced Ultrasound (CEUS)-Based Standardized Algorithms for the Diagnosis of Hepatocellular Carcinoma in High-Risk Patients

Interobserver-Variabilität von Kontrastmittelultraschall (CEUS)-basierten standardisierten Diagnosealgorithmen für die Diagnostik des hepatozellulären Karzinoms bei Hochrisikopatienten

Authors

Barbara Schellhaas, Lukas Pfeifer, Christian Kielisch, Ruediger Stephan Goertz, Markus F. Neurath, Deike Strobel

Affiliation

Department of Internal Medicine 1, University of Erlangen, Germany

Key words

CEUS, LI-RADS, ESCULAP, HCC, interobserver agreement

received 08.09.2017

accepted 09.04.2018

Bibliography

DOI <https://doi.org/10.1055/a-0612-7887>

Published online: June 7, 2018

Ultraschall in Med 2018; 39: 667–674

© Georg Thieme Verlag KG, Stuttgart · New York

ISSN 0172-4614

Correspondence

Dr. Barbara Schellhaas

Department of Internal Medicine 1, University Hospital Erlangen, Ulmenweg 18, 91054 Erlangen, Germany

Tel.: ++49/91 31/8 53 50 00

barbara.schellhaas@uk-erlangen.de

ABSTRACT

Objectives This pilot study aimed at assessing interobserver agreement with two contrast-enhanced ultrasound (CEUS) algorithms for the diagnosis of hepatocellular carcinoma (HCC) in high-risk patients.

Methods Focal liver lesions in 55 high-risk patients were assessed independently by three blinded observers with two standardized CEUS algorithms: ESCULAP (Erlanger Synopsis of Contrast-Enhanced Ultrasound for Liver Lesion Assessment in Patients at risk) and ACR-CEUS-LI-RADSv.2016 (American College of Radiology CEUS-Liver Imaging Reporting and Data System). Lesions were categorized according to size and ultrasound contrast enhancement in the arterial, portal-venous and late phase. Interobserver agreement for assessment of enhancement pattern and categorization was compared between both CEUS algorithms. Additionally, diagnostic accuracy for the definitive diagnosis of HCC was compared. Histol-

ogy and/or CE-MRI and follow-up served as reference standards.

Results 55 patients were included in the study (male/female, 44/ 11; mean age: 65.9 years). 90.9% had cirrhosis. Histological findings were available in 39/55 lesions (70.9%). Reference standard of the 55 lesions revealed 48 HCCs, 2 intrahepatic cholangiocellular carcinomas (ICCs), and 5 non-HCC-non-ICC lesions. Interobserver agreement was moderate to substantial for arterial phase hyperenhancement ($\kappa = 0.53 - 0.67$), and fair to moderate for contrast washout in the portal-venous or late phase ($\kappa = 0.33 - 0.53$). Concerning the CEUS-based algorithms, the interreader agreement was substantial for the ESCULAP category ($\kappa = 0.64 - 0.68$) and fair for the CEUS-LI-RADS® category ($\kappa = 0.3 - 0.39$). Disagreement between observers was mostly due to different perception of washout.

Conclusion Interobserver agreement is better for ESCULAP than for CEUS-LI-RADS®. This is mostly due to the fact that perception of contrast washout varies between different observers. However, interobserver agreement is good for arterial phase hyperenhancement, which is the key diagnostic feature for the diagnosis of HCC with CEUS in the cirrhotic liver.

ZUSAMMENFASSUNG

Ziele Untersuchung der Interobserver-Variabilität zweier Kontrastmittelultraschall (CEUS) -basierter Diagnosealgorithmen für die Diagnostik des hepatozellulären Karzinoms (HCC) bei Hochrisikopatienten.

Methoden Fokale Leberläsionen bei 55 Hochrisikopatienten wurden unabhängig voneinander von 3 verblindeten Untersuchern anhand beider CEUS-Algorithmen klassifiziert (ESCULAP = Erlanger Synopsis of Contrast-enhanced Ultrasound for Liver Lesion Assessment in Patients at risk; ACR-CEUS-LI-RADS®v.2016 = American College of Radiology CEUS-Liver Imaging Reporting and Data System). Die Kategorisierung erfolgte nach Größe und Ultraschallkontrastmittelverhalten in arterieller, portalvenöser und Spätphase. Die Interobserver-Variabilität für Kontrastmittelverhalten und Kategorisierung sowie die diagnostische Genauigkeit für die definitive HCC-Diagnose wurden für beide Algorithmen ver-

glichen. Histologie und/oder CE-MRT und Follow-Up dienten als Referenzstandard.

Ergebnisse 55 Patienten wurden eingeschlossen (männlich/weiblich = 44/11; 90,9 % Leberzirrhose). Histologische Befunde waren für 39/55 Läsionen verfügbar (70,9 %). Der Referenzstandard ergab 48 HCCs, 2 intrahepatische cholangiozelluläre Karzinome (ICCs), und 5 Non-HCC-non-ICC-Läsionen. Die Interobserver-Übereinstimmung war mäßig bis beachtlich für arterielles Hyperenhancement ($\kappa = 0,53 - 0,67$) und ausreichend bis mäßig für Washout ($\kappa = 0,33 - 0,53$). Die Interobserver-Variabilität war beachtlich für die ESCULAP-Kategorie

($\kappa = 0,64 - 0,68$) und ausreichend für die CEUS-LI-RADS®-Kategorie ($\kappa = 0,3 - 0,39$). Diskrepanzen resultierten hauptsächlich aus unterschiedlicher Einschätzung des Kontrastmittel-Auswaschens.

Schlussfolgerungen Die Interobserver-Übereinstimmung für ESCULAP ist besser als für CEUS-LI-RADS®. Hauptursache ist, dass die Wahrnehmung des Kontrastmittel-Auswaschens zwischen einzelnen Untersuchern stark variiert. Gute Übereinstimmung findet sich hingegen bei der Beurteilung des arteriellen Hyperenhancement, des Schlüsselkriteriums für die CEUS-basierte HCC-Diagnose bei Zirrhose-Patienten.

Introduction

Hepatocellular carcinoma (HCC) in high-risk patients can be diagnosed noninvasively due to the characteristic enhancement pattern on contrast-enhanced imaging. A pattern of arterial phase hyperenhancement with subsequent washout of the contrast agent during the portal venous and late phase is considered to be characteristic of HCC.

National guidelines recommend screening high-risk patients with conventional ultrasound at intervals of six months. In the case of detection of a focal liver lesion, contrast-enhanced imaging is required to rule out or confirm HCC. According to German national guidelines, contrast-enhanced ultrasound (CEUS), computed tomography (CE-CT) and magnetic resonance imaging (CE-MRI) are considered equivalent imaging modalities for the noninvasive diagnosis of HCC [1]. The fact that contrast agents in CEUS remain strictly intravascular without diffusion into the intravascular space makes the method very sensitive for the visualization of tumor vascularity in real time. The excellent diagnostic accuracy of CEUS in the differential diagnosis of focal liver lesions has been demonstrated in several large multi-center studies and meta-analyses [2–6, 17]. For CE-CT and CE-MRI, recent studies and meta-analyses report diagnostic accuracies of 75–95 % with slightly superior sensitivity of CE-MRI (especially using hepatobiliary contrast agents) compared to CE-CT [11–13]. The diagnostic performance of all three contrast-enhanced imaging modalities seems comparable [1–20], with, however, some evidence for best cost-effectiveness for CEUS [14]. However, noninvasive diagnosis of HCC is complicated by the fact that not all lesions display the characteristic “hyper-hypo-pattern” of contrast enhancement. In some well-differentiated HCCs, washout and even arterial phase hyperenhancement may be missing. Also, studies report that about 10 % of HCCs are primarily hypovascular [6–10]. Therefore, standardization of CEUS-based criteria for the noninvasive diagnosis of HCC in high-risk patients remains a challenge.

Recently, standardized CEUS-based diagnostic algorithms such as ESCULAP [21, 22] and CEUS-LI-RADS® [23] have been developed. ESCULAP was designed according to the German National Guidelines for HCC [1] (► **Supplemental Fig. 1**). CEUS-LIRADS® was developed by the American College of Radiology (ACR) [23] (<https://www.acr.org/Quality-Safety/Resources/LIRADS>) and dif-

fers from ESCULAP in several points. The main differences include that with ESCULAP, arterial hyperenhancement alone can be sufficient for the diagnosis of HCC in a lesion of ≥ 20 mm in the cirrhotic liver, whereas with CEUS-LI-RADS®, any lesion without washout in the portal-venous or late phase cannot possibly be classified as definite HCC. Furthermore, ESCULAP contains additional features for large, diffusely infiltrating HCCs, as these tumors often do not display the “typical” contrast enhancement pattern [21–28].

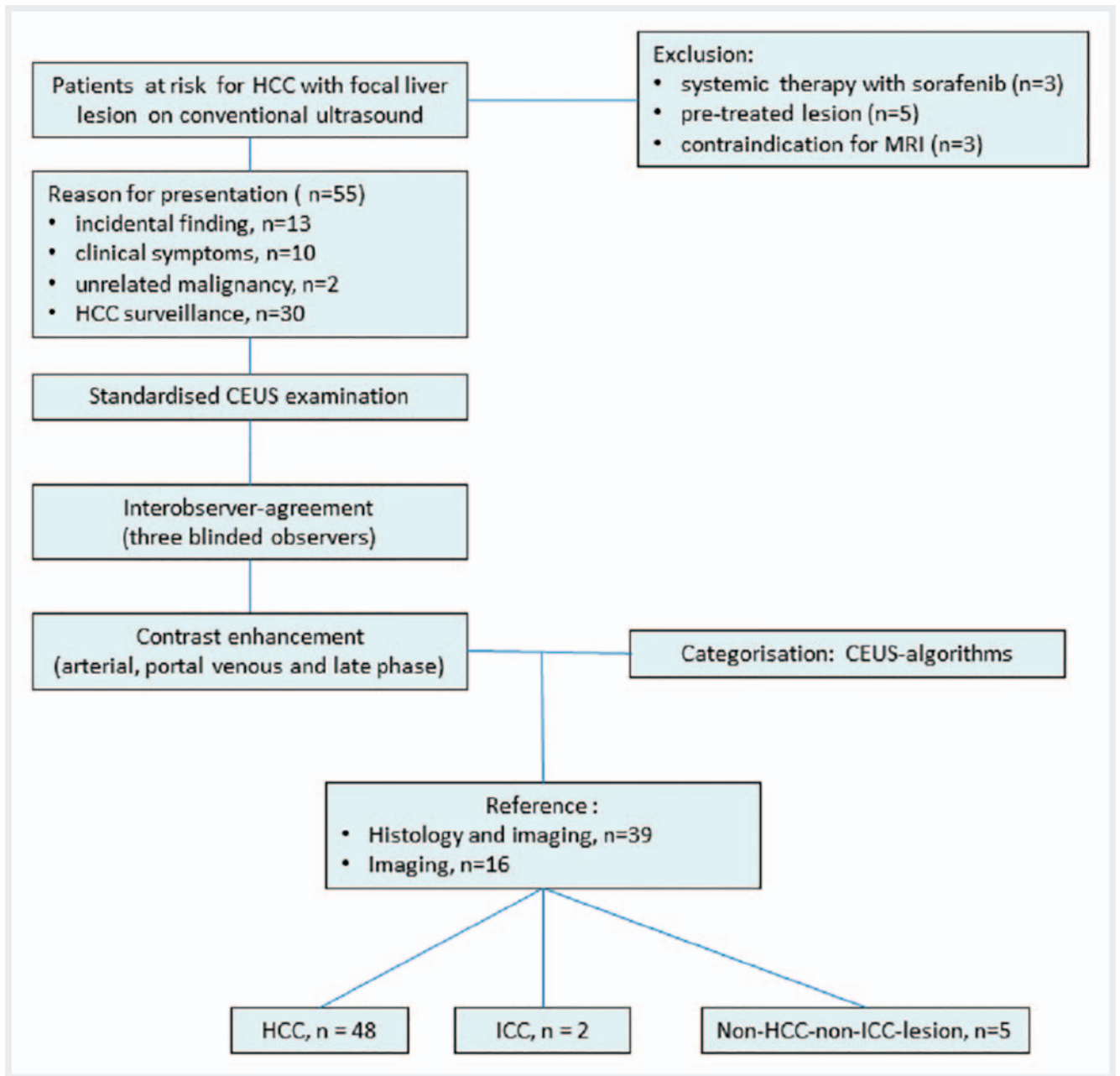
However, these CEUS algorithms are a very recent development and it is not yet clear whether they can be uniformly used by various observers. There are no studies assessing interobserver or intraobserver agreement for standardized CEUS-based algorithms. Therefore, in the current study, we present data on interobserver agreement for two CEUS-based standardized algorithms for the noninvasive diagnosis of HCC in high-risk patients.

Methods

Study Design and Patient Selection

The study design and process of patient selection are shown in ► **Fig. 1**. The risk population for HCC was defined according to the German national guidelines (patients with cirrhosis of any origin, chronic hepatitis B infection, chronic hepatitis C infection with advanced fibrosis, non-alcoholic steatohepatitis (NASH), hereditary hemochromatosis) [1]. The inclusion criteria were at least one focal liver lesion visible on conventional ultrasound and availability of a reference standard (CE-MRI, CE-CT and/or histology) in addition to contrast-enhanced ultrasound (CEUS) of the liver. Patients with systemic treatment for HCC (sorafenib) and those who had received loco-regional treatment were excluded.

Patients at risk for HCC with CEUS examinations of the liver were identified via systematic research of the interdisciplinary liver cancer board and assessment of electronic patient records. In addition, patients who presented for conventional liver ultrasound and/or CEUS at our department were included if they met the inclusion criteria (► **Fig. 1**). Patients had to provide written informed consent for evaluation of anonymized data. The study was approved by the local Ethics Board Committee.



► **Fig. 1** Study design and the process of patient selection.

CEUS

CEUS was performed by three physicians with high expertise in hepatobiliary ultrasound (qualification levels of DEGUM 2 or DEGUM 3 according to the standards of the German society for ultrasound in medicine/DEGUM). This is defined as follows: DEGUM 2, completed specialist medical training; at least 4.5 years of ultrasound experience during routine clinical practice in internal medicine; at least 6000 ultrasound examinations/3000 pathological ultrasound examinations; at least 1000 ultrasound examinations per year; DEGUM 3, completed specialist medical training; at least 6 years of ultrasound experience during routine clinical practice in internal medicine; at least 10 000 ultrasound examina-

tions/7500 pathological ultrasound examinations; at least 1500 ultrasound examinations per year or 750 pathological findings per year). High-end ultrasound devices with a dedicated low-MI technology were used (Siemens Acuson S2000, GE Logiq E9, Toshiba Aplio 500; MI, range: 0.06 – 0.16). For all examinations, a convex probe (range: 4 – 5 MHz) was used. In all patients, conventional B-mode ultrasound examination of the liver preceded CEUS examinations. All CEUS examinations followed a standardized protocol with a low mechanical index according to EFSUMB guidelines for the characterization of focal liver lesions (FLLs) [15]. SonoVue® (Bracco Imaging GmbH, Konstanz, Germany) was used as the contrast agent with bolus injection of 1.5 mL of SonoVue® followed by a 10 mL flush of 0.9 % NaCl. Video clips of the examinations were

recorded for several seconds at defined points in time to capture contrast enhancement patterns in the arterial, portal-venous and late phase. Vascular phases were defined according to EFSUMB guidelines [15]. Time points were defined as follows: arterial phase, from the arrival of first detectable microbubbles until about 30 seconds after injection; portal-venous phase, about 60 seconds after injection; late phase, about 2–3 minutes after injection; late-late phase after 4 and 6 minutes. In case of insufficient contrast enhancement in the late phase, examiners could decide to use a second bolus of SonoVue®. In our patient collective, a second contrast injection was necessary in eight cases.

MRI

MRI was performed following standardized protocols with two different 1.5 Tesla (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 and a spacing of 6 mm. The default protocol for the Magnetom Avanto consisted of a T2-weighted (T2w) half-Fourier acquisition single-shot turbo spin-echo (HASTE) with transverse (tra) image orientation, acquired during one single breath-hold (bh) with a repetition time (TR) of 900 ms and an echo time (TE) of 122 ms, followed by a T2w HASTE tra bh with fat saturation (fs), a TR of 1000 ms and a TE of 90 ms, a T2w HASTE bh in coronal (cor) orientation with a TR of 1000 ms and a TE of 122 ms, 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) volume-interpolated (VIBE) gradient-echo (GRE) sequence in transverse orientation in an opposed-phase and an in-phase with a TR of 7.96 ms and a TE of 2.38 ms for the opposed-phase and a TE of 4.76 ms for the in-phase. Subsequently, a series of T1w VIBE GRE fs tra bh sequences with a TR of 5.42 ms and a TE of 2.38 ms was acquired without contrast and in the arterial, venous, and post-CE phase. Weight-adapted gadobutrol 1.0 mmol/ml (Gadovist®, Bayer Pharma AG, Berlin, Germany) was injected intravenously as a non-liver-specific contrast agent. Further post-CE phases included a T1w VIBE GRE fs cor bh sequence with a TR of 5.6 ms and a TE of 2.56 ms as well as a high resolution T1w fast low angle shot fs tra bh sequence with a TR of 130 ms and a TE of 7.15 ms. The matrix was roughly 260 × 320 pixels for all sequences except the DWI, which was 130 × 160 pixels. In the protocol for the Magnetom Aera, the T2w HASTE fs tra bh was replaced by a T2w turbo-spin-echo (TSE) fs tra bh sequence with a TR of 5217 ms and a TE of 102 ms. The T1w VIBE GRE sequences were replaced by T1w VIBE DIXON sequences with a TR of 6.65 ms and a TE of 2.39/4.77 ms. Furthermore, slightly different TR and TE times were used.

CT

CT was performed following standardized protocols with 20, 64, 128, or 256 row scanners (Somatom Definition AS20, Somatom Sensation, Somatom Definition AS+, Somatom Force, or Definition Flash, Siemens AG Healthcare, Forchheim, Germany) using power injectors (Accutron CT-D, Medtron, Saarbrücken, Germa-

ny) for intravenous contrast media application (Imeron 350, Bracco Imaging, Milan, Italy). An unenhanced baseline scan of the liver was obtained. The arterial phase was identified by test-bolus measurements. The portal-venous phase was set at 70 s after the start of contrast media injection.

Interobserver agreement for CEUS-based diagnostic algorithms

CEUS examinations were evaluated according to the standardized diagnostic algorithms ESCULAP and CEUS-LI-RADS® v.2016 (► **Supplemental Fig. 1**; <https://www.acr.org/Quality-Safety/Resources/LIRADS>). Three observers with 2, 4 and 6 years of experience in CEUS examinations of the liver blinded to patients' clinical data and final diagnoses independently reviewed the CEUS examinations retrospectively. The observers were instructed to focus on the target lesion. The observers had access to both the images from conventional B-mode ultrasound and the CEUS video clips of all vascular phases. Observers had to decide on hyper-, iso- or hypoenhancement of the target lesion relative to the surrounding liver parenchyma in the arterial, portal venous and late phase. Afterwards, they had to assign a category with both ESCULAP and CEUS-LI-RADS® to the target lesion. Beforehand, all observers received both theoretical training of the CEUS-based algorithms and hands-on training with 10 test lesions in order to get used to the application of the CEUS-based algorithms.

Diagnoses resulting from the CEUS-based algorithms were also compared to diagnoses from the on-site observer who had performed standard CEUS examinations following EFSUMB guidelines [15].

Statistical Analysis

Statistical analysis was performed with the programs IBM SPSS Statistics 2013 and Microsoft Excel 2010. Values are expressed as mean and range, or absolute sum and proportion in percent as appropriate. McNemar test for paired samples was used for comparison of means. Chi Square Test and Fisher's exact test were used for assessment of categorical data. Cohen κ statistics were used for evaluation of interobserver agreement. Differences were considered statistically significant for $p < 0.05$.

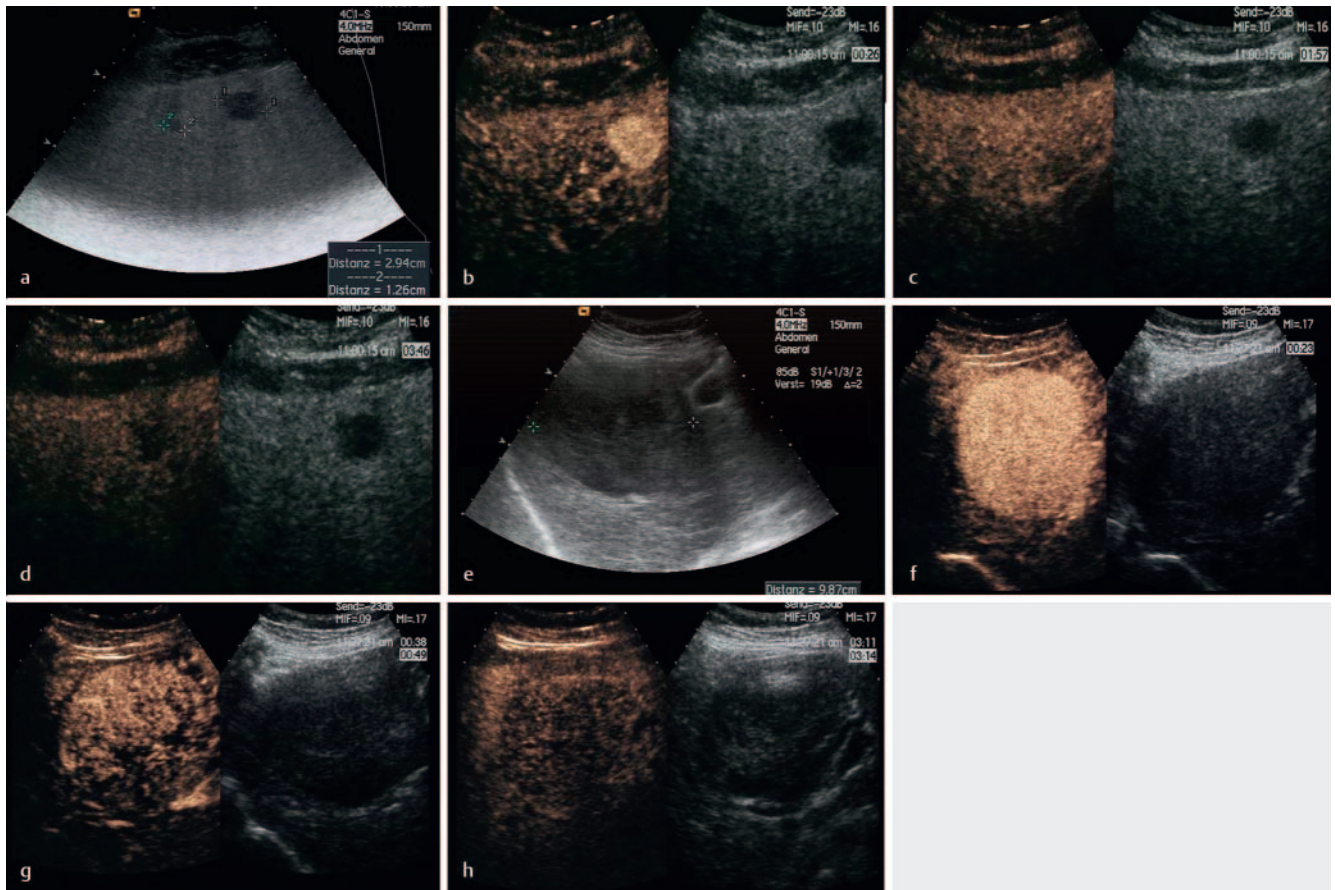
Results

Patient and tumor characteristics

55 patients were included in the study (male/female, 44/11; mean age: 65.9 years; range: 53–86 years). 50 patients (90.9%) had cirrhosis. 2 patients had fibrosis (Ishak 3), 1 had steatosis, and 2 had both steatosis and fibrosis (Ishak 2–3).

Final diagnosis according to the reference standard was HCC in 48/55 lesions (87.3%), and intrahepatic cholangiocellular carcinoma (ICC) in 2 lesions. 5 lesions were revealed as benign findings (3 regenerate nodules, 1 cyst, 1 focal fat sparing).

Histological findings were available in 39/55 lesions (70.9%; 35 HCCs, 2 ICCs, 2 regenerate/dysplastic nodules). In the other cases, final diagnosis was based on findings upon contrast-enhanced



► **Fig. 2** Examples of CEUS images with complete (a–d) versus partial (e–h) agreement of the three observers. **a–d** Complete agreement of all three observers in all vascular phases. **a** conventional B-mode: two hypoechoic lesions of 29 and 12 mm in cirrhotic liver. **b** arterial phase hyperenhancement (agreement of all three observers). **c** late phase after 2 minutes: isoenhancement (agreement of all three observers). **d** late phase: mild washout after 4 minutes (agreement of all three observers). **e–h** Partial agreement. **e** conventional B-mode: hypoechoic lesion with a size of 10 cm in cirrhotic liver. **f** arterial phase hyperenhancement (agreement of all three observers). **g** portal venous phase: two observers perceived discrete washout, and one observer perceived isoenhancement. **h** late phase: all three observers agreed on washout.

imaging and, in case of benign lesions, constant findings over more than 6–12 months during follow-up. Tumor characteristics are summarized in ► **Supplemental Table 1**.

Interobserver agreement for contrast enhancement patterns in the arterial, portal venous and late phase

Results for the assessment of interobserver agreement for major imaging features and both CEUS-based algorithms are shown in ► **Supplemental Table 2**. Interobserver agreement according to Cohen's Kappa was moderate to substantial for arterial phase hyperenhancement and fair to moderate for washout appearance. Arterial phase hyperenhancement was observed in 75%/78%/87% of lesions (Observer 1/2/3); differences were significant only between observers 1 and 3 ($p = 0.039$). Washout was observed in 69%/44%/80% of lesions, with significant differences between observers 1/2 ($p < 0.001$) and observers 2/3 ($p < 0.001$). However, the percentage of lesions with agreement was high with 84–88% for arterial phase hyperenhancement and 64–82% for washout in the portal-venous or late phase.

Interobserver agreement for standardized CEUS algorithms

Results for the categorization of lesions with the CEUS-based algorithms for the three observers are shown in ► **Supplemental Table 3**. Concerning the CEUS-based algorithms, agreement was fair for ACR CEUS-LI-RADS® categories, but substantial for ESCU-LAP categories (► **Supplemental Table 2**). Washout in the portal-venous or late phase was perceived by observer 2 in significantly fewer cases compared to observers 1 and 3. ► **Fig. 2** shows examples of lesions with perfect (► **Fig. 2a–d**) versus partial agreement (► **Fig. 2e–h**) between the three observers due to discrepant perception of washout resulting in different categorization according to CEUS-LI-RADS®.

Interobserver agreement for the diagnosis of HCC with CEUS-based algorithms in histologically proven lesions

Comparison of interobserver agreement for the definite diagnosis of HCC with the two standardized CEUS-based algorithms is shown in ► **Supplemental Table 4**. Results are shown for the

subgroup of histologically proven lesions only ($n = 39$), regarding histology as the gold standard.

For all three observers, the diagnostic accuracy for the definite diagnosis of HCC was superior with ESCULAP compared to CEUS-LI-RADS® (mean values 93.2 % versus 70.1 %). Also, the sensitivity and NPV were markedly superior with ESCULAP, whereas the specificity was slightly better with CEUS-LI-RADS® for observers 2 and 3. However, there were only 4 non-HCC lesions in this subpopulation with histological findings available, so cautious interpretation is necessary.

Interestingly, differences between the three observers are more pronounced with CEUS-LI-RADS® compared to ESCULAP. Particularly, there is a tendency towards strikingly better results for observer 3 compared to observers 1 and 2 with CEUS-LI-RADS® only, although absolute values for diagnostic accuracy are superior with ESCULAP for all three observers.

The proportion of lesions categorized as definite HCC differed significantly between all three observers with CEUS-LI-RADS® (LR-5) ($p < 0.001 - 0.014$), but not with ESCULAP (ESCULAP-4) ($p = 0.06 - 0.74$).

To evaluate the potential additional value of standardized diagnostic algorithms for the diagnosis of HCC with CEUS, we compared the results for both ESCULAP and CEUS-LI-RADS® with the diagnoses made by the on-site observer conducting the primary CEUS examination, who had not used any diagnostic algorithm for final diagnosis. The on-site observer correctly identified 29/35 HCCs (82.9 %) compared to 93.2 % with ESCULAP and 70.2 % with CEUS-LI-RADS. Of the two ICC lesions, one was correctly diagnosed and the other was categorized as “benign lesion”. In four cases (10.3 %), the on-site observer did not determine a final diagnosis, but offered several differential diagnoses.

Discussion

The excellent diagnostic accuracy of contrast-enhanced ultrasound for the noninvasive diagnosis of hepatocellular carcinoma in high-risk patients has been shown in several multi-center studies. According to German National S3 (AWMF) guidelines, CEUS can be considered equivalent to CE-CT and CE-MRI in the diagnostic workup of lesions with suspicion of HCC [1–8, 14–18, 20]. In comparison with MRI, CEUS has the advantage of providing a unique real-time assessment of contrast enhancement patterns using strictly intravascular contrast agents. CEUS is very sensitive for the detection of early arterial phase hypervascularization, which can be missed with CT in case of not optimized contrast for CT scans. With MRI, there are artifacts such as av-shunts, which can be misleading. Contrast agents in CEUS differ from those in MRI, resulting in different vascularization patterns. Malignant lesions show a washout of contrast agent during the late phase, whereas in MRI, diffusion of contrast agents into the tumor interstitium can result in increasing contrast enhancement of malignant lesions in the late phase [36]. Recent studies suggest that combination of CEUS and MRI with liver-specific contrast agents might reduce the rate of false-negative findings in the assessment of focal liver lesions [36].

Another advantage is that CEUS can be carried out immediately after the detection of a focal liver lesion on B-mode ultrasound, thus avoiding further diagnostic delay.

However, CEUS is often accused of being observer-dependent and subjective. Thus, standardized CEUS-based algorithms seem to be a promising approach to facilitate assessment and documentation of lesions with suspicion of HCC in patients at risk. This study assessed the interobserver agreement for two CEUS-based algorithms in direct comparison.

For the assessment of the contrast enhancement pattern, our results show a substantial proportion of agreement between the observers (on average, 86 % of agreement for arterial hyperenhancement and 69 % for washout). However, using Cohen's Kappa as the statistical gold standard for the evaluation of interobserver agreement, we found moderate to substantial interobserver agreement for arterial phase hyperenhancement ($\kappa = 0.53 - 0.67$), but only fair to moderate agreement for the perception of washout ($\kappa = 0.33 - 0.53$). This apparent discrepancy between relatively low Kappa values in spite of a high proportion of agreement is due to the fact that calculation of Cohen's Kappa is influenced by the frequency of a certain feature being observed. If a feature is observed very frequently (as was the case with arterial hyperenhancement in our patient collective), the calculation of Kappa relies on the assumption that a significant proportion of agreement is due to chance, thus resulting in low Kappa values despite a high percentage of agreement between the observers [29]. Our results show that interobserver agreement is better for arterial hyperenhancement, which is the key diagnostic feature of HCC in CEUS, than for the perception of washout.

Direct comparison of the two CEUS-based algorithms shows a significantly higher diagnostic accuracy for the definite noninvasive diagnosis of HCC with ESCULAP than with CEUS-LI-RADS® for all three observers. This is mostly due to the fact that a significant percentage (almost one third) of HCCs was underestimated as LR-4 or LR-3 with CEUS-LI-RADS®. In order to avoid potential misdiagnoses inherent in all kinds of contrast-enhanced imaging, diagnoses from the CEUS-based algorithms were compared with histology as the reference standard and not with diagnoses from other imaging modalities. Taking together the results from all three observers for CEUS-LI-RADS® in histologically proven lesions, 27/29 LR-4 lesions (93.1 %) were HCCs. Poor values for specificity are also due to the small number of non-HCC lesions ($n = 4$) in this subcollective of histologically proven lesions. This reflects the high a-priori probability for HCC for any solid liver lesion seen in high-risk patients, which is well-known from the clinical routine.

Concerning the use of CEUS-based algorithms in the differential diagnosis of HCC and ICC, cautious interpretation is necessary as there were only 2 ICC lesions in the study collective. With ESCULAP, there is a separate category ESCULAP-C to designate ICCs. To be categorized as ESCULAP-C, a lesion has to show both rim-like arterial enhancement and early washout beginning after < 60 seconds. With CEUS-LI-RADS®, a correctly classified ICC would be categorized as LR-M (“definitely or probably malignant, not specific for HCC”). In the study collective, none of the two ICCs was correctly classified as ESCULAP-C or CEUS-LI-RADS® LR-M. This im-

plies the risk that ICCs can be misinterpreted as HCCs or benign lesions.

We found substantially higher interobserver agreement and diagnostic accuracy for ESCULAP compared to CEUS-LI-RADS®, while results with CEUS-LI-RADS® were better for the more experienced reader 3 than for readers 1 and 2. Differences were negligible with ESCULAP. Thus, ESCULAP seems to be the better option for less experienced readers. Yet, even for more experienced readers, there is no benefit of CEUS-LI-RADS® over ESCULAP. One main reason for this superiority of ESCULAP in the CEUS-based diagnosis of HCC in high-risk patients is that in contrast to CEUS-LI-RADS®, washout is not mandatory with ESCULAP for the definite diagnosis of HCC in a lesion ≥ 20 with arterial hyperenhancement in a high-risk patient. Accordingly, in our study, discrepancies in the perception of washout led to classification of a lesion as LR-CEUS-4 by one observer and LR-CEUS-5 by another observer (whereas the same lesion would be classified as ESCULAP-4 = definitive HCC by both observers). There is evidence from the literature that the presence and extent of contrast washout might be related to the grade of differentiation in hepatocellular carcinoma with no washout especially in well-differentiated tumors [2, 8]. In our study collective, most HCCs showed high or moderate differentiation (G1, $n = 9/25.7\%$; G2, $n = 22/62.9\%$), whereas only a minority shows poor differentiation (G3, $n = 4/11.4\%$) (► **Supplemental Table 1**). This might be a reason why classification of washout appearance proved particularly difficult and was not observed in all cases. However, the histological subgroups were too small for statistically significant correlation of tumor grading and washout appearance in our study collective.

Arterial phase hyperenhancement can certainly be considered as the key feature of HCC in CEUS. In our study, 9/48 HCCs lacked arterial hyperenhancement in the perception of at least one observer and were therefore underestimated as LR-3 or LR-4 by CEUS-LI-RADS®. However, this was most often the case in larger HCCs ($\geq 7 - 10$ cm) with necrotic areas or diffuse infiltration of the liver. In most cases, the interobserver agreement was high for arterial phase hyperenhancement and discrepancies were mostly seen in the perception of washout.

To date, only a few studies have assessed interobserver agreement for LI-RADS in MRI or CT. For CT and MRI, studies report κ -values of 0.35 – 0.69. As for CEUS, interobserver agreement seems to be higher for arterial hyperenhancement than for washout appearance and perception of tumor capsule [31 – 35]. However, our study is the first to address this issue for standardized CEUS-based algorithms. Therefore, our results are not directly comparable to the literature.

A limitation of our study is the relatively small patient number and the fact that histological findings were not available in all cases. However, to our knowledge, our study is the first one assessing interobserver agreement in standardized CEUS-based algorithms for the diagnosis of HCC. Therefore, our findings can provide a starting point for further prospective studies in larger patient collectives.

Furthermore, our study population includes $>90\%$ of patients with liver cirrhosis and only one patient with chronic hepatitis B alone (and fibrosis Ishak 2). However, it can be assumed that interobserver agreement for the CEUS-based standardized algo-

rithms is likely to improve along with ongoing implementation of these algorithms in the clinical routine due to training effects of the observers and increasing familiarization with the use of the standardized CEUS algorithms.

Conclusion

CEUS-based algorithms are designed for standardized assessment of contrast enhancement patterns in lesions with suspicion of HCC in high-risk patients. Interobserver agreement is good for arterial phase hyperenhancement, whereas the perception of contrast washout varies between different observers. The omission of washout as a mandatory prerequisite for the definite diagnosis of HCC in CEUS would lead to an improvement of interreader agreement and diagnostic accuracy in patients with liver cirrhosis.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- [1] S3-Leitlinie Deutschland: AWMF-Leitlinie "Hepatozelluläres Karzinom", Registernummer 032 – 053OL. Stand: 01.05.2013, gültig bis 30.04.2018; http://www.awmf.org/uploads/tx_szleitlinien/032-053OL_S3_Hepatozellul%C3%A4res_Karzinom_Diagnostik_Therapie_2013-verlaengert.pdf
- [2] Leoni S, Piscaglia F, Granito A et al. Characterization of primary and recurrent nodules in liver cirrhosis using contrast-enhanced ultrasound: which vascular criteria should be adopted? *Ultraschall in Med* 2013; 34: 280 – 287
- [3] Strobel D, Bernatik T, Blank W et al. Diagnostic accuracy of CEUS in the differential diagnosis of small (≤ 20 mm) and subcentimetric (≤ 10 mm) focal liver lesions in comparison with histology. Results of the DEGUM multicenter trial. *Ultraschall in Med* 2011; 32: 593 – 597
- [4] Strobel D, Seitz K, Blank W et al. Tumor-specific vascularization pattern of liver metastasis, hepatocellular carcinoma, hemangioma and focal nodular hyperplasia in the differential diagnosis of 1349 liver lesions in contrast-enhanced ultrasound (CEUS). *Ultraschall in Med* 2009; 30: 376 – 382
- [5] Seitz K, Greis C, Schuler A et al. Frequency of tumor entities among liver tumors of unclear etiology initially detected by sonography in the non-cirrhotic or cirrhotic livers of 1349 patients. Results of the DEGUM multicenter study. *Ultraschall in Med* 2011; 32: 598 – 603
- [6] Boozari B, Soudah B, Rifai K et al. Grading of hypervascular hepatocellular carcinoma using late phase of contrast enhanced sonography – a prospective study. *Dig Liver Dis* 2011; 43: 484 – 490
- [7] Kim TK, Jang HJ. Contrast-enhanced ultrasound in the diagnosis of nodules in liver cirrhosis. *World J Gastroenterol* 2014; 20: 3590 – 3596
- [8] D'Onofrio M, Crosara S, De Robertis R et al. Contrast-Enhanced Ultrasound of Focal Liver Lesions. *Am J Roentgenol* 2015; 205: W56 – W66
- [9] Shah S, Shukla A, Paunipagar B. Radiological features of hepatocellular carcinoma. *J Clin Exp Hepatol* 2014; 4 (Suppl. 3): S63 – S66
- [10] Choi BI, Lee JM, Kim TK et al. Diagnosing Borderline Hepatic Nodules in Hepatocarcinogenesis: Imaging Performance. *Am J Roentgenol* 2015; 205: 10 – 21
- [11] Roberts LR, Sirlin CB, Zaiem F et al. Imaging for the Diagnosis of Hepatocellular Carcinoma: a Systematic Review and Meta-analysis. *Hepatology* 2018; 67: 401 – 421

- [12] Kim BR, Lee JM, Lee DH et al. Diagnostic Performance of Gadoxetic Acid-enhanced Liver MR Imaging versus Multidetector CT in the Detection of Dysplastic Nodules and Early Hepatocellular Carcinoma. *Radiology* 2017; 285: 134–146
- [13] Bae JS, Kim JH, Yu MH et al. Diagnostic accuracy of gadoxetic acid-enhanced MR for small hypervascular hepatocellular carcinoma and the concordance rate of Liver Imaging Reporting and Data System (LI-RADS). *PLoS One* 2017; 12: e0178495
- [14] Westwood M, Joore M, Grutters J et al. Contrast-enhanced ultrasound using SonoVue® (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* 2013; 17: 1–243
- [15] Claudon M, Dietrich CF, Choi BI 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. *Ultraschall in Med* 2013; 34: 11–29
- [16] Bota S, Piscaglia F, Marinelli S et al. Comparison of international guidelines for noninvasive diagnosis of hepatocellular carcinoma. *Liver Cancer* 2012; 1: 190–200
- [17] Friedrich-Rust M, Klopffleisch T, Nierhoff J et al. Contrast-Enhanced Ultrasound for the differentiation of benign and malignant focal liver lesions: a meta-analysis. *Liver Int* 2013; 33: 739–755
- [18] Wildner D, Bernatik T, Greis C et al. CEUS in Hepatocellular Carcinoma and Intrahepatic Cholangiocellular Carcinoma in 320 Patients – Early or Late Washout Matters: A Subanalysis of the DEGUM Multicenter Trial. *Ultraschall in Med* 2015; 36: 132–139
- [19] McEvoy SH, McCarthy CJ, Lavelle LP et al. Hepatocellular carcinoma: illustrated guide to systematic radiologic diagnosis and staging according to guidelines of the American Association for the Study of Liver Diseases. *Radiographics* 2013; 33: 1653–1668
- [20] Giorgio A, Montesarchio L, Gatti P et al. 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* 2016; 25: 205–211
- [21] Schellhaas B, Wildner D, Pfeifer L et al. LI-RADS-CEUS – Proposal for a Contrast-Enhanced Ultrasound Algorithm for the Diagnosis of Hepatocellular Carcinoma in High-Risk Populations. *Ultraschall in Med* 2016; 37: 627–634
- [22] Schellhaas B, Görtz RS, Pfeifer L et al. Diagnostic accuracy of contrast-enhanced ultrasound for the differential diagnosis of hepatocellular carcinoma: ESCULAP versus CEUS-LI-RADS. *Eur J Gastroenterol Hepatol* 2017; 29: 1036–1044
- [23] American College of Radiology. Liver Imaging Reporting and Data System version 2014. Accessed 04/2017, from <http://www.acr.org/Quality-Safety/Resources/LIRADS>
- [24] Mitchell DG, Bruix J, Sherman M et al. LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. *Hepatology* 2015; 61: 1056–1065
- [25] Tang A, Valasek MA, Sirlin CB. Update on the Liver Imaging Reporting and Data System: What the Pathologist Needs to Know. *Adv Anat Pathol* 2015; 22: 314–322
- [26] Elsayes KM, Kielar AZ, Agrons MM et al. Liver Imaging Reporting and Data System: an expert consensus statement. *J Hepatocell Carcinoma* 2017; 4: 29–39
- [27] Piscaglia F, Wilson SR, Lyschik A et al. 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 in Med* 2017. doi:10.1055/s-0042-124661
- [28] Karlas T. LI-RADS-CEUS for the classification of HCC risk in liver lesions. *Z Gastroenterol* 2017. doi:10.1055/s-0043-103327
- [29] McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012; 22: 276–282
- [30] Forner A, Vilana R, Bianchi L et al. Lack of arterial hypervascularity at contrast-enhanced ultrasound should not define the priority for diagnostic work-up of nodules <2 cm. *J Hepatol* 2015; 62: 150–155
- [31] Bashir MR, Huang R, Mayes N 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* 2015; 42: 305–314
- [32] Ehman EC, Behr SC, Umetsu SE et al. 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)* 2016; 41: 963–969
- [33] Barth BK, Donati OF, Fischer MA et al. Reliability, Validity, and Reader Acceptance of LI-RADS-An In-depth Analysis. *Acad Radiol* 2016; 23: 1145–1153
- [34] Davenport MS, Khalatbari S, Liu PS et al. Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging. *Radiology* 2014; 272: 132–142
- [35] Zhang YD, Zhu FP, Xu X 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* 2016; 43: 373–383
- [36] Beyer LP, Wassermann F, Pregler B et al. Characterization of Focal Liver Lesions using CEUS and MRI with Liver-Specific Contrast Media: Experience of a Single Radiologic Center. *Ultraschall in Med* 2017; 38: 619–625