

Answering Reviewers

Answer to Reviewer #1

Dear reviewer, thank you for your comments.

In response to your specific issues:

- 1) The feasibility to calculate the diagnostic ability of LR-4 or LR-3. Take LR-4 for example. A nodule should be regarded as HCC if it meets the feature of LR-4. If a nodule does not meet the feature of LR-4, should it be taken as non-HCC even if it meets the criteria for LR-5? The same question exists for LR-3. That is why in most studies, only percentage of HCC was presented in these categories.

LR-4 class was originally built to include “probably HCC” nodules. So, we think it is appropriate to test its accuracy in the diagnosis of this tumor. Obviously, its sensitivity is expected to be low, since there is another class, LR-5, that expresses a higher risk of HCC. The purpose of the evaluation of LR-4 accuracy alone was to check it against the accuracy of LR-4 and 5 merging class. By doing this we can affirm that LR-4 alone is not a suitable class for the diagnosis of HCC, due to its extremely low sensitivity. LR-5 is more effective, however perfectible, as its sensitivity is still quite low. LR-4 and 5 merging class, on the contrary, is much more sensitive and the loss of specificity is acceptable. Other studies evaluated the accuracy of LR-4 in the diagnosis of HCC. For example, in [Terzi E, et al. Contrast ultrasound LI-RADS LR-5 identifies hepatocellular carcinoma in cirrhosis in a multicenter retrospective study of 1,006 nodules. J Hepatol 2018](#), table S5 summarize the diagnostic accuracy for HCC of LR-5, LR-4 and LR-3 classes. In [Cioaltea A, et al. Role of Contrast-Enhanced Ultrasonography in Hepatocellular Carcinoma by Using LI-RADS and Ancillary Features: A Single Tertiary Centre Experience. Diagnostics \(Basel\). 2021 Nov 29](#), the authors tested the accuracy of LR-4, LR-5 and LR-4 + 5 in the diagnosis of HCC.

LR-3 class was built to express an “intermediate risk of malignancy”, so we tested its accuracy in the identification of all malignancies (not only HCC) to quantify this risk. By doing this, we also aimed to investigate the difference in terms of risk of malignancy for the different patterns of LR-3 class.

- 2) How to explain the low PPV of LR-M for the diagnosis of ICC?

Our PPV of LR-M for the diagnosis of ICC was actually higher compared to other studies, as you probably intended to point out and as we understand from your further comments on this issue. The problem of the heterogeneity of LR-M class is already known in the literature. For example, in [Terzi E, et al. Contrast ultrasound LI-RADS LR-5 identifies hepatocellular carcinoma in cirrhosis in a multicenter retrospective study of 1,006 nodules. J Hepatol 2018](#) the authors reported 37.8% PPV. That is the reason why a lot of effort has been put to try to refine LR-M criteria, as we explain in the Discussion, see:

- [Li F, et al. Distinguishing intrahepatic cholangiocarcinoma from hepatocellular carcinoma in patients with and without risks: the evaluation of the LR-M criteria of contrast-enhanced ultrasound liver imaging reporting and data system version 2017. *Eur Radiol* 2020; 30: 461-470 DOI: 10.1007/s00330-019-06317-2]

- [Ding J, et al. Impact of Revision of the LR-M Criteria on the Diagnostic Performance of Contrast-Enhanced Ultrasound LI-RADS. *Ultrasound Med Biol* 2021; published online ahead of print, 2021 Sep 28 DOI: 10.1016/j.ultrasmedbio.2021.08.007]
- [Zeng D, et al. Using new criteria to improve the differentiation between HCC and non-HCC malignancies: clinical practice and discussion in CEUS LI-RADS 2017. *Radiol Med* 2021; published online ahead of print, 2021 Oct 19 DOI: 10.1007/s11547-021-01417-w]
- [Chen LD, et al. Comparison between M-score and LR-M in the reporting system of contrast-enhanced ultrasound LI-RADS. *Eur Radiol* 2019; 29: 4249-4257 DOI: 10.1007/s00330-018-5927-8]

The composition of pathologic entities in LR-M has enormous impact for the diagnosis of ICC. This study not only has a relatively lower percentage of HCC, but also high percentage of ICC in LR-M group. So, the diagnostic power of LR-M for diagnosing ICC may be exaggerated and the finding needs outer validation before clinical application.

We did observe some difference in the percentage of ICC and HCC in the LR-M class compared to some other studies. There are many factors that could explain this gap, including differences in inclusion/exclusion criteria, differences in cirrhosis etiology, geographical factors, experience of the operator in CEUS execution and in CEUS LI-RADS class attribution. We agree that a validation of our findings in future prospective multicentric studies would be advisable should our results be extended to the global population of patients with cirrhosis, as we declared in Discussion.

- 3) In this study, there was a much higher proportion of ICC in LR-M category compared with previous studies (including those with large sample) which HCC composed the majority of LR-M lesions. The authors should explain this discrepancy. See answer to point 2.
- 4) The issue of reference standard should be addressed in detail since most of the cases were diagnosed by CT/MRI. Are there any cases that diagnosed as CE/MRI LI-RADS 2 or 3 or 4? If so, the reference standard is not as robust as it should be.
Non-invasive diagnosis of HCC was made only if the nodule showed typical features after CT and/or MRI (arterial phase hyperenhancement, washout in venous phase, size > 10 mm), that implicate a CT/MRI LR-5 class. When these criteria were not satisfied a biopsy was performed (CT/MRI LR-2, 3 and 4), unless the nodule was clearly benign (CT/MRI LR-1 class).
- 5) In the methods part, how long was the CEUS procedure observed? As we observed in clinical practice, washout could be identified as late as 5 minutes after contrast agent injection. If the procedure was not observed long enough, some of the LR-5 cases could be taken as LR-4.

All CEUS examinations were conducted for at least 5 minutes starting from the injection of the contrast agent. We have added this to the main text for better highlight.

- 6) If patients have multiple nodules, especially those with more than 3 lesions, are all lesions included for analysis? Multiple target nodules might impact the effective evaluation of liver nodules in CEUS examination. The author should elucidate this issue.

Nodules located in different liver segments were analyzed separately with individual boluses of contrast. Within the same segment, only one target nodule was included for analysis based on best visualization criteria. We have added this to the main text.

- 7) Tables should be presented in form of three-line table.

We have formatted the tables.

Answer to Reviewer #2

Dear reviewer, thank you for your comments.

In response to your specific issues:

- 1) Although the manuscript was structured, it seemed to be lengthy.

We have edited the manuscript in order to make it more fluid.

- 2) The probability of HCC in CEUS LR-4 category was 97.4% in this study, which was better than the estimate value of LR-5 in other studies concerning CEUS and CT/MRI LI-RADS category (Shin et al. *Liver International*.2020,DOI: 10.1111/liv.14617; van der Pol et al, *Gastroenterology*, 2019,DOI: 10.1053/j.gastro.2018.11.020). That might be the cause that the specificity of CEUS LR-4/5 for diagnosing HCC remained 94.3% specificity and 98.8% PPV. The results cannot be generalized to other populations.

There are several other studies in which the PPV of LR-4 class for HCC was similar to our results. In particular, in *Schellhaas B, 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 Sep*, the proportion of HCC in LR-4 category was 17 out of 17 (100% PPV). In addition, in *Cioaltea A, et al. Role of Contrast-Enhanced Ultrasonography in Hepatocellular Carcinoma by Using LI-RADS and Ancillary Features: A Single Tertiary Centre Experience. Diagnostics (Basel). 2021 Nov 29*, 37 out of 39 LR-4 nodules were HCC (94.9% PPV). Both papers suggested a real benefit in terms of accuracy by merging LR-4 and LR-5 classes. There are many possible explanations for these discrepancies among different studies: retrospective or prospective nature, differences in inclusion/exclusion criteria, differences in diagnostic standard, etc. We agree that a validation of our findings in future prospective multicentric studies would be advisable should our results be extended to the global population of patients with cirrhosis, as we declared in Discussion. Regarding the paper from van der Pol et al, we think it is not appropriate to compare CEUS LI-RADS with CT/MRI LI-RADS, as they refer to very different techniques.

- 3) The process of patient selection was not clearly elaborated. Lesions that can easily be confused with HCC (eg, mixed HCC-CCA, FNH, hepatic adenoma, inflammatory

pseudotumor, etc) were not included in the cohort, which might lead to overestimation of diagnostic accuracy.

We enrolled patients with cirrhosis who developed a new nodule during their surveillance program. This is the reason why we had a high percentage of malignant nodules. Mixed HCC-CCA, FNH, hepatic adenoma, and inflammatory pseudotumor are rare lesions that were not found in our case series, although we acknowledge that histology was not available for all nodules.

- 4) Pathological diagnosis was only available for 102 (20%) cases.
This is a limitation of our study, as we declared in the text. However, we believe it is not ethical to pursue a biopsy where national and international guidelines allow a non-invasive diagnosis by CT/MRI. An alternative option was to limit our study to the nodules for which pathology was available. However, we rejected this option not only because of the potential loss in the statistical power of the study, but also to avoid a selection bias. In fact, nodules subjected to biopsy are more likely to have an atypical feature at CT/MRI and for this reason are not representative of the totality of nodules developed in cirrhosis. As this is a real-life study, we decided to include both invasively and non-invasively diagnosed nodules.
- 5) It was mentioned before, the abstract is lengthy. Please edit the text to make it as concise as possible.
We shortened the abstract as per the Journal guidelines, which set up a minimal length for some sections.
- 6) Pg 4 Ln12 The phrase 'even though' was not used appropriately.
We amended the text.
- 7) Key words When available, please use controlled vocabularies, such as medical subject headings (MeSH).
We have corrected keyword as per MeSH
- 8) Pg 4 Ln25. Use 'cirrhosis' instead of 'hepatic cirrhosis'.
Corrected.
- 9) Introduction The authors did not give a full account of the innovativeness of the study. The Introduction has been revised.
- 10) Material and Methods Pg 7 Ln4. How to confirm the presence of cirrhosis? Was the diagnosis of cirrhosis established by pathological diagnosis or by imaging with MRI or elastography in conjugation with laboratory and clinical findings?
Cirrhosis was diagnosed on the basis of clinical data, biochemical parameters, imaging criteria and elastosonographic measurements. We have added this to the main text.
- 11) Pg 7 Ln3-5. How many patients with multiple lesions? And how to deal with cases with multiple intrahepatic foci in this study? How were target lesions selected? How many target lesions per patient were allowed?
148 patients had only one nodule while 121 patients had more than one nodule. Nodules located in different liver segments were analyzed separately with individual

boluses of contrast. Within the same segment, only one target nodule was included for analysis based on best visualization criteria. We have added this to the main text.

12) Pg 7 Ln6-7. Were patients consecutively or selectively included?

The study was retrospective and patients were enrolled consecutively.

13) Pg 7 Ln10-11. 'CT and/or MRI, when typical for HCC or definitely benign, were used as the gold standard imaging modalities, as per HCC international guidelines'. Indeed, different guidelines have slightly different imaging reference standard. Does 'definitely benign' refer to hemangioma?

"Typical for HCC" at CT/MRI refers to arterial phase hyperenhancement followed by venous phase washout. "Definitely benign" refers to hemangioma, hepatic fat deposition/sparing and hypertrophic pseudomass. These criteria are in accordance with both AASLD and AISF guidelines.

14) Pg 7 Ln15-17. What does it mean that 'The American Association for the Study of Liver Diseases (AASLD) guidelines were followed.....until the end of our study'? Does it mean the imaging reference standard adopted before 2013 was different from that after 2013? If so, what's the differences?

Until 2011, according to 2005 AASLD guidelines, the non-invasive diagnosis of HCC required both CT and MRI typical feature for nodules of 1-2 cm and just one technique for nodules > 2 cm. Since 2011 until the end of the study, according to 2011 AASLD and 2013 AISF guidelines, the diagnosis of HCC required only one technique with typical features for nodules > 1 cm.

15) Pg 7 Ln 23-25. How many cases with deep-seated lesions or severe fatty liver were excluded? In these cases, it is difficult to review the main features. As we specified in the manuscript and in Figure 1, 23 cases were excluded from the study due to poor quality nodule visualization at CEUS.

16) Pg 7 Ln 23-25. Were lesions with prior treatment for HCC included?

No, they were not.

17) Pg 8 Ln 21-22. 'The reviewers were blinded to patient identity and to the final diagnosis after CT, MRI or biopsy'. Were the reviewers blinded to study design? Only the external reviewers (G.I and M.A.Z.) were blinded to study design, but all the reviewers were blinded to patient identity and to the final diagnosis.

18) Pg 9 Ln 8-10. How these 50 cases were selected? Randomly or artificially?

Randomly.

19) Pg 9 Ln 20-22. How these estimates of diagnostic accuracy were calculated? Was generalized estimating equations (GEE) used for adjusting aggregation effects? Sensitivity was calculated as $TP/(TP+FN)$, specificity as $TN/(TN+FP)$, PPV as $TP/(TP+FP)$, NPV as $TN/(TN+FN)$, diagnostic accuracy as $(TP+TN)/(TP+TN+FP+FN)$, Youden's index as $sensitivity+specificity-1$, PLR as

sensitivity/(1-specificity), NLR as (1-sensitivity)/specificity and OR as PLR/NLR. 95% CIs were based on binomial distribution.

We did not use a GEE-based approach because the results of CEUS examinations were not correlated. In fact, for each nodule only a single CEUS LI-RADS class was attributed. As we have specified in the Methods section, in case of disagreement between the two internal raters, the class indicated by the more experienced operator was assigned.

20) Results Please show the probability of HCC in each LR category.

We have added a table in the supplementary section (LR3 22.7%, LR4 97.4%, LR5 99.3%, LRM 29.7%)

21) Pg 11 Ln29- Pg12 Ln2. The probability of HCC in CEUS LR-4 category was 97.4%, which was close to the estimate value of LR-5 in other studies concerning CEUS and CT/MRI LI-RADS category (Shin et al. Liver International.2020,DOI: 10.1111/liv.14617;van der Pol et al, Gastroenterology, 2019,DOI: 10.1053/j.gastro.2018.11.020). That might be the cause that the specificity of CEUS LR-4/5 for diagnosing HCC remained 94.3% specificity and 98.8% PPV.

See answer to point n.2

22) Pg 13 Ln 2-4. In this study, all nodules were greater than 10mm. Five nodules (1%) were categorized as CEUS LR-2 rather than LR-3. Please explain it.

Not all nodules were greater than 10mm. As specified in table 1, diameter range was 5-200 mm. For this reason, a few nodules were categorized as CEUS LR-2.

23) Pg 13 Ln 6-21. The interobserver agreement was substantial or almost perfect concerning LI-RADS category in this study. But in a study by Zhou et al (Ultraschall Med. 2020. DOI: 10.1055/a-1168-6321), the inter-reader agreement was not satisfactory concerning CEUS LI-RADS category and washout appearance.

We agree that our interobserver agreement was indeed more satisfactory compared to the study you mentioned. Here are two studies where the interobserver agreement is closer to our findings:

[Li W, et al. Inter-reader agreement of CEUS LI-RADS among radiologists with different levels of experience. Eur Radiol. 2021 Sep;31(9):6758-6767. doi: 10.1007/s00330-021-07777-1]

[Schellhaas B, et al. Interobserver Agreement for Contrast-Enhanced Ultrasound (CEUS)-Based Standardized Algorithms for the Diagnosis of Hepatocellular Carcinoma in High-Risk Patients. Ultraschall Med. 2018 Dec;39(6):667-674. English. doi: 10.1055/a-0612-7887]

24) Discussions Pg 15 Ln 17-19. In the current study, the most frequent CEUS pattern in the arterial phase was rim APHE. But it was only observed in 16% (55/354) cases in a study by Zheng et al (Radiology 2020; 294:299-307. DOI: 10.1148/radiol.2019190878).

In our study, rim APHE was observed in 24.5% of all nodules and in 62.2% of LR-M nodules. We recognize that these percentage are higher compared to the cited study. Nevertheless, the great differences in inclusion/exclusion criteria make difficult to compare the two studies. In fact, we enrolled patients with cirrhosis from different

etiologies, while Zheng et al. only included patient with HBV chronic infection, with or without cirrhosis.

Answer to reviewer #3

Dear reviewer, thank you for your comments.

In response to your specific issues:

- 1) **The fact that only 20% of the tissues were collected degrades the quality of this study.** This is a limitation of our study, as we declared in the text. However, we believe it is not ethical to pursue a biopsy where national and international guidelines allow a non-invasive diagnosis by CT/MRI. An alternative option was to limit our study to the nodules for which pathology was available. However, we rejected this option not only because of the potential loss in the statistical power of the study, but also to avoid a selection bias. In fact, nodules subjected to biopsy are more likely to have an atypical feature at CT/MRI and for this reason are not representative of the totality of nodules developed in cirrhosis. As this is a real-life study, we decided to include both invasively and non-invasively diagnosed nodules.

How CT and MRI were used to classify HCC and ICC should be described in the section of Contrast-enhanced ultrasound examination and CEUS LI-RADS classification.

Non-invasive diagnosis of HCC was obtained with CT and/or MRI. Specifically, nodules showing a dynamic pattern with an hypervascular aspect during the arterial phase followed by wash out in the portal or late phase, were diagnosed as HCC. This diagnostic algorithm is included in both AASLD and AISF guidelines. Until 2011, according to 2005 AASLD guidelines, the non-invasive diagnosis of HCC required both CT and MRI typical feature for nodules of 1-2 cm and just one technique for nodules > 2 cm. Since 2011 until the end of the study, according to 2011 AASLD and 2013 AISF guidelines, the diagnosis of HCC required only one technique with typical features for nodules > 1 cm.

The diagnosis of ICC was based on pathology findings in all cases. ICC was suspected whenever a nodule showed peripheral rim enhancement with progressive homogeneous contrast uptake during different dynamic phases of CT scanning. Ancillary findings included delayed enhancement and capsular retraction. On MR, ICC appear hypointense on T1 weighted and hyperintense on T2 weighted images. Dynamic images show peripheral enhancement in the arterial phase followed by progressive and concentric filling with contrast material. Pooling of contrast on delayed images is indicative of fibrosis.

We have added this additional information to the text.

- 2) **The typical images of typical HCC and ICC from this study should be included in figure as the CT, MRI, and pathology.**
We have added some figures in the supplementary section.
- 3) **In addition, the combined type of HCC and ICC and CoCC are likely to be mixed in this case, and although it is difficult to change the study design from now on, it should be added in the Discussion.**

Our series included a patient with a nodule located in the right lobe and a second nodule in the left lobe. Both nodules were biopsied and pathology results indicated that the first nodule was an HCC, and the second nodule was an ICC. This was the only identified case of combined HCC/ICC of our series, although there is no certainty about further occult cases of combined HCC/ICC among the non-invasively diagnosed nodules. We accept this limit as current guidelines do not recommend routine biopsy of nodules showing typical pattern of HCC at CT/MRI. We have added this to the supplementary section.

Answer to reviewer #4

Dear reviewer, thank you for your comments.
In response to your specific issues:

- 1) **The article is good, but it is too long. I suggest it be shortened.**
Thank you for your revision. We have edited the manuscript in order to make it more fluent.