# Interobserver agreement for contrast-enhanced ultrasound of liver imaging reporting and data system: A systematic review and meta-analysis

Jun Li, Ming Chen, Zi-Jing Wang, Shu-Gang Li, Meng Jiang, Long Shi, Chun-Li Cao, Tian Sang, Xin-Wu Cui, Christoph F Dietrich

**Jun Li, Ming Chen, Zi-Jing Wang, Chun-Li Cao, Tian Sang,** Department of Medical Ultrasound, The First Affiliated Hospital of Medical College, Shihezi University, Shihezi 832008, Xinjiang Uygur Autonomous Region, China

**Shu-Gang Li,** Department of Child, Adolescent Health and Maternal Health, School of Public Health, Capital Medical University, Beijing 100069, Beijing, China

**Meng Jiang,** Department of Medical Ultrasound, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei Province, China

**Long Shi,** Department of Medical Ultrasound, The Second People's Hospital of Jiangmen, Jingmen 448000, Hubei Province, China

**Xin-Wu Cui,** Department of Medical Ultrasound, Sino-German Tongji-Caritas Research Center of Ultrasound in Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei Province, China

**Christoph F Dietrich,** Department of Internal Medicine, Hirslanden Clinic, Berne 27804, Switzerland

**Corresponding author:** Xin-Wu Cui, MD, PhD, Professor, Department of Medical Ultrasound, Sino-German Tongji-Caritas Research Center of Ultrasound in Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1095 Jiefang Avenue, Wuhan 430030, Hubei Province, China. cuixinwu@live.cn

## Abstract

### BACKGROUND
Hepatocellular carcinoma is the most common primary liver malignancy. From the results of previous studies, Liver Imaging Reporting and Data System (LI-RADS) on contrast-enhanced ultrasound (CEUS) has shown satisfactory diagnostic value. However, a unified conclusion on the interobserver stability of this innovative ultrasound imaging has not been determined. The present meta-analysis examined the interobserver agreement of CEUS LI-RADS to provide some reference for subsequent related research.

### AIM

To evaluate the interobserver agreement of LI-RADS on CEUS and analyze the sources of heterogeneity between studies.

*METHODS*
Relevant papers on the subject of interobserver agreement on CEUS LI-RADS published before March 1, 2020 in China and other countries were analyzed. The studies were filtered, and the diagnostic criteria were evaluated. The selected references were analyzed using the "meta" and "metafor" packages of R software version 3.6.2.

*RESULTS*
Eight studies were ultimately included in the present analysis. Meta-analysis results revealed that the summary Kappa value of included studies was 0.76 [95% confidence interval, 0.67-0.83], which shows substantial agreement. Higgins $I^2$ statistics also confirmed the substantial heterogeneity ($I^2$ = 91.30%, 95% confidence interval, 85.3%-94.9%, $P$ < 0.01). Meta-regression identified the variables, including the method of patient enrollment, method of consistency testing, and patient race, which explained the substantial study heterogeneity.

*CONCLUSION*
CEUS LI-RADS demonstrated overall substantial interobserver agreement, but heterogeneous results between studies were also obvious. Further clinical investigations should consider a modified recommendation about the experimental design.

**Key Words:** Contrast-enhanced ultrasound; Liver imaging reporting and data system; Interobserver agreement; Systematic review; Diagnosis; Meta-analysis

**Core Tip:** From the results of previous studies, Liver Imaging Reporting and Data System (LI-RADS) on contrast-enhanced ultrasound (CEUS) has shown a satisfactory diagnostic value. However, a unified conclusion on the interobserver stability of this innovative ultrasound imaging has not been determined. In this article, we included 8 relevant articles to exploring interobserver agreement of LI-RADS on CEUS by making a meta-analysis. Lastly, meta-analysis results revealed that the summary Kappa value of included studies showed substantial agreement, and meta-regression identified the variables, including the method of patient enrolment, method of consistency testing, and patient race, which explained the substantial study heterogeneity.

# INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common primary liver malignancy, and it is the second-most frequent cause of cancer-related deaths[1,2]. HCC often occurs in patients with risk factors, such as chronic hepatitis and cirrhosis[3].

The prognosis of patients with advanced HCC is poor, and curative treatments in patients with early stage HCC are needed[4]. Unlike other systemic malignancies, HCC is diagnosed noninvasively based on imaging characteristics without mandatory pathological confirmation in at-risk patients[5]. Countries with high incidence rates of HCC traditionally screen for high-risk patients using imaging examination, and these patients are closely followed. The focal liver lesions (FLLs) found in HCC screening on various imaging examinations are definitively diagnosed using contrast-enhanced

diagnostic imaging examinations, including contrast-enhanced ultrasound (CEUS), contrast-enhanced computed tomography (CECT), and contrast-enhanced magnetic resonance imaging (CEMRI)[6].

In the abovementioned contrast-enhanced imaging examinations, CEUS allows noninvasive assessments of the contrast enhancement model of HCC without the use of ionizing radiation and with a much higher temporal resolution than computed tomography (CT) and magnetic resonance imaging (MRI)[7,8]. CEUS continues to gain traction as a technique that complements traditional B-mode and Doppler ultrasound in the evaluation of the liver and other organs[9]. CEUS shows changes in microvascular flow mechanics in the focus using real-time imaging of tissue perfusion, which also yields supplementary information, including flow in the microvasculature, slow flow, and perfusion kinetics[10]. CEUS exhibits high accuracy in the differential diagnosis of FLLs in cirrhosis and non-cirrhotic livers[11,12]. CEUS exhibits the same sensitivity and specificity for the differential diagnosis of FLLs, but it is more economical and effective than CECT and CEMRI[13].

Because imaging is important for the diagnosis and treatment decisions in HCC, it is necessary to standardize the imaging diagnosis of HCC and improve its diagnostic accuracy[14,15]. Considering this background, a group of international experts convened by the American College of Radiology (ACR) proposed the Liver Imaging Reporting and Data System (LI-RADS) to standardize the interpretation and reporting of HCC in 2014[16]. The ACR released the CEUS LI-RADS in 2016, with revisions in 2017, and it has become a standardized system for the technique, interpretation, reporting, and data collection for CEUS exams in patients who are at risk for developing HCC[17]. CEUS LI-RADS integrates with the previously released CT/MRI LI-RADS, which provides the criteria for ordinal categories and definitions of the major and ancillary features for HCC[18]. CEUS allows radiologists to (1) use consistent terminology, (2) reduce variability and mistakes in imaging interpretation, (3) promote communication with referring clinicians, and (4) facilitate research and quality assurance[18]. Therefore, the standardized diagnosis helps promote standardization and reproducibility across institutions and radiologists[19].

However, a unified conclusion on the interobserver stability of this innovative ultrasound imaging has not been determined. Several previous studies reported research on the repeatability of LI-RADS on CEUS. The results of some clinical trials revealed strong controversies in interobserver agreement. Schellhaas *et al*[20] demonstrated that the Kappa value was just 0.39, and Tan *et al*[21] showed a Kappa value of 0.94 using LI-RADS on CEUS.

Based on the abovementioned studies, the present meta-analysis examined the interobserver agreement of CEUS LI-RADS to provide some reference for subsequent related research.

## MATERIALS AND METHODS

### *Literature retrieval*
The meta-analysis was performed on relevant literature that was published as late as March 1, 2020 in the databases of PubMed, Web of Science, Embase, China Biology Medicine disc, Cochrane Library, Google Scholar, China National Knowledge Infrastructure, WANFANG databases, and ClinicalTrials.gov. No restriction on language was applied. Search keywords were HCC, CEUS, LI-RADS, and their synonyms: 'liver neoplasm' or 'liver cancer' or 'liver malignancy' or 'hepatocellular carcinoma' or 'HCC'; 'Liver Imaging Reporting and Data System' or 'LI-RADS'; and 'contrast-enhanced ultrasound' or 'CEUS'.

### *Inclusion and exclusion criteria*
**Inclusion criteria:** Research that met all of the following criteria were included: (1) Study types: Observational studies, such as retrospective or prospective; (2) Population: patients at-risk of HCC who needed regular observation, such as patients with cirrhosis and hepatitis B virus carrier; (3) Index tests: CEUS; and (4) Outcomes: sufficiently detailed information to evaluate interobserver agreement for CEUS of LI-RADS.

**Exclusion criteria:** Papers were excluded if they met the following conditions: (1) Editorials, comments, letters, cases reports, and reviews; (2) Experimentation on animals; (3) Repetitive studies and research topic of documents without meeting the requirements; (4) Studies that were not related to the field of interest of the present

research; and (5) Laboratory studies.

### Data extraction

Two reviewers independently screened the appropriate articles according to the inclusion and exclusion criteria detailed above. Discrepancies in opinion between the two reviewers were resolved *via* consult with an additional researcher for reevaluation at a consensus meeting.

The following data were distilled from the included studies using predefined data formats: (1) Article characteristics, including authors, publication years, and study designs (prospective or retrospective); (2) Process characteristics, including enrollment method of patients (selective or consecutive), the number of patients, age and gender ratio, the number of FLLs, and ratio of benign to malignant; (3) Ultrasound (US) system; (4) LI-RADS version; and (5) Reference standard, including pathology and synthesized clinical reference standard (SCRS). The Kappa value for categorical variables was extracted for each major feature and LI-RADS categorization.

### Literature quality evaluation

The bias risk of the included papers was assessed using QUADAS-2 domains[22]. The answers to the symbolic questions of each of the five sections were either "yes", "no", or "unclear" corresponding to the judgment of the risk level of bias as "low", "high", or "uncertain". If the answers to every question were "yes", the study was at "low risk", meanwhile, if the answers to all questions were "no" or "unclear", the study was judged as "high risk". If one of these answers was "no" or "unclear", the study was placed under "uncertain". Revman 5.3, special software for Cochrane collaborative network was used to output the results of QEDAST.

### Statistical analysis

To calculate meta-analysis summary estimations, the Kappa value with standard error categorization was summarized. We estimated standard error from the 95% confidence interval (CI) if it was not mentioned in the original studies. The meta-analysis pooled Kappa value with 95%CI was calculated using the DerSimonian-Laird model with Knapp and Hartung adjustment[23]. According to Landis and Koch, Kappa value was categorized as follows: < 0.20, poor; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, substantial; and 0.81-1.00, almost perfect agreement[24]. Substantial heterogeneity existed if the value of $I^2$ statistics exceeded 50% and the $P$ value did not exceed 0.10 using the Cochran Q-test. A sensitivity analysis or subgroup analysis was performed when heterogeneity was noted, and data synthesis was selected for the random effect model. If the reasons for heterogeneity required further exploration, then a meta-regression analysis was performed using covariates in the bivariate model. Statistical significance was denoted at $P < 0.05$. Funnel plot was used for diagnostic meta-analysis to assess the publication bias of included articles, and significant asymmetry was denoted at $P < 0.10$ for the slope coefficient. We used the "meta" package and "metafor" package in R software version 3.6.2 for analysis and synthesis (R Foundation, Vienna, Austria).

## RESULTS

### Results of the literature search

A detailed flow chart of the study selection process is shown in Figure 1. An aggregate of 129 articles was originally identified using the search strategy, and 54 articles were filtered after excluding duplicates. Twenty-two of the remaining studies were removed, including 19 articles that were unrelated to the field of interest and three review articles. The full texts of the remaining 13 studies were obtained. After review of the full transcripts, five other articles were excluded. Eight studies were ultimately eligible for meta-analysis.

### Characteristics of eligible studies

The meta-analysis included eight studies (five in English, three in Chinese) with a total of 1177 patients and 1379 FLLs[3,6,20,21,25-28]. The major characteristics and basic information of the included articles are detailed in Table 1. The publication dates were 2017 to 2020. The patients who conformed to the inclusion and exclusion criteria from seven studies were Easterners[3,6,21,25-28], and only one article included Westerners[20]. Only one article was a cohort study[26], and the others were retrospective

**Table 1 Characteristics of included studies**

| Ref. | Publication year | Country | Study design | Enrollment method | US system | Number of patients (male/female) | Age in yr, mean ± SD | Number of FLLs (benign/malignant) | LI-RADS version | Reference standard |
|---|---|---|---|---|---|---|---|---|---|---|
| Chen *et al*[25] | 2017 | China | Retrospective | Consecutive | Hitachi HI Vision Preirus; Siemens S3000 | 377 (207/170) | 43.56 11.48 | 429 (96/333) | 2016 | Pathology and SCRS |
| Schellhaas *et al*[20] | 2018 | Germany | Retrospective | Consecutive | Siemens S2000; GE Logiq E9; Toshiba Aplio 500 | 55 (44/11) | 65.90 (53-86)[1] | 55 (5/50) | 2016 | Pathology and SCRS |
| Ling *et al*[28] | 2018 | China | Retrospective | Consecutive | Philips IU22 | 56 (44/12) | 52.5[2] | 56 (10/46) | 2016 | Pathology |
| Liu *et al*[26] | 2019 | China | Prospective | Consecutive | Philips IU22; Toshiba Aplio 500 | 82 (61/21) | 58.15 10.97 | 82 (34/48) | 2017 | Pathology and SCRS |
| Zhou *et al*[27] | 2019 | China | Retrospective | Consecutive | GE Logiq E9 | 241 (166/75) | 56 10 | 272 (44/228) | 2016 | Pathology and SCRS |
| Tan *et al*[21] | 2020 | Singapore | Retrospective | Selective | GE Logiq E9; Toshiba Aplio 500 | 45 (32/13) | 63.1 (34-84) | 46 (9/37) | 2017 | Pathology and SCRS |
| Wang *et al*[6] | 2020 | China | Retrospective | Consecutive | Hitachi Hi Vision Avius | 258 (200/58) | 52 11 | 355 (235/120) | 2017 | Pathology and SCRS |
| Cui *et al*[3] | 2020 | China | Retrospective | Selective | Hitachi Hi Vision Avius | 63 (47/16) | 56 8 | 84 (35/49) | 2017 | Pathology and SCRS |

[1]This study only included age range and mean.

[2]This study used the median of age.

[3]This is our study. SCRS: Synthesized clinical reference standard; FLLS: Focal liver lesions; LI-RADS: Liver Imaging Reporting and Data System; SD: Standard deviation; US: Ultrasound.

studies[3,6,20,21,25,27,28]. The enrollment method of one article was selective cohort[21], and the other articles used consecutive cohorts[3,6,20-28]. The classification standard for CEUS of the four articles was the LI-RADS 2016 version[20,25,27,28], and the other studies selected the LI-RADS 2017 version[3,6,21,26]. The two articles published in the same year were from the same first author, but the samples and methods selected were not the same[3,6]. To distinguish these studies in the present meta-analysis, the article published in the *Journal of Ultrasound in Medicine* is identified using the first author (Wang *et al*[6]), and the article published in *Ultrasound Medicine Biology* is identified using the corresponding author (Cui *et al*[3]). Three articles were dissertations[25-27]. Only one study used pathology as a reference standard[28], and the other seven studies used a combination of pathology and SCRS[3,6,20,21,25-27].

### Literature quality evaluation

Results of the quality assessment of the included articles are shown in Figure 2. The results revealed that relatively acceptable quality evaluations could be acquired from the involved studies.

### Pooled interobserver agreement of LI-RADS for CEUS

The 95%CI of Kappa values of each study and the combined Kappa value of all included studies were estimated using analysis software, and the results are indicated
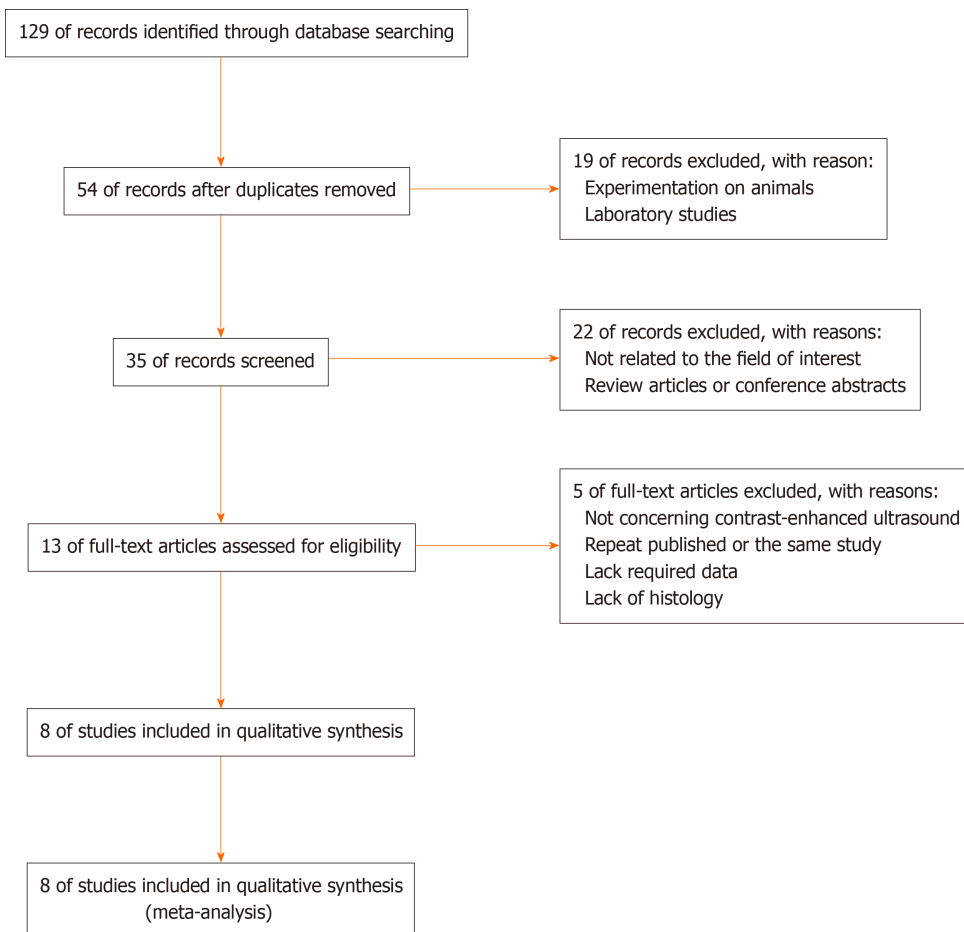
```
┌─────────────────────────────────────────────┐
│ 129 of records identified through database    │
│ searching                                     │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────┐        ┌────────────────────────────────────┐
│ 54 of records after duplicates   │───────▶│ 19 of records excluded, with reason: │
│ removed                          │        │   Experimentation on animals         │
└─────────────────────────────────┘        │   Laboratory studies                 │
                      │                      └────────────────────────────────────┘
                      ▼
┌─────────────────────────────────┐        ┌────────────────────────────────────┐
│ 35 of records screened           │───────▶│ 22 of records excluded, with reasons:│
└─────────────────────────────────┘        │   Not related to the field of interest│
                      │                      │   Review articles or conference      │
                      │                      │   abstracts                          │
                      │                      └────────────────────────────────────┘
                      ▼
┌─────────────────────────────────┐        ┌────────────────────────────────────┐
│ 13 of full-text articles assessed│───────▶│ 5 of full-text articles excluded,    │
│ for eligibility                  │        │ with reasons:                        │
└─────────────────────────────────┘        │   Not concerning contrast-enhanced   │
                      │                      │   ultrasound                         │
                      │                      │   Repeat published or the same study │
                      │                      │   Lack required data                 │
                      │                      │   Lack of histology                  │
                      │                      └────────────────────────────────────┘
                      ▼
┌─────────────────────────────────┐
│ 8 of studies included in         │
│ qualitative synthesis            │
└─────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────┐
│ 8 of studies included in         │
│ qualitative synthesis            │
│ (meta-analysis)                  │
└─────────────────────────────────┘
```

**Figure 1  Study flow chart.**

in Table 2 and Table 3. Tan *et al*[21] concluded that the interobserver agreement of LI-RADS was in near perfect agreement (Kappa value = 0.94; 95%CI, 0.89-0.97). The Higgins $I^2$ statistics indicated substantial heterogeneity in the summary Kappa value ($I^2$ = 91.30%, 95%CI, 85.3%-94.9%, $P < 0.01$). Therefore, the pooled calculation used a random effect model. The pooled Kappa value from the random effect model was 0.76 (95%CI, 0.67-0.83), which showed substantial agreement. The forest plot of summary Kappa values is shown in Figure 3.

**Publication bias**
The *P* value of the linear regression test of funnel plot asymmetry was 0.39, and we determined that no publication bias existed in the statistics. The Egger's funnel plot is shown in Figure 4.

**Sensitivity analysis**
One study at a time was omitted from analysis, and the results of the sensitivity analysis are shown in Table 4, which showed no literature influence. Exclusion of the included articles one by one revealed that the Higgins $I^2$ also did not change significantly. The consolidation result was relatively stable.

**Meta-regression analysis**
Because of the strong heterogeneity of pooled studies in the merger statistics, a meta-regression was used. The meta-regression analysis analyzed some clinically relevant variables of investigation, including study design (retrospective or prospective), method of patient enrollment (consecutive or selective), LI-RADS version (2016 or 2017), number of interobservers (two or three), number of US systems used (one or more), race of chosen patients (Easterners or Westerners), number of diagnosing FLLs (less than 100 or more than 100), and reference standard (only pathology or combination of pathology and SCRS). The results of regression analysis are exhibited in Table 5. The variables of method of patient enrollment ($P < 0.01$), number of

## Table 2 Analysis of interobserver agreement for included studies

| Ref. | Kappa value (95%CI) | Weight in fixed model, % | Weight in random model, % |
|---|---|---|---|
| Chen *et al*[25] | 0.81 (0.77, 0.84) | 31.40 | 14.00 |
| Schellhaas *et al*[20] | 0.39 (0.14, 0.59) | 3.80 | 11.40 |
| Ling *et al*[28] | 0.69 (0.52, 0.81) | 3.90 | 11.40 |
| Liu *et al*[26] | 0.61 (0.45, 0.73) | 5.80 | 12.30 |
| Zhou *et al*[27] | 0.87 (0.84, 0.90) | 19.90 | 13.80 |
| Tan *et al*[21] | 0.94 (0.89, 0.97) | 3.20 | 10.90 |
| Wang *et al*[6] | 0.77 (0.72, 0.81) | 26.00 | 13.90 |
| Cui *et al*[1] | 0.67 (0.53, 0.77) | 6.00 | 12.30 |

[1]This is our study. CI: Confidence interval.

## Table 3 Interobserver agreement of pooled included studies

| Summary type | Kappa value (95%CI) | Z | P value |
|---|---|---|---|
| Fixed effects model | 0.79 (0.77, 0.81) | 39.68 | < 0.01 |
| Random effects model | 0.76 (0.67, 0.83) | 9.99 | < 0.01 |

CI: Confidence interval.

## Table 4 Sensitivity analysis eliminating studies one by one

| Ref. | Kappa value (95%CI) | $I^2$ (95%CI), % | P value |
|---|---|---|---|
| Chen *et al*[25] | 0.75 (0.62; 0.84) | 92.42 (87.07; 95.57) | < 0.01 |
| Schellhaas *et al*[20] | 0.79 (0.71; 0.85) | 89.44 (80.83; 94.22) | < 0.01 |
| Ling *et al*[28] | 0.77 (0.67; 0.84) | 92.30 (86.72; 95.58) | < 0.01 |
| Liu *et al*[26] | 0.78 (0.69; 0.85) | 91.44 (84.80; 95.13) | < 0.01 |
| Zhou *et al*[27] | 0.74 (0.63; 0.82) | 89.78 (81.61; 94.36) | < 0.01 |
| Tan *et al*[21] | 0.73 (0.62; 0.80) | 90.21 (82.51; 94.60) | < 0.01 |
| Wang *et al*[6] | 0.76 (0.64; 0.85) | 92.41 (86.90; 95.59] | < 0.01 |
| Cui *et al*[1] | 0.77 (0.67; 0.85) | 92.04 (86.02; 95.43) | < 0.01 |

[1]This is our study. CI: Confidence interval.

interobservers ($P < 0.01$), and race of the chosen patients ($P < 0.01$) had significant statistical significance for the Kappa value. The Kappa value was 0.94 (95%CI, 0.89-0.97) in the study that used selective patient enrollment[21], and the pooled Kappa value was 0.73 (95%CI, 0.62-0.80) in the articles that used consecutive patient enrollment[3,6,20,25-28]. Estimations of the agreement of three interobservers indicated fair agreement (Kappa value = 0.39; 95%CI, 0.14-0.59)[20], and the consistency of two interobservers showed substantial agreement[3,6,21,25-28]. The race of chosen patients also contributed to the strong heterogeneity. The covariates of study design ($P = 0.26$), LI-RADS version ($P = 0.66$), number of US systems used ($P = 0.93$), number of FLLs ($P = 0.16$), and reference standard ($P = 0.58$) did not cause statistical significance in the heterogeneity test.

**Table 5 Results of meta-regression analysis of interobserver agreement of Liver Imaging Reporting and Data System in contrast-enhanced ultrasound**

| Covariate and subgroup | Number of included studies | Kappa value (95%CI) | *P* value |
| --- | --- | --- | --- |
| Study design | | | 0.26 |
| Retrospective | 7 | 0.78 (0.69, 0.85) | |
| Prospective | 1 | 0.61 (0.45, 0.73) | |
| Enrollment method | | | < 0.01 |
| Consecutive | 7 | 0.73 (0.62, 0.80) | |
| Selective | 1 | 0.94 (0.89, 0.97) | |
| LI-RADS version | | | 0.66 |
| 2016 | 4 | 0.74 (0.58, 0.85) | |
| 2017 | 4 | 0.78 (0.62, 0.88) | |
| Number of interobservers | | | < 0.01 |
| 2 | 7 | 0.79 (0.71, 0.85) | |
| 3 | 1 | 0.39 (0.14, 0.59) | |
| Number of US systems | | | 0.93 |
| 1 | 4 | 0.77 (0.65, 0.85) | |
| > 1 | 4 | 0.76 (0.51, 0.89) | |
| Race of patients | | | < 0.01 |
| Westerners | 1 | 0.79 (0.71, 0.85) | |
| Easterners | 7 | 0.39 (0.14, 0.59) | |
| Number of FLLs | | | 0.16 |
| < 100 | 5 | 0.72 (0.48, 0.86) | |
| > 100 | 3 | 0.82 (0.76, 0.87) | |
| Reference standard | | | 0.58 |
| Pathology | 1 | 0.69 (0.52, 0.81) | |
| Pathology and SCRS | 7 | 0.77 (0.67, 0.84) | |

FLL: Focal hepatic lesion; LI-RADS: Liver Imaging Reporting and Data System; SCRS: Synthesized clinical reference standard; US: Ultrasound.

## DISCUSSION

Meta-analysis of interobserver agreement of LI-RADS on CEUS was not reported previously. The summary Kappa value for the eight included studies was 0.76 (95%CI, 0.67-0.83) in our study, which showed substantial inter-reader agreement for the use of LI-RADS on CEUS.

The CEMRI is another common noninvasive imaging method to assess benign and malignant FLLs. Notably, the LI-RADS on MRI to evaluate FLLs was developed in 2011 and recently updated in 2018[29]. The meta-analysis of Kang *et al*[30] revealed that summary interobserver agreement of LI-RADS on MRI was 0.70 (95%CI, 0.56-0.85). Another multicenter international study, which used a large number of readers and a mixture of all LI-RADS category assignments, obtained a similar result (Kappa value = 0.73; 95%CI, 0.68-0.77)[19]. The interobserver agreement of LI-RADS on CEUS seems better than that of LI-RADS on MRI. Notably, CEUS avoids the disadvantages of MRI, such as high expense and a long inspection time. Some researchers demonstrated that the sensitivity of CEUS in the observation of arterial hypervascularity from nodules in liver cirrhosis was significantly higher than that of MRI[31-33]. Two recent meta-analyses of CEUS showed excellent diagnostic accuracy in differentiating malignant from benign FLLs with a summary sensitivity of 0.92 and summary specificity of 0.87, and the sensitivity of CEMRI was slightly weaker than that of CEUS with a pooled sensitivity of 0.86 and pooled specificity of 0.89[34,35].
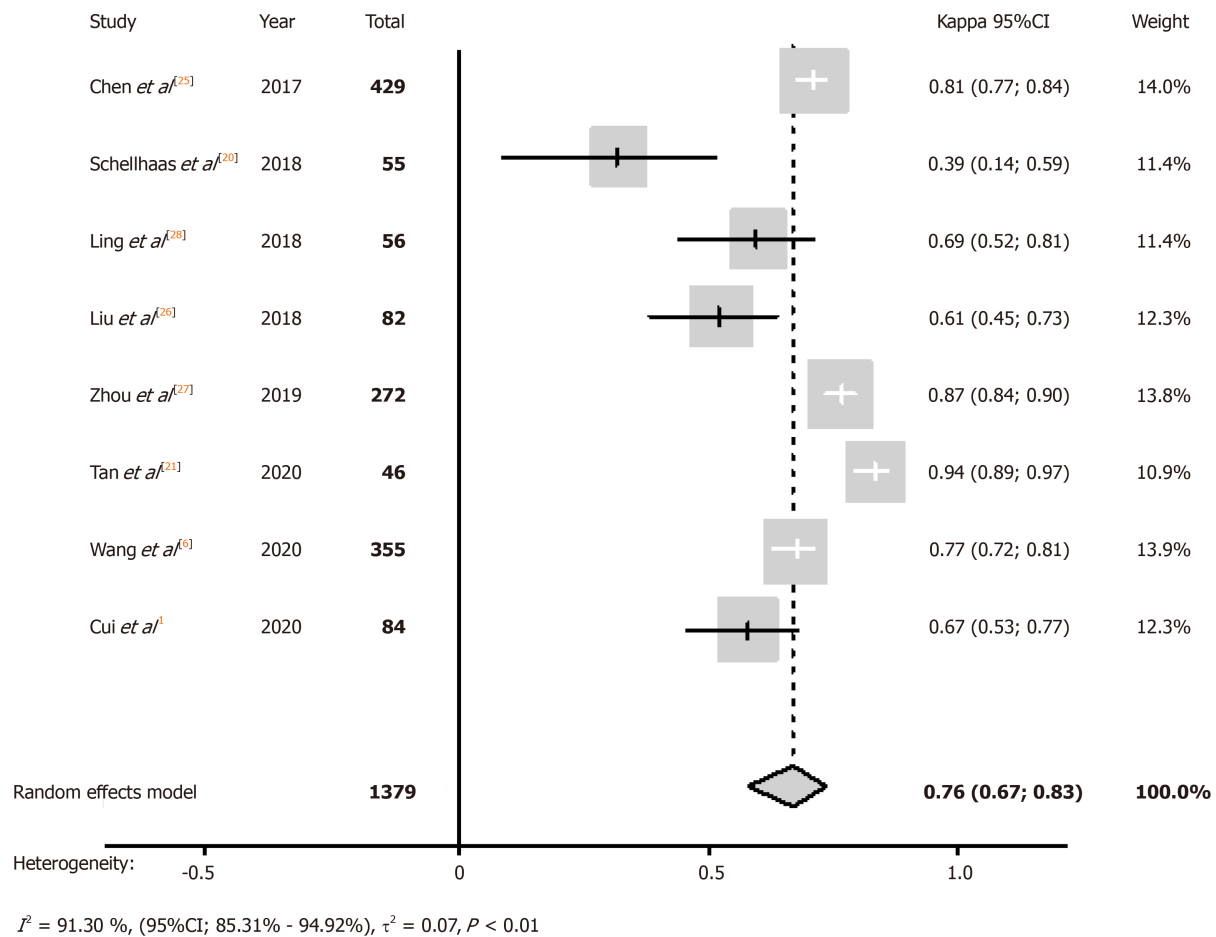
**Figure 2 Bias risk of the included studies (QUADAS 2 criteria).** [1]This is our study.

The majority of HCCs are not suitable for curative resection at the time of treatment, and difficulties of surgical resection may be related to size, site, and number of tumors, vascular and extrahepatic involvement as well as liver function of the patient[36]. Radiofrequency ablation (RFA) is another effective treatment for liver cancer, and it has emerged in clinical practice to expand the pool of patients considered for liver-directed therapies[37]. Traditionally, RFA is usually performed under B-mode US guidance. In recent years, some scholars have reported the treatment technique of RFA guided by CEUS for HCC. Miyamoto *et al*[38] exhibited the complete ablation rate after a single treatment session was significantly higher in CEUS group than in the B-mode US group. Moreover, Masuzaki *et al*[39] reported in a large-scale study that the detectability of tumor nodules was 83.5% in B-mode US and 93.2% in CEUS ($P = 0.04$). Therefore, the use of CEUS guidance in RFA for liver cancer is an efficient approach.

The eight studies included in the present meta-analysis also exhibited some problems. For example, most studies did not list the interobserver agreement for the major features of CEUS in detail. Notably, the CEUS LI-RADS criteria requires the combination of two major features, including arterial phase hyperenhancement (APHE) and washout, to distinguish benign and malignant FLLs[17]. Unfortunately, just two articles mentioned the Kappa value of APHE and washout[3,20]. Therefore, it is recommended that further research on the interobserver agreement of CEUS LI-RADS add an extra consistency test of the major features of CEUS to increase the persuasiveness of the research.

The present meta-analysis used meta-regression to compensate for the high heterogeneity and analyzed existing covariates due to several potential causes. The method of patient enrollment had a significant impact on the Kappa value. Tan *et al*[21] used selective screening as an inclusion standard rather than consecutive screening, which was different from the other seven articles, and the Kappa value of his research was obviously higher than the summary values of the other articles (0.94 *vs* 0.73, $P <$ 0.01). However, selective screening of patients caused a small sample size of FLLs, which may introduce potential confounders and bias.

Consistency tests of the eight included studies revealed that the number of reviewers also affected heterogeneity. Schellhaas *et al*[20] used two reviewers for interobserver agreement and achieved a satisfactory Kappa value using pairwise comparisons, and this Kappa value was lower than in the other articles (0.39 *vs* 0.79, $P$

| Study | Year | Total | | Kappa 95%CI | Weight |
|---|---|---|---|---|---|
| Chen *et al*[25] | 2017 | **429** | | 0.81 (0.77; 0.84) | 14.0% |
| Schellhaas *et al*[20] | 2018 | **55** | | 0.39 (0.14; 0.59) | 11.4% |
| Ling *et al*[28] | 2018 | **56** | | 0.69 (0.52; 0.81) | 11.4% |
| Liu *et al*[26] | 2018 | **82** | | 0.61 (0.45; 0.73) | 12.3% |
| Zhou *et al*[27] | 2019 | **272** | | 0.87 (0.84; 0.90) | 13.8% |
| Tan *et al*[21] | 2020 | **46** | | 0.94 (0.89; 0.97) | 10.9% |
| Wang *et al*[6] | 2020 | **355** | | 0.77 (0.72; 0.81) | 13.9% |
| Cui *et al*[1] | 2020 | **84** | | 0.67 (0.53; 0.77) | 12.3% |
| Random effects model | | 1379 | | 0.76 (0.67; 0.83) | **100.0%** |

Heterogeneity:

$I^2$ = 91.30 %, (95%CI; 85.31% - 94.92%), $\tau^2$ = 0.07, $P$ < 0.01

**Figure 3 Forest plot of pooled interobserver agreement for Liver Imaging Reporting and Data System on contrast-enhanced ultrasound.** [1] This is our study. CI: Confidence interval.
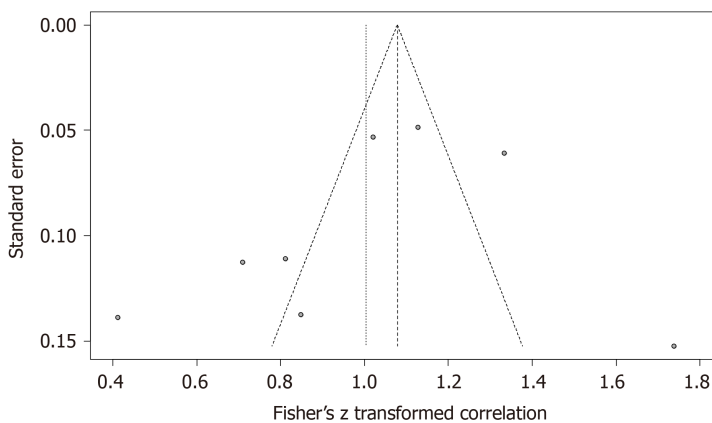


**Figure 4  Publication bias of the included studies.**

< 0.01). The low Kappa value may be because the calculation of Kappa relies on the assumption that a significant proportion of agreement is due to chance, and if a feature is observed very frequently, then a low Kappa value between the observers results[40]. The author indicated that the reason for the low Kappa value was that the calculation of Cohen's Kappa was influenced by the frequency of a certain feature being observed[20]. Therefore, the use of intraclass coefficient correlation rather than Cohen's Kappa to represent interobserver agreement in multiple observers would be more satisfactory.

The race of enrolled patients may contribute to the heterogeneity in pooled analysis. Epidemiological surveys showed that chronic hepatitis B virus and hepatitis C virus

infection led to HCC in eastern Asia and sub-Saharan Africa, and non-alcoholic fatty liver disease was the major precipitating factor of HCC in Western countries[41]. Non-alcoholic fatty liver diseases (NAFLD) are one of the most prevalent causes of chronic liver diseases in Western countries, with an estimated prevalence of 20%-40%[42]. In the context of fatty liver, the diagnosis of HCC that progressed from NAFLD may be more difficult. However, there is no relevant literature on this issue to research in the databases. Therefore, the diagnosing of HCC progressed from NAFLD and viral stimulation on CEUS is a hot issue.

Because most of the included articles were retrospective studies, the US system and reference standard of these studies were partially diversified. However, the meta-regression analysis showed that the study design (0.78 *vs* 0.61, *P* = 0.26), the number of US systems (0.77 *vs* 0.76, *P* = 0.93), the number of FLLs (0.72 *vs* 0.82, *P* = 0.16), and reference standard (0.69 *vs* 0.77, *P* = 0.58) did not reach statistical significance for the heterogeneity, which confirmed that the use of LI-RADS on CEUS for the diagnosing of FLLs was stable laterally.

Notably, reliable interobserver agreements for LI-RADS categorization on CEUS were also observed for LI-RADS version 2016 and version 2017 (0.74 *vs* 0.78, *P* = 0.66). Because LI-RADS included lexicons, minute definitions and illustrations for imaging features, a high interobserver agreement may be achieved. The two versions of LI-RADS on CEUS define APHE and washout clearly. For example, a feature of APHE may be considered present if it is demonstrated in the entire nodule or only a portion of the nodule[43]. CEUS characterization of washout requires assessment of its onset (late *vs* early) and degree (mild *vs* marked), not just its presence[17]. Generally, early (< 60 s) and/or marked washout is a major feature for LR-M, and late (≥ 60 s) and mild washout is a major feature for HCC[44,45]. The diversity of version did not cause heterogeneity of statistical significance, and it was sufficient to authenticate the diagnosis stability of the two versions of LI-RADS on interobserver agreement.

Measurements may differ due to bias between reviewers, which may be explained by the use of the reader's judgment of test results of a measurement[46]. The bias may result from differences in training, learning, and experience between reviewers. The consistency test demonstrated that reviewers in similar working environments used diagnostic reasoning, which may be reduced with continuous education and updated definitions in LI-RADS on CEUS.

The present meta-analysis has several limitations. First, substantial research heterogeneity was mentioned. However, three significant factors for study heterogeneity were found using meta-regression. Second, the different study designs and participants from diverse geographic locations led to the heterogeneous distribution of the disease.

## CONCLUSION

In conclusion, the present meta-analysis produced some results. Summary results showed substantial interobserver agreement for LI-RADS on CEUS. The heterogeneity factors included the method of enrolling patients, the method of consistency testing, and the race of patients, which should be considered in subsequent study design. Certainly, a large, prospective, and multicenter study is also needed to confirm our results.

## ARTICLE HIGHLIGHTS

### Research background

From the results of previous studies, Liver Imaging Reporting and Data System (LI-RADS) on contrast-enhanced ultrasound (CEUS) for diagnosing hepatocellular carcinoma (HCC) has shown a satisfactory diagnostic value. However, a unified conclusion on the interobserver stability of this innovative ultrasound imaging has not been determined. The present meta-analysis examined the interobserver agreement of CEUS LI-RADS to provide some reference for subsequent related research.

### Research motivation

According to the inclusion and exclusion criteria, we included eight relevant articles to explore interobserver agreement of LI-RADS on CEUS by making a meta-analysis. Finally, meta-analysis results revealed that the summary Kappa value of included

studies showed substantial agreement. The heterogeneity factors included the method of enrolling patients, the method of consistency testing, and the race of patients, which should be considered in subsequent study design.

### Research objectives

The main objective of the present article is to explore interobserver agreement of LI-RADS on CEUS for diagnosing HCC. Results of meta-analysis showed interobserver agreement is substantial and the heterogeneity factors included the method of enrolling patients, the method of consistency testing, and the race of patients, which should be considered in subsequent study design.

### Research methods

The method of this article is to calculate Kappa value to estimate interobserver agreement of LI-RADS on CEUS for diagnosing HCC by using the "meta" package and "metafor" package in R software version 3.6.2 for analysis and synthesis (R Foundation, Vienna, Austria). The result of consistency test has a vital reference value for the stability of LI-RADS.

### Research results

This article exhibited substantial interobserver agreement for LI-RADS on CEUS. In addition, meta-regression identified several heterogeneity factors, including the method of enrolling patients, the method of consistency testing, and the race of patients, which should be considered in subsequent study design. Meanwhile, a large, prospective, and multicenter related study is also needed to confirm our results.

### Research conclusions

This study reported that interobserver agreement for LI-RADS on CEUS was substantial and that the method of enrolling patients, the method of consistency testing, and the race of patients perhaps interfere with interobserver agreement, which should be considered in subsequent study design.

### Research perspectives

The method of enrolling patients, the method of consistency testing, and the race of patients perhaps interfere interobserver agreement and should be considered in future research about LI-RADS on CEUS for diagnosing HCC.

## REFERENCES

1       **El-Serag HB**. Hepatocellular carcinoma. *N Engl J Med* 2011; **365**: 1118-1127 [PMID: 21992124 DOI: 10.1056/NEJMra1001683]

2       **Chernyak V**, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, Kono Y, Do RK, Mitchell DG, Singal AG, Tang A, Sirlin CB. Liver Imaging Reporting and Data System (LI-RADS) Version 2018: Imaging of Hepatocellular Carcinoma in At-Risk Patients. *Radiology* 2018; **289**: 816-830 [PMID: 30251931 DOI: 10.1148/radiol.2018181494]

3       **Wang JY**, Feng SY, Yi AJ, Zhu D, Xu JW, Li J, Cui XW, Dietrich CF. Comparison of Contrast-Enhanced Ultrasound *vs* Contrast-Enhanced Magnetic Resonance Imaging for the Diagnosis of Focal Liver Lesions Using the Liver Imaging Reporting and Data System. *Ultrasound Med Biol* 2020; **46**: 1216-1223 [PMID: 32115307 DOI: 10.1016/j.ultrasmedbio.2020.01.023]

4       **Lu DS**, Yu NC, Raman SS, Limanond P, Lassman C, Murray K, Tong MJ, Amado RG, Busuttil RW. Radiofrequency ablation of hepatocellular carcinoma: treatment success as defined by histologic examination of the explanted liver. *Radiology* 2005; **234**: 954-960 [PMID: 15681691 DOI: 10.1148/radiol.2343040153]

5       **Bruix J**, Sherman M; American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma: an update. *Hepatology* 2011; **53**: 1020-1022 [PMID: 21374666 DOI: 10.1002/hep.24199]

6       **Wang JY**, Feng SY, Xu JW, Li J, Chu L, Cui XW, Dietrich CF. Usefulness of the Contrast-Enhanced Ultrasound Liver Imaging Reporting and Data System in Diagnosing Focal Liver Lesions by Inexperienced Radiologists. *J Ultrasound Med* 2020; Online ahead of print [PMID: 32078173 DOI: 10.1002/jum.15242]

7       **Bartolotta TV**, Taibbi A, Midiri M, La Grutta L, De Maria M, Lagalla R. Characterisation of focal liver lesions undetermined at grey-scale US: contrast-enhanced US *vs* 64-row MDCT and MRI with liver-specific contrast agent. *Radiol Med* 2010; **115**: 714-731 [PMID: 20082225 DOI: 10.1007/s11547-010-0506-3]

8       **Bartolotta TV**, Vernuccio F, Taibbi A, Lagalla R. Contrast-Enhanced Ultrasound in Focal Liver Lesions: Where Do We Stand? *Semin Ultrasound CT MR* 2016; **37**: 573-586 [PMID: 27986175 DOI: 10.1053/j.sult.2016.10.003]

9       **Pang EHT**, Chan A, Ho SG, Harris AC. Contrast-Enhanced Ultrasound of the Liver: Optimizing Technique and Clinical Applications. *AJR Am J Roentgenol* 2018; **210**: 320-332 [PMID: 29220210 DOI: 10.2214/AJR.17.17843]

10      **Paefgen V**, Doleschel D, Kiessling F. Evolution of contrast agents for ultrasound imaging and ultrasound-

mediated drug delivery. *Front Pharmacol* 2015; **6**: 197 [PMID: 26441654 DOI: 10.3389/fphar.2015.00197]

11   **Wildner D**, Bernatik T, Greis C, Seitz K, Neurath MF, Strobel D. 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* 2015; **36**: 132-139 [PMID: 25812115 DOI: 10.1055/s-0034-1399147]

12   **Darnell A**, Forner A, Rimola J, Reig M, GarcÃa-Criado Ã, Ayuso C, Bruix J. Liver Imaging Reporting and Data System with MR Imaging: Evaluation in Nodules 20 mm or Smaller Detected in Cirrhosis at Screening US. *Radiology* 2015; **275**: 698-707 [PMID: 25658038 DOI: 10.1148/radiol.15141132]

13   **Westwood M**, Joore M, Grutters J, Redekop K, Armstrong N, Lee K, Gloy V, Raatz H, Misso K, Severens J, Kleijnen J. 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 [PMID: 23611316 DOI: 10.3310/hta17160]

14   **Clavien PA**, Lesurtel M, Bossuyt PM, Gores GJ, Langer B, Perrier A; OLT for HCC Consensus Group. Recommendations for liver transplantation for hepatocellular carcinoma: an international consensus conference report. *Lancet Oncol* 2012; **13**: e11-e22 [PMID: 22047762 DOI: 10.1016/S1470-2045(11)70175-9]

15   **Wald C**, Russo MW, Heimbach JK, Hussain HK, Pomfret EA, Bruix J. New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. *Radiology* 2013; **266**: 376-382 [PMID: 23362092 DOI: 10.1148/radiol.12121698]

16   **Kim TK**, Noh SY, Wilson SR, Kono Y, Piscaglia F, Jang HJ, Lyshchik A, Dietrich CF, Willmann JK, Vezeridis A, Sirlin CB. Contrast-enhanced ultrasound (CEUS) liver imaging reporting and data system (LI-RADS) 2017 - a review of important differences compared to the CT/MRI system. *Clin Mol Hepatol* 2017; **23**: 280-289 [PMID: 28911220 DOI: 10.3350/cmh.2017.0037]

17   **Bartolotta TV**, Terranova MC, Gagliardo C, Taibbi A. CEUS LI-RADS: a pictorial review. *Insights Imaging* 2020; **11**: 9 [PMID: 32020352 DOI: 10.1186/s13244-019-0819-2]

18   **Elsayes KM**, Hooker JC, Agrons MM, Kielar AZ, Tang A, Fowler KJ, Chernyak V, Bashir MR, Kono Y, Do RK, Mitchell DG, Kamaya A, Hecht EM, Sirlin CB. 2017 Version of LI-RADS for CT and MR Imaging: An Update. *Radiographics* 2017; **37**: 1994-2017 [PMID: 29131761 DOI: 10.1148/rg.2017170098]

19   **Fowler KJ**, Tang A, Santillan C, Bhargavan-Chatfield M, Heiken J, Jha RC, Weinreb J, Hussain H, Mitchell DG, Bashir MR, Costa EAC, Cunha GM, Coombs L, Wolfson T, Gamst AC, Brancatelli G, Yeh B, Sirlin CB. Interreader Reliability of LI-RADS Version 2014 Algorithm and Imaging Features for Diagnosis of Hepatocellular Carcinoma: A Large International Multireader Study. *Radiology* 2018; **286**: 173-185 [PMID: 29091751 DOI: 10.1148/radiol.2017170376]

20   **Schellhaas B**, Pfeifer L, Kielisch C, Goertz RS, Neurath MF, Strobel D. Interobserver Agreement for Contrast-Enhanced Ultrasound (CEUS)-Based Standardized Algorithms for the Diagnosis of Hepatocellular Carcinoma in High-Risk Patients. *Ultraschall Med* 2018; **39**: 667-674 [PMID: 29879746 DOI: 10.1055/a-0612-7887]

21   **Tan Z**, Teoh WC, Wong KM, Wansaicheong GK, Sandrasegaran K. Analysis of comparative performance of CEUS and CECT/MR LI-RADS classification: Can CEUS dichotomize LI-RADS indeterminate lesions on CT or MRI? *Clin Imaging* 2020; **62**: 63-68 [PMID: 32070868 DOI: 10.1016/j.clinimag.2020.02.002]

22   **Whiting PF**, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529-536 [PMID: 22007046 DOI: 10.7326/0003-4819-155-8-201110180-00009]

23   **IntHout J**, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014; **14**: 25 [PMID: 24548571 DOI: 10.1186/1471-2288-14-25]

24   **Landis JR**, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159-174 [PMID: 843571 DOI: 10.2307/2529310]

25   **Chen HQ**. The Clinical Application of Liver Ultrasound Imaging Reporting and Date System Classification Method of Liver Occupying Lesions. M.M. Thesis, Fujian Medical University. 2017. Available from: https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201802&filename=1017857964.nh&v=nawQoHiuqktbqtjkgUKGV%25mmd2BB8ItXjiPz8mYJScOmbY1yhfEV4FIK6dK20i1c9z8jk

26   **Liu QY**. The clinical application Value of liver imaging reporting and data system with contrast-enhanced ultrasound in the diagnosis of hepatocellular carcinoma. M.M. Thesis, Zhengzhou University. 2019. Available from: https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201902&filename=1019115498.nh&v=7swi6YCLhrWorznJSACxtiK4RphRy5jP4GUpIWjJ5nAVc25Q8eX6BuadlSLi08Wm

27   **Zhou YN**. Contrast-enhanced ultrasound (CEUS) liver imaging reporting and data system (L1-RADS) LR-5 for Hepatocellular carcinoma-diagnostic value in clinical practice. M.M. Thesis, Jilin University. 2019. Available from: https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201902&filename=1019159698.nh&v=P1allGVwHZq2ROBMMzOdWaIQrR0KodmJYjDVmYZAQoDjdnyV8soLTmPXqH2aVd25

28   **Ling W**, Wang M, Ma X, Qiu T, Li J, Lu Q, Luo Y. The preliminary application of liver imaging reporting and data system (LI-RADS) with contrast-enhanced ultrasound (CEUS) on small hepatic nodules (≤ 2cm). *J Cancer* 2018; **9**: 2946-2952 [PMID: 30123363 DOI: 10.7150/jca.25539]

29   **Mitchell DG**, Bruix J, Sherman M, Sirlin CB. 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 [PMID: 25041904 DOI: 10.1002/hep.27304]

30   **Kang JH**, Choi SH, Lee JS, Park SH, Kim KW, Kim SY, Lee SS, Byun JH. Interreader Agreement of Liver Imaging Reporting and Data System on MRI: A Systematic Review and Meta-Analysis. *J Magn Reson Imaging* 2020; **52**: 795-804 [PMID: 31984578 DOI: 10.1002/jmri.27065]

31    **Maruyama H**, Takahashi M, Ishibashi H, Yoshikawa M, Yokosuka O. Contrast-enhanced ultrasound for characterisation of hepatic lesions appearing non-hypervascular on CT in chronic liver diseases. *Br J Radiol* 2012; **85**: 351-357 [PMID: 21224305 DOI: 10.1259/bjr/20440141]

32    **Takahashi M**, Maruyama H, Shimada T, Kamezaki H, Sekimoto T, Kanai F, Yokosuka O. Characterization of hepatic lesions (≤ 30 mm) with liver-specific contrast agents: a comparison between ultrasound and magnetic resonance imaging. *Eur J Radiol* 2013; **82**: 75-84 [PMID: 23116806 DOI: 10.1016/j.ejrad.2012.05.035]

33    **Sugimoto K**, Moriyasu F, Shiraishi J, Saito K, Taira J, Saguchi T, Imai Y. Assessment of arterial hypervascularity of hepatocellular carcinoma: comparison of contrast-enhanced US and gadoxetate disodium-enhanced MR imaging. *Eur Radiol* 2012; **22**: 1205-1213 [PMID: 22270142 DOI: 10.1007/s00330-011-2372-3]

34    **Wu M**, Li L, Wang J, Zhang Y, Guo Q, Li X, Zhang X. Contrast-enhanced US for characterization of focal liver lesions: a comprehensive meta-analysis. *Eur Radiol* 2018; **28**: 2077-2088 [PMID: 29189932 DOI: 10.1007/s00330-017-5152-x]

35    **Chou R**, Cuevas C, Fu R, Devine B, Wasson N, Ginsburg A, Zakher B, Pappas M, Graham E, Sullivan SD. Imaging Techniques for the Diagnosis of Hepatocellular Carcinoma: A Systematic Review and Meta-analysis. *Ann Intern Med* 2015; **162**: 697-711 [PMID: 25984845 DOI: 10.7326/M14-2509]

36    **Minami Y**, Kudo M. Review of dynamic contrast-enhanced ultrasound guidance in ablation therapy for hepatocellular carcinoma. *World J Gastroenterol* 2011; **17**: 4952-4959 [PMID: 22174544 DOI: 10.3748/wjg.v17.i45.4952]

37    **Shiina S**, Teratani T, Obi S, Hamamura K, Koike Y, Omata M. Nonsurgical treatment of hepatocellular carcinoma: from percutaneous ethanol injection therapy and percutaneous microwave coagulation therapy to radiofrequency ablation. *Oncology* 2002; **62** Suppl 1: 64-68 [PMID: 11868788 DOI: 10.1159/000048278]

38    **Miyamoto N**, Hiramatsu K, Tsuchiya K, Sato Y. Contrast-enhanced sonography-guided radiofrequency ablation for the local recurrence of previously treated hepatocellular carcinoma undetected by B-mode sonography. *J Clin Ultrasound* 2010; **38**: 339-345 [PMID: 20572066 DOI: 10.1002/jcu.20720]

39    **Masuzaki R**, Shiina S, Tateishi R, Yoshida H, Goto E, Sugioka Y, Kondo Y, Goto T, Ikeda H, Omata M, Koike K. Utility of contrast-enhanced ultrasonography with Sonazoid in radiofrequency ablation for hepatocellular carcinoma. *J Gastroenterol Hepatol* 2011; **26**: 759-764 [PMID: 21054516 DOI: 10.1111/j.1440-1746.2010.06559.x]

40    **McHugh ML**. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012; **22**: 276-282 [PMID: 23092060 DOI: 10.11613/bm.2012.031]

41    **Cheung OK**, Cheng AS. Gender Differences in Adipocyte Metabolism and Liver Cancer Progression. *Front Genet* 2016; **7**: 168 [PMID: 27703473 DOI: 10.3389/fgene.2016.00168]

42    **Chalasani N**, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, Sanyal AJ; American Gastroenterological Association; American Association for the Study of Liver Diseases; American College of Gastroenterologyh. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. *Gastroenterology* 2012; **142**: 1592-1609 [PMID: 22656328 DOI: 10.1053/j.gastro.2012.04.001]

43    **Wilson SR**, Lyshchik A, Piscaglia F, Cosgrove D, Jang HJ, Sirlin C, Dietrich CF, Kim TK, Willmann JK, Kono Y. CEUS LI-RADS: algorithm, implementation, and key differences from CT/MRI. *Abdom Radiol (NY)* 2018; **43**: 127-142 [PMID: 28819825 DOI: 10.1007/s00261-017-1250-0]

44    **Bhayana D**, Kim TK, Jang HJ, Burns PN, Wilson SR. Hypervascular liver masses on contrast-enhanced ultrasound: the importance of washout. *AJR Am J Roentgenol* 2010; **194**: 977-983 [PMID: 20308500 DOI: 10.2214/AJR.09.3375]

45    **Jang HJ**, Kim TK, Burns PN, Wilson SR. Enhancement patterns of hepatocellular carcinoma at contrast-enhanced US: comparison with histologic differentiation. *Radiology* 2007; **244**: 898-906 [PMID: 17709836 DOI: 10.1148/radiol.2443061520]

46    **Bartlett JW**, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol* 2008; **31**: 466-475 [PMID: 18306169 DOI: 10.1002/uog.5256]