This appendix presents completed QUADAS-2 assessments for all included studies.

Blondin 2011[48]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

Retrospective selection of patients with liver cirrhosis from a database (radiological information system) of patients who underwent CEMRI and CEUS

Was a consecutive or random sample of patients enrolled?	No
Was a case–control design avoided?	No
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Patients with liver cirrhosis and FLLs diagnosed with CEUS and CEMRI

Is there concern that the included patients do not match the review question?	**Concern: high**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

CEUS and CEMRI results were interpreted by two experts who were blinded (no more details given on blinding); index and comparator tests were conducted with maximum 4 weeks in between

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

CEUS and CEMRI results were interpreted by two experts who were blinded (no more details given on blinding); index and comparator test were conducted with maximum 4 weeks in between

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Histology was carried out for all FLLs, before imaging results were analysed

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Yes
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: low**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

All patients received each test.

Describe the time interval and any interventions carried out between index, comparator(s) and reference standard:

Time between index and comparator tests and reference standard was not reported. Time between index and comparator tests was maximum 4 weeks.

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: low**

Catala 2007[52]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

Patients ≥18 years with FLLs detected on standard US. A total of 213 patients assessed for inclusion, with 77 enrolled

Excluded if pregnant or nursing, if more than 1 month between CEUS and spiral computed tomography (unclear if these patients may be systematically different) and if positive lesions not confirmed by pathology

Was a consecutive or random sample of patients enrolled?	No
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Adult patients with FLLs detected at standard US. Not clear if standard US was diagnostic

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Independently, by experienced radiologists, who were unaware of the diagnosis and the results of other imaging tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Independently, by experienced radiologists, who were unaware of the diagnosis and the results of other imaging tests

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

All index test-positive FLLs were confirmed pathologically following biopsy or surgery. Index test-negative lesions were confirmed by MRI and a minimum of 12 months' follow-up

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

A total of 213 patients were originally recruited; 77 were included in the analysis. Patients were excluded if there was >1 month between CEUS and SCT, or if positive lesions were not confirmed by pathology

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Time between index test and comparator was ≤1 month; time between tests and pathology reference standard not specified; follow-up period appropriate

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	No
Did patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: high**

Clevert 2009[51]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

Prospective cohort of 100 consecutive patients with suspected hepatic tumours

Exclusion criteria were lesion >5 cm, more than five lesions, strong allergic reactions, liver or kidney disease with confirmed elevation of laboratory parameters, acute heart failure, acute myocardial infarction, subcutaneous emphysema, meteorism, tachypnea and aerobilia

The majority of test-positive patients were diagnosed with liver metastases, but previous investigations and diagnostic status with respect to primary tumours were unclear

Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Risk: low**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Previous investigations and diagnostic status with respect to primary tumours were unclear

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

CEUS interpreters blinded. Reference standard performed after both tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Interpreted blind to index test; reference standard performed after both tests

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

No details of blinding or interpretation reported

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

100 patients, with one lesion per patient. Positive tests were confirmed histologically and negative tests by imaging follow-up over 2 years. A total of 21 patients were excluded from the CT analysis (eight did not undergo CT and 13 had non-diagnostic CT results)

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Imaging tests were performed on the same day. Follow-up was >6 months but time between imaging and histological confirmation was not reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: high**

Dai 2008[43]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: 498 consecutive patients with cirrhosis; study included 72 patients with 103 indeterminate liver nodules detected on surveillance US Nine patients had been previously treated for HCC	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Risk: low**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Adult patients with cirrhosis and indeterminate FLLs detected at surveillance US	
Is there concern that the included patients do not match the review question?	**Concern: low**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: In consensus, by two experienced sonologists, who were unaware of the diagnosis and the results of other imaging tests	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: In consensus, by two experienced radiologists, who were unaware of the diagnosis and the results of other imaging tests	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

All patients underwent biopsy (malignant and benign FLLs) within 15 days after CEUS; a negative biopsy was followed up for at least 6 months, including US, CT and testing for AFP

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

498 patients with cirrhosis; 72 with indeterminate liver nodules on US were included in the study

Describe the time interval and any interventions between index, comparator(s) and reference standard:

All patients underwent biopsy within 15 days after CEUS; all patients underwent CECT within 15 days before or after CEUS

Was there an appropriate interval between index test and reference standard?	Yes
Was there an appropriate interval between comparator test and reference standard?	Yes
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: low**

Feng 2007[57]

Chinese-language paper.

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Prospective cohort of 23 patients with 26 malignant lesions (23 HCC and three metastases) undergoing cryosurgery	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: unclear**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Patients being assessed for treatment response	
Is there concern that the included patients do not match the review question?	**Concern: low**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details of interpretation reported. Reference standard followed imaging	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details of interpretation reported. Reference standard followed imaging	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:	
Unclear if those making the diagnosis were aware of imaging results	
Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):	
All patients underwent imaging tests within 2 weeks of each other and within 1 week to 3 months after treatment. All diagnoses were confirmed by histopathology	
Describe the time interval and any interventions between index, comparator(s) and reference standard:	
Time between imaging tests and reference standard was not reported	
Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: unclear**

Flor 2010[39]

Abstract only.

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Prospective cohort of 18 patients with known primary cancer and indeterminate liver lesions (< 1.5 cm) detected at MDCT	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: unclear**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Patients with known primary cancer and indeterminate liver lesions (< 1.5 cm) detected at MDCT	
Is there concern that the included patients do not match the review question?	**Concern: low**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details reported	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Not applicable	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	
Were the comparator test results interpreted without knowledge of the index test?	
If a threshold was used, was it prespecified?	
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: not applicable**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Biopsy or 3- to 6-month follow-up was used as the reference standard. No further details were reported

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Not applicable
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

All 18 patients appear to have received a reference standard. Numbers confirmed by biopsy/follow-up were not reported

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Time between index test and biopsy was not reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Not applicable
Was there an appropriate interval between index test and comparator test?	Not applicable
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Unclear
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: unclear**

Forner 2008[44]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Prospective cohort of 89 patients with Child–Pugh A/B cirrhosis and a new solid (5–20 mm) nodule detected on surveillance US No patients had history of HCC	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Risk: unclear**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Adult patients with cirrhosis and new FLLs detected at surveillance US. Diagnostic status following conventional US was not specified	
Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Interpreted by two experienced radiologists. Article states 'blindly', but nature of blinding is unspecified	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Interpreted by two experienced radiologists who were unaware of biopsy results	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

All index test-positive FLLs were confirmed pathologically following biopsy or surgery. Index test-negative lesions were confirmed by MRI and a minimum of 12 months' follow-up

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

89 patients all received index test, comparator and a reference standard

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Times between index test, comparator and reference standard were not reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: unclear**

Georgio 2007[45]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Prospective cohort of 73 consecutive patients with cirrhosis and a single nodule (≤30 mm) detected on US Patients with a history of heart disease excluded (because of a rare side effect of SonoVue)	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Risk: low**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Adult patients with cirrhosis and single FLL detected at US. Diagnostic status following conventional US was not specified	
Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Interpreted by one operator with 20 years' experience. Index test performed before comparator and reference standard	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Interpreted by one radiologist who was unaware of index test results. Comparator test performed before reference standard	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Biopsy performed in all patients the day after both imaging studies were complete. No details of blinding were reported

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

73 patients all received the index test, comparator and a reference standard; same reference standard was used in all patients

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Comparator was performed the day after the index test and the reference standard the day after that

Was there an appropriate interval between index test and reference standard?	Yes
Was there an appropriate interval between comparator test and reference standard?	Yes
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: low**

Gierbliński 2008[53]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

Prospective cohort of 100 patients with incidentally detected liver lesions and inconclusive unenhanced US and/or CT. Patients with current or previous malignancy, with lesions with features of haemangioma or who were unable to undergo biopsy were excluded

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Adult patients with incidentally detected FLLs in whom US and/or CT could not rule out malignancy. Not clear how many patients had CT

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Interpreted by two experienced gastroenterologists; blinding unspecified

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Not applicable

Were the comparator test results interpreted without knowledge of the results of the reference standard?	
Were the comparator test results interpreted without knowledge of the index test?	
If a threshold was used, was it prespecified?	
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: not applicable**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

All FLLs were confirmed pathologically following biopsy. Biopsy-negative lesions were confirmed by clinical and imaging follow-up

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Not applicable
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: low**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

89 patients all received index test, comparator and a reference standard

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Time between index test and reference standard was not reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Not applicable
Was there an appropriate interval between index test and comparator test?	Not applicable
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: unclear**

Jonas 2011[50]

Abstract only.

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

Prospective cohort of 20 consecutive patients with CRC liver metastases who could be rendered tumour free by a single-stage surgical intervention and who underwent complete preoperative workup

Note: study states aim as determining the sensitivity and specificity for detection of metastases, but all included patients appear to have metastases

Patients with concomitant resectable extrahepatic disease and previous hepatobiliary surgery, other than cholecystectomy, were excluded

Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Adult patients with CRC liver metastases. Initial diagnostic status unclear (see previous note)

Is there concern that the included patients do not match the review question?	**Concern: high**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

No details of blinding or interpretation reported

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:	
No details of blinding or interpretation reported	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:	
No details of blinding or interpretation reported	
Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):	
20 patients, 48 lesions, by lesion analysis. All patients appear to have received index test and both comparators. All resected, imaging-positive lesions were confirmed histologically and all patients had at least 36 months' imaging follow-up. Per 2×2 patient data were not reported/derivable and the number of lesions per patient was unclear	
Describe the time interval and any interventions between index, comparator(s) and reference standard:	
No details of the timing of the tests were reported	
Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Could the patient flow have introduced bias?	**Risk: unclear**

Leoni 2010[42]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:	
Prospective consecutive cohort of cirrhotic patients with one to three hepatic nodules between 1 and 3 cm on US surveillance. Included both newly detected and recurrence of nodules	
Patients in whom the nodules to be included in the study had been pretreated, those with contraindications to imaging and those with neoplastic portal thrombosis or extrahepatic metastases were excluded	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):	
Diagnostic status following unenhanced imaging unclear	
Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:	
Unclear if those interpreting CEUS had knowledge of other imaging test results	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:	
Interpreted blind to other imaging test results, and biopsy/follow-up occurred after imaging	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Non-invasive positive diagnoses were interpreted without knowledge of other imaging studies. No details of interpretation of biopsy and follow-up were reported

Is the reference standard likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: high**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

Prospective cohort of 60 cirrhotic patients with at least one to three hepatic nodules (1–3 cm) on US (75 nodules). Positive nodules confirmed by two concordant imaging test results, FNB or follow-up at 3-month intervals. Negative nodules confirmed by FNB or follow-up at 3-month intervals. Seven nodules (<10%) were not examined by SPIO-MRI and were excluded from the analysis of test performance

Describe the time interval and any interventions between index, comparator(s) and reference standard:

No details of the timing of examinations were reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: high**

Li 2007[54]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Prospective cohort of 109 patients examined with unenhanced US and unenhanced CT. Exclusions not specified	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: unclear**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Diagnostic status following baseline imaging unclear	
Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Interpreted blind to comparator; reference standard performed after both tests	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Interpreted blind to index test; reference standard performed after both tests	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:	
No details of blinding or interpretation reported	
Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):	
109 patients, one lesion per patient. All patients appear to have received the index test, comparator and reference standard. Reference standard was histology in all patients. Seven lesions could not be visualised by CECT and three could not be visualised by CEUS. For our analysis, non-visualised lesions were classified as negative (false-negative or true-negative according to final diagnosis)	
Describe the time interval and any interventions between index, comparator(s) and reference standard:	
Reference standard was performed within 2 weeks of the index test and comparator	
Was there an appropriate interval between index test and reference standard?	Yes
Was there an appropriate interval between comparator test and reference standard?	Yes
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: low**

Lüttich 2006[40]

Abstract only.

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Cohort of 15 patients with HCC lesions undergoing RFA treatment	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: unclear**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Patients being assessed for response to treatment	
Is there concern that the included patients do not match the review question?	**Concern: low**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details of interpretation reported. Reference standard followed CEUS	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details of interpretation reported. Reference standard followed CEUS	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted: Unclear if those making the diagnosis were aware of imaging results	
Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s): All patients underwent both imaging tests within 4 weeks of treatment. All patients had results confirmed by biopsy Describe the time interval and any interventions between index, comparator(s) and reference standard: Time between tests and reference standard was not reported	
Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: unclear**

Mainenti 2010[49]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:
Prospective cohort of 34 consecutive patients with histologically proven CRC who were scheduled for surgery
Patients who refused to participate and those who had contraindications to one of the examinations were excluded

Was a consecutive or random sample of patients enrolled?	Yes
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	**Risk: low**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):
Diagnostic status following unenhanced imaging unclear

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:
Interpreted blind to comparator; reference standard performed after both tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:
Interpreted blind to index test; reference standard performed after both tests

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

No details of blinding or interpretation reported

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of Bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

34 patients, 57 lesions; both per-lesion and per-patient data reported. Positive tests were confirmed by biopsy or resection. All patients were followed up for 6 and 12 months, either to confirm negative tests or to detect newly developed metastases

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Surgery was performed within 10 days of imaging and imaging tests were performed over a 4- to 8-day period

Was there an appropriate interval between index test and reference standard?	Yes
Was there an appropriate interval between comparator test and reference standard?	Yes
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: low**

Quaia 2009[46]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:
Prospective cohort of cirrhotic patients with at least one hepatic nodule on US surveillance
Only those nodules ≤3 cm that underwent biopsy after CT were included
Nodules with peripheral enhancement at CECT were excluded because of a high probability of haemangioma diagnosis

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):
Diagnostic status following unenhanced imaging unclear

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:
Interpreted blind to comparator, reference standard and clinical details

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: low**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:
Interpreted blind to index test, reference standard and clinical details

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: low**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

No details of blinding or interpretation reported

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

Prospective cohort of 180 cirrhotic patients with at least one hepatic nodule on US surveillance (195 nodules)

74 nodules were excluded because of a lack of histological diagnosis ($n = 60$), technical inadequacy of CT ($n = 10$) or inadequacy of CEUS examination ($n = 4$); 106 patients with 121 nodules finally included

Reference standard biopsy in all nodules

Describe the time interval and any interventions between index, comparator(s) and reference standard:

CT was performed 2–30 days after CEUS. Biopsy was within 15 days of CT

Was there an appropriate interval between index test and reference standard?	Yes
Was there an appropriate interval between comparator test and reference standard?	Yes
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	No
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: high**

Sangiovanni 2010[47,61]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

Prospective cohort of cirrhotic patients with at least one hepatic nodule on US surveillance

Only 1- to 2-cm nodules were included in the analysis

Patients with a pre-existing liver nodule, poor liver function indicating transplantation regardless of HCC, or no defined nodule were excluded

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Diagnostic status following unenhanced imaging unclear

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Interpreted blind to reference standard

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

Interpreted blind to reference standard

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:	
Reference standard interpreted without knowledge of clinical or imaging results	
Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Yes
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: low**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):	
Prospective cohort of 64 cirrhotic patients with at least one hepatic nodule (67 nodules). All nodules confirmed by biopsy	
Describe the time interval and any interventions between index, comparator(s) and reference standard:	
Biopsy was performed within 2 months of nodule detection	
Was there an appropriate interval between index test and reference standard?	Yes
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: high**

Seitz 2009[55]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

The study used a cohort of 267 out of 1349 patients of a prospective study of consecutive patients with newly detected FLLs identified on US. The 267 patients were divided into subgroups A and B. Subgroup A had mainly benign diagnoses and subgroup B had mainly malignant diagnosis; 2×2 data with an appropriate reference standard were extractable only for subgroup B

Patients with specific liver lesions diagnosed by typical US echomorphology, such as cysts or haemangiomas, in a non-steatotic liver without clinical signs and symptoms, as well as malignant tumours with infiltration into hepatic vessels, were excluded; patients who were critically ill or who suffered from pulmonary hypertension or unstable angina, as well as pregnant and nursing women, were also excluded.

Was a consecutive or random sample of patients enrolled?	No
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Patients with newly detected FLLs on US; primary diseases not specified

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

The definitive CEUS diagnosis was made at the time of the US examination by the physician performing CEUS; US carried out by the local investigators; US investigator not blinded to the results of the preceding CT in eight cases

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Yes
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

No details of blinding reported. Reporting radiologists had access to the patients' clinical information

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Subgroup B: diagnosis was based on US-guided FNB; no definitive diagnosis could be obtained in four patients

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

Four patients with inconclusive histology were excluded from the analyses (<10% of patients)

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Times between index and comparator tests and reference standard were not reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: unclear**

Seitz 2010[56]

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection:

The study used a cohort of 269 out of 1349 patients of a prospective study of consecutive patients with newly detected FLLs identified on US. The 269 patients were divided into subgroups A and B. Subgroup A had mainly benign diagnoses and subgroup B had mainly malignant diagnosis; 2×2 data with an appropriate reference standard were extractable only for subgroup B

Patients with specific liver lesions diagnosed by typical US echomorphology, such as cysts or haemangiomas, in a non-steatotic liver without clinical signs and symptoms, as well as malignant tumours with infiltration into hepatic vessels, were excluded

Was a consecutive or random sample of patients enrolled?	No
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting):

Patients with newly detected FLLs on US; primary diseases not specified

Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

The definitive CEUS diagnosis was made at the time of the US examination by the physician performing CEUS; US carried out by the local investigators

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted:

No details of blinding reported. Reporting radiologists had access to the patients' clinical information

Were the comparator test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

All index test-positive and -negative FLLs were confirmed pathologically following biopsy in subgroup B

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

Two patients with inconclusive histology were excluded from the analyses (<10% of patients)

Describe the time interval and any interventions between index, comparator(s) and reference standard:

Times between index and comparator tests and reference standard were not reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: unclear**

Solbiati 2006[41]

Abstract only.

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Retrospective cohort of patients with incidentally detected FLLs on unenhanced US	
Was a consecutive or random sample of patients enrolled?	No
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: high**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Diagnostic status following unenhanced imaging unclear	
Is there concern that the included patients do not match the review question?	**Concern: unclear**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Unclear if those interpreting CEUS had knowledge of other imaging test results. Biopsy performed after imaging	
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: Unclear if those interpreting CECT had knowledge of other imaging test results. Biopsy performed after imaging	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Unclear
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Unclear
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Reference standard was a combination of CEUS and CT in most cases. No details of interpretation of biopsy and follow-up were reported

Is the reference standard likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index test?	No
Were the reference standard results interpreted without knowledge of the results of the comparator test?	No
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: high**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

Retrospective cohort of 694 lesions in 686 patients. Reference standard was concordant imaging test results in most ($n = 656$) lesions and FNB in case of discordance ($n = 38$). One lesion was missing from the analysis

Describe the time interval and any interventions between index, comparator(s) and reference standard:

No details of the timing of examinations were reported

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	**Risk: high**

Zhou 2007[58]

Chinese-language paper.

Domain 1: patient selection

A. Risk of bias

Describe methods of patient selection: Retrospective analysis of data from 56 patients with 64 HCC lesions undergoing non-surgical treatment	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case–control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	**Risk: unclear**

B. Concerns regarding applicability

Describe included patients (previous testing, presentation, intended use of index test and setting): Patients being assessed for response to treatment	
Is there concern that the included patients do not match the review question?	**Concern: low**

Domain 2a: index test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details of interpretation reported. Reference standard followed imaging	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Were the index test results interpreted without knowledge of the comparator?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the index test have introduced bias?	**Risk: unclear**

Domain 2b: comparator test

A. Risk of bias

Describe how the index test and any comparator tests were conducted and interpreted: No details of interpretation reported. Reference standard followed imaging	
Were the comparator test results interpreted without knowledge of the results of the reference standard?	Yes
Were the comparator test results interpreted without knowledge of the index test?	Unclear
If a threshold was used, was it prespecified?	Yes
Could the conduct or interpretation of the comparator test have introduced bias?	**Risk: unclear**

Domain 3: reference standard

A. Risk of bias

Describe the reference standard and how it was conducted and interpreted:

Unclear if those making the diagnosis were aware of imaging results; 3-month follow-up may not be adequate to confirm tumour response

Is the reference standard likely to correctly classify the target condition?	Unclear
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Were the reference standard results interpreted without knowledge of the results of the comparator test?	Unclear
Could methods used to conduct or interpret the reference standard have introduced bias?	**Risk: unclear**

Domain 4: flow and timing

A. Risk of bias

Describe any patients who did not receive the index test, comparator(s) and/or reference standard or who were excluded from the 2×2 table(s):

All patients underwent both imaging tests within 1 week of treatment. Patients with a positive response on imaging were followed up for 3 months. Patients with a negative response on imaging (residual tumour detected) had diagnosis confirmed by FNB

Describe the time interval and any interventions between index, comparator(s) and reference standard:

See above. Note: 3-month follow-up may not be adequate to confirm tumour response

Was there an appropriate interval between index test and reference standard?	Unclear
Was there an appropriate interval between comparator test and reference standard?	Unclear
Was there an appropriate interval between index test and comparator test?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	**Risk: unclear**