



Cochrane
Library

Cochrane Database of Systematic Reviews

Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease (Review)

Nadarevic T, Colli A, Giljaca V, Fraquelli M, Casazza G, Manzotti C, Štimac D, Miletic D

Nadarevic T, Colli A, Giljaca V, Fraquelli M, Casazza G, Manzotti C, Štimac D, Miletic D.
Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease.
Cochrane Database of Systematic Reviews 2022, Issue 5. Art. No.: CD014798.
DOI: [10.1002/14651858.CD014798.pub2](https://doi.org/10.1002/14651858.CD014798.pub2).

www.cochranelibrary.com

Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease
(Review)

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	7
Figure 1.	9
OBJECTIVES	11
METHODS	11
RESULTS	14
Figure 2.	15
Figure 3.	16
Figure 4.	17
Figure 5.	21
Figure 6.	22
Figure 7.	23
DISCUSSION	23
AUTHORS' CONCLUSIONS	26
ACKNOWLEDGEMENTS	27
REFERENCES	28
CHARACTERISTICS OF STUDIES	40
DATA	100
Test 1. MRI	101
Test 2. Secondary outcome	101
ADDITIONAL TABLES	101
APPENDICES	104
HISTORY	109
CONTRIBUTIONS OF AUTHORS	109
DECLARATIONS OF INTEREST	109
SOURCES OF SUPPORT	110
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	110
INDEX TERMS	110

[Diagnostic Test Accuracy Review]

Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease

Tin Nadarevic¹, Agostino Colli², Vanja Giljaca³, Mirella Fraquelli⁴, Giovanni Casazza⁵, Cristina Manzotti⁴, Davor Štimac⁶, Damir Miletic⁷

¹Department of Radiology, Clinical Hospital Centre Rijeka, Rijeka, Croatia. ²Department of Transfusion Medicine and Haematology, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano, Italy. ³Department of Gastroenterology, Heart of England NHS Foundation Trust, Birmingham, UK. ⁴Gastroenterology and Endoscopy Unit, Fondazione IRCCS Ca' Granda - Ospedale Maggiore Policlinico, Milan, Italy. ⁵Dipartimento di Scienze Biomediche e Cliniche "L. Sacco", Università degli Studi di Milano, Milan, Italy. ⁶Department of Gastroenterology, Clinical Hospital Centre Rijeka, Rijeka, Croatia. ⁷Department of Radiology, Clinical Hospital Centre Rijeka, Rijeka, Croatia

Contact: Agostino Colli, colliagostino@gmail.com.

Editorial group: Cochrane Hepato-Biliary Group.

Publication status and date: New, published in Issue 5, 2022.

Citation: Nadarevic T, Colli A, Giljaca V, Fraquelli M, Casazza G, Manzotti C, Štimac D, Miletic D. Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease. *Cochrane Database of Systematic Reviews* 2022, Issue 5. Art. No.: CD014798. DOI: [10.1002/14651858.CD014798.pub2](https://doi.org/10.1002/14651858.CD014798.pub2).

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Hepatocellular carcinoma occurs mostly in people with chronic liver disease and ranks sixth in terms of global incidence of cancer, and third in terms of cancer deaths. In clinical practice, magnetic resonance imaging (MRI) is used as a second-line diagnostic imaging modality to confirm the presence of focal liver lesions suspected as hepatocellular carcinoma on prior diagnostic test such as abdominal ultrasound or alpha-fetoprotein, or both, either in surveillance programmes or in clinical settings. According to current guidelines, a single contrast-enhanced imaging study (computed tomography (CT) or MRI) showing typical hallmarks of hepatocellular carcinoma in people with cirrhosis is considered valid to diagnose hepatocellular carcinoma. The detection of hepatocellular carcinoma amenable to surgical resection could improve the prognosis. However, a significant number of hepatocellular carcinomas do not show typical hallmarks on imaging modalities, and hepatocellular carcinoma may, therefore, be missed. There is no clear evidence of the benefit of surveillance programmes in terms of overall survival: the conflicting results can be a consequence of inaccurate detection, ineffective treatment, or both. Assessing the diagnostic accuracy of MRI may clarify whether the absence of benefit could be related to underdiagnosis. Furthermore, an assessment of the accuracy of MRI in people with chronic liver disease who are not included in surveillance programmes is needed for either ruling out or diagnosing hepatocellular carcinoma.

Objectives

Primary: to assess the diagnostic accuracy of MRI for the diagnosis of hepatocellular carcinoma of any size and at any stage in adults with chronic liver disease.

Secondary: to assess the diagnostic accuracy of MRI for the diagnosis of resectable hepatocellular carcinoma in adults with chronic liver disease, and to identify potential sources of heterogeneity in the results.

Search methods

We searched the Cochrane Hepato-Biliary Group Controlled Trials Register, the Cochrane Hepato-Biliary Group Diagnostic Test of Accuracy Studies Register, the Cochrane Library, MEDLINE, Embase, and three other databases to 9 November 2021. We manually searched articles retrieved, contacted experts, handsearched abstract books from meetings held during the last 10 years, and searched for literature in OpenGrey (9 November 2021). Further information was requested by e-mails, but no additional information was provided. No data was obtained through correspondence with investigators. We applied no language or document-type restrictions.

Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease (Review)

1

Copyright © 2022 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Selection criteria

Studies assessing the diagnostic accuracy of MRI for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease, with cross-sectional designs, using one of the acceptable reference standards, such as pathology of the explanted liver and histology of resected or biopsied focal liver lesion with at least a six-month follow-up.

Data collection and analysis

At least two review authors independently screened studies, extracted data, and assessed the risk of bias and applicability concerns, using the QUADAS-2 checklist. We presented the results of sensitivity and specificity, using paired forest plots, and we tabulated the results. We used a hierarchical meta-analysis model where appropriate. We presented uncertainty of the accuracy estimates using 95% confidence intervals (CIs). We double-checked all data extractions and analyses.

Main results

We included 34 studies, with 4841 participants. We judged all studies to be at high risk of bias in at least one domain because most studies used different reference standards, often inappropriate to exclude the presence of the target condition, and the time interval between the index test and the reference standard was rarely defined. Regarding applicability, we judged 15% (5/34) of studies to be at low concern and 85% (29/34) of studies to be at high concern mostly owing to characteristics of the participants, most of whom were on waiting lists for orthotopic liver transplantation, and due to pathology of the explanted liver being the only reference standard.

MRI for hepatocellular carcinoma of any size and stage: sensitivity 84.4% (95% CI 80.1% to 87.9%) and specificity 93.8% (95% CI 90.1% to 96.1%) (34 studies, 4841 participants; low-certainty evidence).

MRI for resectable hepatocellular carcinoma: sensitivity 84.3% (95% CI 77.6% to 89.3%) and specificity 92.9% (95% CI 88.3% to 95.9%) (16 studies, 2150 participants; low-certainty evidence).

The observed heterogeneity in the results remains mostly unexplained. The sensitivity analyses, which included only studies with clearly prespecified positivity criteria and only studies in which the reference standard results were interpreted without knowledge of the results of the index test, showed no variation in the results.

Authors' conclusions

We found that using MRI as a second-line imaging modality to diagnose hepatocellular carcinoma of any size and stage, 16% of people with hepatocellular carcinoma would be missed, and 6% of people without hepatocellular carcinoma would be unnecessarily treated. For resectable hepatocellular carcinoma, we found that 16% of people with resectable hepatocellular carcinoma would improperly not be resected, while 7% of people without hepatocellular carcinoma would undergo inappropriate surgery. The uncertainty resulting from the high risk of bias in the included studies and concerns regarding their applicability limit our ability to confidently draw conclusions based on our results.

PLAIN LANGUAGE SUMMARY

How accurate are magnetic resonance imaging (MRI) scans for detecting liver cancer?

Key messages

In people with chronic liver disease, magnetic resonance imaging (MRI: cross-sectional scans inside the body) probably misses liver cancer in 16% of people, who would not receive timely or appropriate treatment, and incorrectly finds liver cancer in 6% of people, who would receive unnecessary treatment.

MRI probably misses liver cancer in 16% of people with liver cancer who could have surgery to remove part of their liver, and incorrectly finds liver cancer in 7% of people who undergo inappropriate surgery.

The studies were at high risk of bias and too different from each other to allow us to draw firm conclusions based on the evidence.

Why is it important to diagnose liver cancer accurately?

Liver cancer, or 'hepatocellular carcinoma', occurs mostly in people with chronic liver disease, regardless of the cause. It is the sixth most common cancer in the world and the third most common cause of deaths due to cancer. It is difficult to diagnose because early symptoms are similar to those of liver disease. People with blood test or ultrasound results that suggest liver cancer may go on to have further tests, such as scans that produce images of the liver, or biopsy where a small piece of the liver is removed and examined. If liver cancer is detected early, people may be treated with surgery to remove part of the liver (called a liver resection) or with a liver transplant. If the liver cancer is more advanced, they may need chemotherapy. If liver cancer is missed at the diagnostic test, people will not receive appropriate treatment. However, incorrectly diagnosing liver cancer when it is not present means that people may undergo unnecessary testing or treatment.

What is magnetic resonance imaging (MRI) and how might it diagnose liver cancer?

MRI produces images that show a cross-section or 'slice' of the bones, blood vessels, and tissues inside the body. The images are a series of signal intensities that are directed and combined by a computer. MRI scans can detect the presence of abnormalities in the liver that might be cancer. Current guidelines recommend using either MRI or another type of imaging, computed tomography, or a combination to confirm the presence of liver cancer in people who might have liver cancer.

What did we want to find out?

We wanted to find out if MRI is accurate enough to diagnose liver cancer in adults with chronic liver disease. We were interested first, in liver cancers of any size and stage and second, in liver cancers that were suitable for resection.

What did we do?

We searched for studies that assessed the accuracy of MRI scans compared to the best available tests to confirm liver cancer in adults with chronic liver disease. The best available tests are examination of the liver, or part of the liver under a microscope.

What did we find?

We found 34 studies assessing 4841 people.

Around 560 of 1000 (56%) adults with chronic liver disease have confirmed liver cancer. Of these 1000 people, MRI may:

- correctly detect liver cancer in 473 people;
- miss liver cancer in 87 people;
- incorrectly detect liver cancer in 27 cancer-free people;
- correctly detect no liver cancer in 413 people.

Based on the studies, around 560 of 1000 (56%) adults with chronic liver disease have confirmed resectable liver cancer. Of these 1000 people, MRI may:

- correctly detect resectable liver cancer in 472 people;
- miss resectable liver cancer in 88 people;
- incorrectly detect resectable liver cancer in 31 people;
- correctly detect no resectable liver cancer in 409 people.

What are the limitations of the evidence?

Our confidence in the evidence is limited because the studies used different methods to select study participants and used different definitions for the presence of liver disease. This means MRI scans could be more or less accurate than suggested by our analyses of the evidence.

How up to date is this evidence?

The evidence is up to date to 9 November 2021.

SUMMARY OF FINDINGS

Summary of findings 1. Diagnostic accuracy of magnetic resonance imaging for the diagnosis of hepatocellular carcinoma

Review question: what is the diagnostic accuracy of MRI for the diagnosis of HCC in people with chronic liver disease?

Population: adults with chronic liver disease

Setting: clinical setting (secondary or tertiary care setting) or surveillance programmes

Study design: cross-sectional studies

Index test: MRI

Target condition: HCC of any size, any stage

Reference standards

- Pathology of the explanted liver in case of transplantation
- Histology of resected focal liver lesion(s), or the histology of biopsied focal liver lesion(s) with a follow-up of ≥ 6 months to exclude the presence of focal lesions not detected by the index test

Limitations in the evidence: risk of bias and applicability concerns

- Participant selection: high/unclear risk of bias 26 studies (76%); high concern for applicability 20 studies (59%)
- Index tests: high/unclear risk of bias 4 studies (12%); high concern for applicability 1 study (3%)
- Reference standard: high/unclear risk of bias 28 studies (82%); high concern for applicability 16 studies (47%)
- Flow and timing: high/unclear risk of bias 31 studies (91%)
- Overall: high risk of bias all included studies; high concern for applicability 29 studies (85%)

Findings

Index test	Number of studies (participants)	Sensitivity (95% CI)	Specificity (95% CI)	Implications in a hypothetical cohort of 1000 people					
				Prevalence ^a %	True positives will receive appropriate treatment (surgery or local ablative therapy or systemic chemotherapy)	False negatives will be misdiagnosed and not receive appropriate treatment	True negatives will not undergo inappropriate treatment or unnecessary further testing	False positives will undergo inappropriate treatment	Certainty of the evidence
MRI	34 (4841)	84.4% (80.1% to 87.9%)	93.8% (90.1% to 96.1%)	36	304	56	600	40	Low ^b
				56	473	87	413	27	

66 557 103 319 21

CI: confidence interval; **HCC:** hepatocellular carcinoma; **MRI:** magnetic resonance imaging.

^aWe chose for exemplification three values of hepatocellular carcinoma prevalence: 36% for a population with low clinical suspicion, 56% as a median derived from our study analysis, and 66% for population with high clinical suspicion (assessment of nodules detected by ultrasound).

^bDowngraded two levels for risk of bias and indirectness.

Summary of findings 2. Diagnostic accuracy of magnetic resonance imaging for the diagnosis of resectable hepatocellular carcinoma

Review question: what is the diagnostic accuracy of MRI for the diagnosis of resectable HCC in people with chronic liver disease?

Population: adults with chronic liver disease

Setting: clinical setting (secondary or tertiary care setting) or surveillance programmes

Study design: cross-sectional studies

Index test: MRI

Target condition: resectable HCC

Reference standards

- Pathology of the explanted liver in case of transplantation
- Histology of resected focal liver lesion(s), or the histology of resected or biopsied focal liver lesion(s) with a follow-up of ≥ 6 months to exclude the presence of focal lesions not detected by the index test

Limitations in the evidence: risk of bias and applicability concerns (total 16 studies which had all participants with resectable HCC)

- Participant selection: high/unclear risk of bias 10 studies (63%); high concern for applicability 11 studies (69%)
- Index tests: high/unclear risk of bias 0 studies (0%); high concern for applicability 0 studies (0%)
- Reference standard: high/unclear risk of bias 10 studies (63%); high concern for applicability 11 studies (69%)
- Flow and timing: high/unclear risk of bias 14 studies (88%)
- Overall: high risk of bias; all included studies; high concern for applicability 11 studies (69%)

Findings

Index test	Number of studies (participants)	Sensitivity (95% CI)	Specificity (95% CI)	Implications in a hypothetical cohort of 1000 people
------------	----------------------------------	----------------------	----------------------	--

				Prevalence^a%	True positives will receive appropriate treatment (surgical resection)	False negatives will be misdiagnosed and not undergo surgical resection	True negatives will not undergo inappropriate further testing or surgical resection	False positives will undergo inappropriate further testing or surgical resection	Certainty of the evidence
MRI	16	84.3%	92.9%	36	303	57	595	45	Low ^b
	(2150)	(77.6% to 89.3%)	(88.3% to 95.9%)	56	472	88	409	31	
				66	556	104	316	24	

CI: confidence interval; **HCC:** hepatocellular carcinoma; **MRI:** magnetic resonance imaging.

^aWe chose for exemplification three values of hepatocellular carcinoma prevalence: 36% for a population with low clinical suspicion, 56% as a median derived from our study analysis, and 66% for population with high clinical suspicion (assessment of nodules detected by ultrasound).

^bDowngraded two levels for risk of bias and indirectness.

BACKGROUND

Hepatocellular carcinoma (HCC) is the most common primary liver tumour, usually developing in the setting of chronic liver disease. It represents the third most common cause of death from cancer worldwide, with high rates in East and Southeast Asia, several areas of Africa, and Southern Europe (Bertuccio 2017). From the early 2010s, HCC was one of the few cancers that showed increasing incidence and mortality trends in several areas of the world including Europe, and North and Latin America (Bosetti 2013; Hashim 2016; Ryerson 2016). Mortality rates, even with a recently downward reported trend, are reported to remain two to five times higher in Japan, Hong Kong, and Korea than in most European countries, and North and South America (Bertuccio 2017). Most common risk factors include liver cirrhosis, severe liver fibrosis, chronic infections with hepatitis B and C, heavy alcohol intake, tobacco use, diabetes, metabolic syndrome, aflatoxins (poisonous carcinogens produced by *Aspergillus flavus* and *Aspergillus parasiticus*, which grow in soil, decaying vegetation, hay, and grains), non-alcoholic fatty liver disease, and being overweight (Yang 2011; Bosetti 2014; Stanaway 2016; Bertuccio 2017). People with HCC but without known risk factors have also been reported (Bralet 2000; Young 2012). HCC is rare among adolescents, with an incidence of 0.3 to 0.45 occurrences per million per year and accounts for less than 1% of all malignant neoplasms among children aged less than 20 years (Mann 1990). The reported HCCs were associated with hepatitis B infection or with inherited metabolic disorders, specifically hereditary tyrosinaemia, alpha-1-antitrypsin deficiency, and glycogen storage disease type 1. Only approximately 30% of HCC in children are associated with cirrhosis, and the carcinogenesis and the clinical course are considered peculiar (Ni 2004; Omata 2017; Mogul 2018).

Clinically, HCC is frequently diagnosed in the late stages of liver disease because of the absence of specific symptoms, other than those related to chronic liver disease. Only less than 20% of patients are eligible for curative treatment – such as liver resection, transplantation, or ablation – due to advanced tumour stage, liver dysfunction, or shortage of liver donors (Davila 2012). Furthermore, curative treatment options are unfeasible in most people due to severe clinical deterioration at the moment of diagnosis, or due to the inaccuracy of the preoperative clinical evaluation and staging procedure.

Despite the poor initial prognosis (the mortality-to-incidence overall ratio has been reported as 0.95 (Ferlay 2015)), a five-year survival of more than 50% can be achieved if HCC is detected at an early stage and relevant surgery conducted (Forner 2018a). According to the Barcelona Clinic Liver Cancer (BCLC) staging system, only people with early-stage HCC are eligible for curative treatment (Llovet 1999). Therefore, accurate and early diagnosis of HCC is of high importance.

Prior to advancements in medical imaging, biopsy and cytological examination of the liver specimen were used to make a definitive diagnosis of HCC (Tao 1984). With the development of advanced imaging techniques, HCC has become unique among tumours in that its characteristics can be accurately detected using imaging, thus reducing the need for invasive liver biopsy (Forner 2008; Sangiovanni 2010; Manini 2014). Currently, biopsy is not preferred for the diagnosis of HCC due to concerns regarding tumour seeding, risks of bleeding, and high rate of false-negative results (Silva

2008; Pomfret 2010). Therefore, biopsy is reserved for lesions with atypical appearance and when imaging results are equivocal (Bruix 2011).

Computed tomography (CT) and contrast-enhanced magnetic resonance imaging (MRI) have been established as the non-invasive imaging modalities for detection and evaluation of liver lesions (Lee 2012a; O'Neill 2015). In comparison with CT, MRI offers many advantages such as lack of ionising radiation, higher spatial resolution, ability to use both extracellular and hepatocellular contrast, and potentially better accuracy (Hartwig 2009; Grover 2015). Disadvantages are higher cost, longer imaging time, the need for patient co-operation, patient claustrophobia, and contraindications related to paramagnetic implanted devices (O'Neill 2015). The ability of MRI to detect HCC rests on characterising the contrast enhancement patterns in arterial, portal venous, and subsequent phases relative to the surrounding liver tissue. The differences in blood flow and extracellular volume between HCC and normal liver tissue lead to main radiological hallmarks of HCC (LI-RADS 2018).

According to the American Association for the Study of Liver Diseases (AASLD) and the European Association for the Study of the Liver (EASL) guidelines, a single contrast enhanced imaging study (CT or MRI), showing typical radiological hallmarks in people with cirrhosis, is valid to diagnose HCC (EASL 2018; Heimbach 2018). However, if a detected lesion presents with some (but not all) of the hallmarks of HCC, another imaging study or biopsy is warranted (EASL 2018; Heimbach 2018).

According to current relevant guidelines, there are some differences in recommendations for management with regards to the size of a suspected focal liver lesion. In AASLD guidelines, lesions with a diameter 1 cm or less and those with a diameter more than 1 cm without HCC hallmarks are labelled as indeterminate lesions and require follow-up (Heimbach 2018). EASL guidelines propose a diagnostic algorithm for management of suspected focal liver lesions and group lesions in two categories with a diameter of 1 cm or less, and more than 1 cm (EASL 2018). Asian Pacific Association for the Study of the Liver (APASL) diagnostic pathways focus more on lesion characteristics than on their size (Omata 2017). AASLD, EASL, and APASL guideline recommendations do not encompass children and adolescents (Omata 2017; EASL 2018; Heimbach 2018).

Previous systematic reviews and reviews have assessed the performance of MRI in detecting HCC, and they have included different studies and yielded different results (Colli 2006; Kim 2008; Xie 2011; Chen 2013; Floriani 2013; Chen 2014; Junqiang 2014; Chou 2015; Lee 2015; Li 2015a; Ye 2015; Guo 2016; Hanna 2016; Kierans 2016; Roberts 2018; Li 2019). These reviews assessed MRI either as a stand-alone test or compared MRI with CT and ultrasonography. Evaluation of risk of bias and inclusion criteria, type of studies, and reference standards were often inconsistent and questionable. Furthermore, these reviews did not put the index tests into context and did not clearly define their role. Instead, they compared all the available tests as they were used simultaneously. The aim of this systematic review and meta-analysis is to use Cochrane methodology to determine the accuracy of MRI using either extracellular or hepatocellular contrast agent for the diagnosis of HCC of any size, as well as to identify resectable HCC in adults with chronic liver disease.

Target condition being diagnosed

Hepatocellular carcinoma

HCC is the most common primary liver cancer which occurs mostly in people with chronic liver disease. The incidence of HCC increases in people with hepatitis B and C, alcohol use, and non-alcoholic fatty liver disease, and those with liver cirrhosis of various aetiologies (Bruix 2011). There is no definite threshold in the definition of lesion size, although literature tends to classify lesions with a diameter of 2 cm or less as 'small' (Hussain 2002; Choi 2014a; Park 2017).

In clinical practice and according to pertinent guidelines, multiphase CT or MRI with intravascular contrast allow for a highly accurate diagnosis of HCC without an invasive biopsy. The diagnosis of HCC is usually obtained on the basis of cross-sectional CT or MRI features, and liver histology is required only for undefined lesions (Omata 2017; EASL 2018; Heimbach 2018; LI-RADS 2018).

Several staging systems for HCC have been proposed and developed; however, there is no globally applicable staging system (Kinoshita 2015). Among different staging protocols, the BCLC staging system has a notable feature of treatment recommendations for each stage based on the best treatment options currently available (Llovet 1999; Llovet 2003; Llovet 2008). It is comprised of four elements: tumour extension, liver functional reserve, physical status, and cancer-related symptoms. According to the BCLC, only people with early-stage HCC are eligible for curative treatment such as surgical resection or percutaneous treatment. Orthotopic liver transplantation (OLT) is reserved for people with decompensated cirrhosis, and it is considered a definite curative treatment for HCC. The early experience with OLT for HCC in the 1980s included initial poor five-year survival and high recurrence rates, leading to OLT being contraindicated in HCC (Yokoyama 1990). In 1996, specific criteria were developed for selection of people with HCC for OLT, which became known as the Milan criteria (Mazzaferro 1996). These criteria have been repeatedly validated and their value is considerable (EASL 2018). With their implementation, the overall five-year survival of people after OLT exceeded 70% (Mazzaferro 2011). The criteria for patients eligible for OLT include single HCC lesion with diameter of 5 cm or less; or up to three HCC lesions, each with diameter of 3 cm or less; no vascular invasion, and no extrahepatic involvement (no metastasis) (Mazzaferro 1996; Omata 2017; EASL 2018).

Index test(s)

MRI is an advanced imaging modality that uses magnetic fields, magnetic field gradients, and radio waves to produce images of tissues and organs. Since the early 2010s, it was established as a powerful clinical tool for liver imaging offering relevant answers to specific clinical questions (Edelman 2014; Xian 2015). Magnetic field strength of 1.5 or 3.0 Tesla is currently sufficient for standard clinical practice. In the context of liver imaging and focal liver lesion characterisation, the morphology is assessed by analysing specific features on different MRI sequences.

In MRI, the use of contrast agents is frequently necessary. Most commonly used types of contrast agents are gadolinium-based

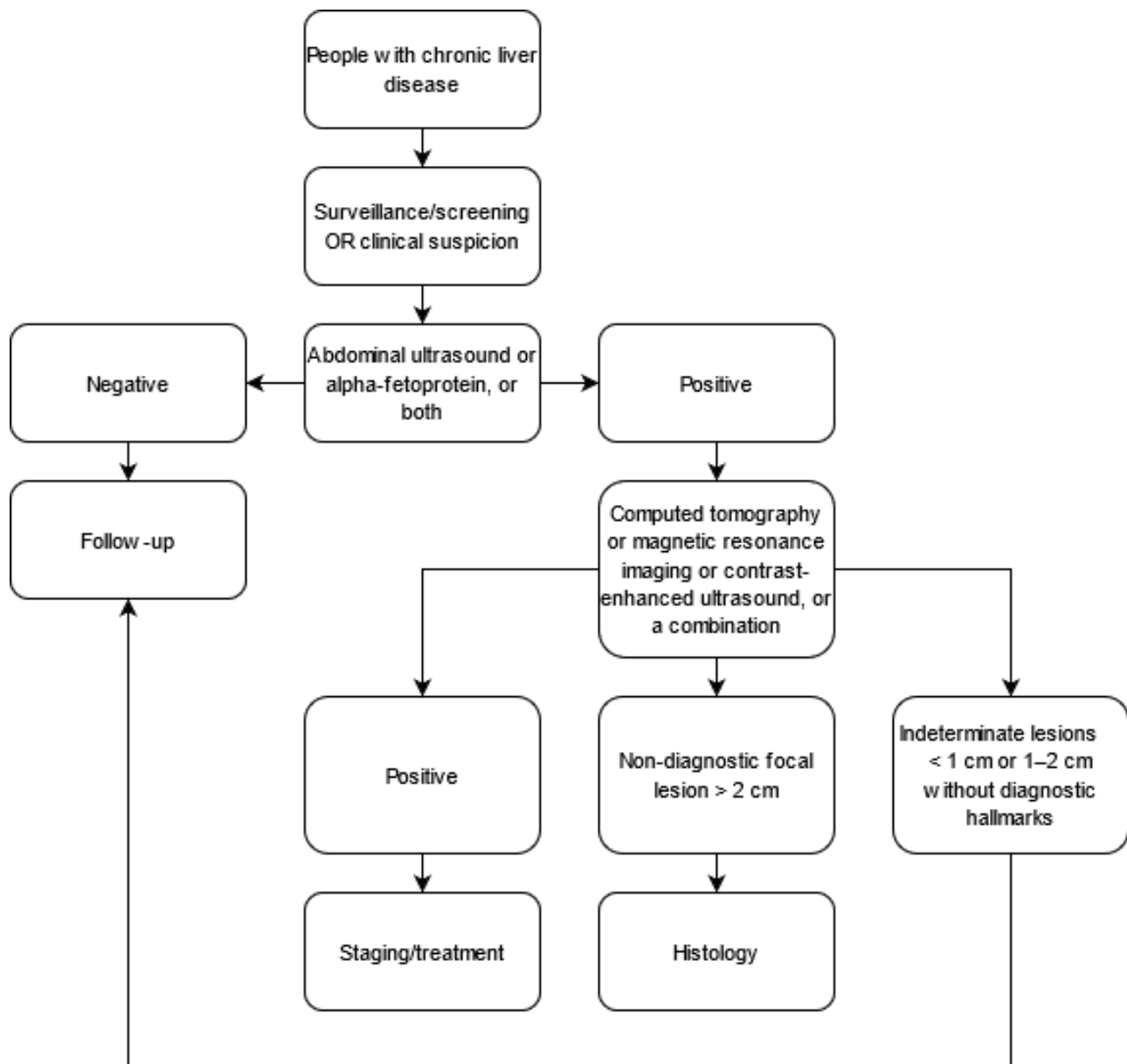
compounds, while other infrequently used contrasts include manganese-based, iron oxide, and iron platinum agents (Xiao 2016). In clinical practice, two main types of gadolinium-based contrast agents are used in liver imaging: extracellular contrast agents (ECA) and hepatobiliary contrast agents (HBA). ECAs are most widely used, as they allow the acquisition of arterial, portal venous, and delayed phases. HBAs provide similar information as ECAs, with the unique functional information on hepatocyte uptake provided in additional delayed hepatotropic phase (Choi 2014b; O'Neill 2015). When performing an MRI examination using ECA, morphological criteria for a definite HCC include non-rim-like hyperenhancement in the late arterial phase, and subsequent non-peripheral washout in the portal-venous phase or delayed phase. These criteria assess the vascular pattern of the lesions, emphasising the presence of hypervascularity of the tumour tissue. In the context of an MRI examination using HBA, the additional hepatotropic phase provides information on the parenchymal status of the lesion. Altered hepatocytes lose the ability to take up the contrast, so the lesion is hypointense relative to enhanced normal liver parenchyma (Omata 2017; EASL 2018).

In 2011, the American College of Radiology introduced a comprehensive system and criteria for diagnosing HCC, due to the need for accurate and structured non-invasive interpretation of suspected liver lesions (Elsayes 2019). According to the latest version from 2018, eight categories exist ranging from definite benign lesions to a definite HCC lesion (LI-RADS 2018). Major features for HCC include non-rim-like arterial hyperenhancement, non-peripheral washout, enhancing capsule, lesion size (cut-off values of 10 mm or 20 mm), and threshold growth (size increase of a mass by 50% or greater in six months or less) (Chernyak 2018). Depending on the presence of major and ancillary features, the lesion is characterised as Liver Imaging Reporting and Data System (LI-RADS) 1 to 5, with LR-TIV representing tumour in vein, and LR-M representing probable or definite malignancy, but not specific HCC (LI-RADS 2018). LI-RADS 4 refers to probable HCC, and LI-RADS 5 refers to definite HCC (LI-RADS 2018). In practice, lesion diameter and threshold growth are not universally accepted, contributing to heterogeneity of the use of positivity criteria in centres worldwide (EASL 2018).

Clinical pathway

Surveillance for HCC (i.e. screening performed at regular intervals) in the at-risk population (people with chronic liver disease regardless of aetiology) is carried out by abdominal ultrasound (US) for detection of liver nodules (Kanwal 2019). Once a suspected nodule has been detected, other imaging methods are considered, according to the size of the nodule and appropriate guidelines (Omata 2017; EASL 2018; Heimbach 2018). Figure 1 presents the clinical pathway showing how the tests are used. US and alpha-fetoprotein (AFP) serum measurement, alone or in combination, are used as a triage test before CT and MRI. CT and MRI play a role of add-on tests to confirm the diagnosis and to stage the disease. Only in the case of focal lesion greater than 2 cm and without diagnostic hallmarks for HCC on CT and MRI, is a biopsy recommended. Lesions with a diameter less than 1 cm and those with a diameter of 1 cm to 2 cm without HCC hallmarks are labelled as indeterminate lesions and require follow-up (Omata 2017; EASL 2018; Heimbach 2018).

Figure 1. Flow diagram of the diagnostic pathway for the diagnosis of hepatocellular carcinoma.



The diagnostic pathway after the detection of a focal liver lesion is only minimally variable among the different scientific societies as reviewed below.

American Association for the Study of Liver Disease diagnostic guidelines

According to AASLD guidelines, it is recommended that further diagnostic workup of people suspected of having HCC is performed with either multiphase CT or multiphase MRI because of their similar diagnostic performance. There is no agreement about which diagnostic test to use: multiphase CT with extracellular agents, multiphase MRI with extracellular agents, and multiphase MRI with hepatocellular contrast agent. Although it is not widely used in North America, contrast-enhanced ultrasound (CEUS) can be used to diagnose HCC. In case of indeterminate imaging findings on CT and MRI, several options are available, such as follow-

up imaging, imaging with an alternative modality or alternative contrast agent, or biopsy, but no option can be recommended over another (Heimbach 2018).

European Association for the Study of the Liver diagnostic guidelines

In cirrhosis or advanced chronic liver disease, the EASL proposed diagnostic algorithm divides suspected focal liver lesions into two categories: lesions smaller than 1 cm, and those larger than 1 cm in diameter. Lesions smaller than 1 cm are to be followed up by US every four months: if the size of the lesion did not increase, then further US follow-up is recommended, otherwise multiphase contrast-enhanced CT, multiphase contrast-enhanced MRI, or gadoteric-enhanced MRI is required. Lesions larger than 1 cm directly require to be evaluated by CT or MRI. If at least one of these imaging modalities is positive (i.e. confirms the

existence of HCC hallmarks), diagnosis of HCC is considered certain. If the results are equivocal, the use of other multiphase imaging modality is required: multiphase contrast-enhanced CT or multiphase contrast-enhanced MRI, gadoteric-enhanced MRI, or CEUS. If these studies confirm the hallmarks of HCC, the diagnosis is certain, otherwise biopsy is warranted. If biopsy appears to be unclear, repeat biopsy is to be considered or a repeat US follow-up every four months (EASL 2018).

Asian Pacific Association for the Study of the Liver diagnostic guidelines

Under the APASL guidelines, a single dynamic contrast-enhanced MRI or CT is warranted regardless of the size of suspected liver nodule. If typical hallmarks of HCC are shown (presence of arterial hyperenhancement, followed by washout in the portal venous phase or delayed phase, or both), diagnosis is confirmed. If the lesion is hypervascular but shows no washout, another contrast-enhanced MRI is needed. If the lesion proves to be hypointense, HCC diagnosis is confirmed; however, if the lesion is iso- or hyperintense, biopsy is warranted. If the lesion on the first dynamic MRI or CT is non-hypervascular, a dynamic MRI study in hepatobiliary phase is needed. If the lesion is iso- or hyperintense, surveillance by US is recommended every six months, and if the lesion is hypointense, CEUS of the liver nodule is warranted. Depending on lesion features on CEUS, biopsy, or another dynamic CT or MRI study is recommended every three to six months (Omata 2017).

Prior test(s)

US is recommended as a triage test in people at risk for developing HCC in surveillance programmes or suspected of having HCC in clinical settings (Omata 2017; EASL 2018; Heimbach 2018). When US detects a focal lesion suspected of HCC, MRI should be performed to confirm the diagnosis. Moreover, when US, CT, or CEUS detect liver nodules that are not diagnostic for HCC, MRI can be used for further diagnosis prior to histology. AFP, a glycoprotein assessed in serum as a tumour marker, can also be used prior to MRI to assess the malignancy of a focal liver lesion.

The diagnosis of the underlying chronic liver disease is based on clinical judgement derived from history, laboratory testing, physical examination, imaging, liver stiffness measurement, liver histology, or a combination of these. Due to the accuracy of non-invasive tests, liver histology is reserved to only a minority of patients with unclear diagnosis and a non-invasive diagnosis of advanced chronic liver disease is considered equivalent to a histological diagnosis of cirrhosis (de Franchis 2015).

Alternative test(s)

Contrast-enhanced ultrasound

CEUS is an advanced form of US examination in which images are acquired using intravenously injected microbubble contrast agent (Pang 2018). Dynamic CEUS images are obtained similarly to contrast-enhanced CT and MRI studies: depending on the time of image acquisition after intravenous contrast injection, the diagnostic examination differentiates arterial and portal venous phases in which sonographic hallmarks for HCC, such as arterial hyperenhancement and subsequent washout appearance, are investigated (Chung 2015; LI-RADS 2017). Unlike CT and MRI contrasts, US contrast agent is a purely intravascular agent;

therefore, it is highly accurate in detecting tumour angiogenesis (Schirner 2004). The use of US contrast agents, in particular sulphur hexafluoride microbubbles, is generally considered safe, with the reported incidence of adverse reactions of less than 0.02% (Piscaglia 2006; Tang 2017). In comparison, the adverse event proportion for iodine-based contrast agent ranges from 1% to 12%, and for gadolinium-based contrast agents ranges from 0.07% to 2.4% (Bottinor 2013; McDonald 2019; ACR 2021). In the context of liver lesion imaging, CEUS is considered more cost-effective than CT or MRI (Sirlı 2010; Westwood 2013; Smajerova 2016).

Computed tomography

Contrast-enhanced multiphase multidetector CT is a non-invasive imaging modality for detection and evaluation of liver lesions (Federle 2001). The ability to detect HCC rests on characterising the enhancement patterns in arterial, portal venous, and subsequent phases relative to the surrounding liver tissue (Navin 2019). The differences in blood flow and extracellular volume between HCC tissue and normal liver tissue lead to main radiological hallmarks such as homogeneous (non-rim-like) arterial phase hyperenhancement suggesting tumoural neo-angiogenesis and subsequent non-peripheral washout with enhancing capsule in later phases, suggesting the presence of arteriovenous communications (Hennedige 2013; Choi 2014a; LI-RADS 2018). CT is a commonly used modality for diagnosing HCC due to its short acquisition time and high spatial resolution. The obvious downfall of CT is the use of ionising radiation, which is harmful for tissues and organs on a molecular level. Although the damage is quickly repaired, occasional misrepair can induce mutations, gene fusion, and chromosomal translocations, all of which could lead to the development of cancer (Mitelman 2018). Iodine-based contrast agents may also be damaging to tissues and organs resulting in acute or late adverse reactions of different severity (Beckett 2015).

Rationale

A suspected HCC liver lesion is currently detected by liver US in people with normal or high AFP levels during surveillance programmes in people with chronic liver disease. Following US, the diagnosis of HCC is usually confirmed with CEUS, CT, or MRI. CT and MRI are also appropriate for staging of HCC and allow the choice of the most appropriate treatment. There is no clear evidence of the benefits of surveillance programmes in terms of overall survival: the conflicting results can be a consequence of an inaccurate detection, ineffective treatment, or both. Assessing the diagnostic accuracy of MRI and CT as two confirmatory tests after triage tests (US, AFP, or their combination), may clarify whether the absence of benefit in surveillance programmes might be related to underdiagnosing or understaging. Furthermore, an assessment of the accuracy of MRI for the diagnosis of HCC is also needed for ruling out, diagnosing, or supporting further testing in people with chronic liver disease who are not included in surveillance programmes.

This review represents part of a series of systematic reviews about the diagnostic accuracy of the most commonly used modalities for diagnosing HCC in adults with chronic liver disease. The first review includes assessment of the diagnostic accuracy of US and AFP levels, which are used as triage tests in surveillance (Colli 2021). The second review will focus on the diagnostic accuracy of CEUS in characterising suspected lesions as HCC as a second-line diagnostic modality (Fraquelli 2019). The third review focuses on the assessment of CT as a third-line imaging modality in assessing

focal liver lesions detected on US suspected for HCC (Nadarevic 2021a). The current fourth review analyses the accuracy of MRI for diagnosing HCC using different types of contrast media. Once these reviews are completed and published, we plan to design an overarching review comparing the accuracy of CEUS, CT, and MRI for the diagnosis and staging of HCC.

OBJECTIVES

To assess the diagnostic accuracy of magnetic resonance imaging for the diagnosis of hepatocellular carcinoma of any size and at any stage in adults with chronic liver disease.

Secondary objectives

To assess the diagnostic accuracy of MRI for the diagnosis of resectable HCC in adults with chronic liver disease.

To identify potential sources of heterogeneity, we plan to investigate the effects of the following variables: study date; inclusion of people without cirrhosis; study location (population differences); participant selection; different HCC stages; different reference standards; different liver cirrhosis aetiologies; differences in prior testing; predefinition of magnetic resonance (MR) positivity criteria; and type of contrast media (see [Investigations of heterogeneity](#)).

METHODS

Criteria for considering studies for this review

Types of studies

We included studies that, irrespective of publication status and language, evaluated the diagnostic accuracy of MRI for the diagnosis of HCC in adults with chronic liver disease. These studies should have used one of the acceptable reference standards (see [Reference standards](#)).

We considered studies of cross-sectional design only if they included participants with clinical suspicion of HCC. We excluded two-gate design studies ([DTA Handbook 2021](#)), that compared people with known HCC to matched controls as these studies are considered to have high risk of bias due to inflated accuracy estimates ([Colli 2014](#)). We included studies assessing MRI if all the participants had undergone testing with at least one of the acceptable reference standards. We excluded studies that analysed data only per lesion, rather than per participant, unless study authors made available per-participant data.

Participants

We included studies with adults (aged 16 years and above), with chronic liver disease, irrespective of aetiology, severity of disease, and duration of illness, with suspicion of having HCC based on prior tests, US or AFP, or both. The review focused on diagnostic questions related to adults with a first diagnosis of HCC.

Exclusion criteria

Adults with previous diagnosis and treatment of HCC make up a distinct group for which the diagnosis or natural history of HCC has been modified. These people were not the focus of this review; therefore, we excluded studies that included such participants unless they represented less than 5% of all the included participants, or if investigators had presented data in such

a way as to allow this group of participants to be isolated from the remaining included participants.

Index tests

MRI for the detection of HCC with the use of ECA or HBA, or both. Regarding positivity criteria, we accepted any definition of positivity explained in the studies.

Target conditions

- HCC of any size and at any stage.
- Resectable HCC (see [Secondary objectives](#)). The definition of resectable HCC is a neoplasm amenable to surgical radical resection according to the current guidelines ([Omata 2017](#); [EASL 2018](#); [Heimbach 2018](#)): a single lesion with a maximum diameter of less than 5 cm, or fewer than three lesions with a maximum diameter of 3 cm ([Mazzaferro 1996](#)).

Reference standards

We accepted as a reference standard for the diagnosis of HCC one of the following:

- the pathology of the explanted liver in case of transplantation;
- the histology of resected focal liver lesion(s), or the histology of resected or biopsied focal liver lesion(s) with a follow-up period of at least six months using periodic testing with US, AFP CT, or MRI. to exclude the presence of focal lesions not detected by the index test.

These two reference standards are not perfect. The pathology of the explanted liver is possible only in the case when all the included participants undergo liver transplantation; therefore, the setting does not correspond to the clinical question that only people with advanced and decompensated liver disease are candidates for OLT ([EASL 2016](#)). In the case of histology of resected focal lesion and histology of biopsied liver lesions, the negative result can be confirmed only with an adequate follow-up using periodic testing with US, AFP, CT, or MRI ([Nathani 2021](#)). Therefore, differential verification is unavoidable in this context ([Lijmer 1999](#)).

Search methods for identification of studies

Electronic searches

We searched the Cochrane Hepato-Biliary Group (CHBG) Controlled Trials Register and the CHBG Diagnostic Test of Accuracy Studies Register (the CHBG Information Specialist searched both registers via the Cochrane Register of Studies Web on 9 November 2021), the Cochrane Library (2021, Issue 11), MEDLINE Ovid (1946 to 9 November 2021), Embase Ovid (1974 to 9 November 2021), LILACS (Bireme; 1982 to 9 November 2021), Science Citation Index – Expanded (Web of Science; 1900 to 9 November 2021), and Conference Proceedings Citation Index – Science (Web of Science; 1990 to 9 November 2021) using the search strategies shown in [Appendix 1](#). We ran the searches on 24 February 2021, and then we reran them on 9 November 2021. We used the Cochrane Register of Studies Web for managing search results from the electronic searches and for identifying duplicates.

We applied no restrictions on language or document type.

Searching other resources

We tried to identify additional references by manually searching articles retrieved from digital databases and relevant review articles. We sought information on unpublished studies by contacting experts in the field. In addition, we handsearched abstract books from meetings of the AASLD, EASL, and APASL held during the 10 years prior to the search date (9 November 2021). We also searched for other types of grey literature in the System for Information on Grey Literature in Europe 'OpenGrey' (www.opengrey.eu/) on 9 November 2021. Further information was requested by contacting authors of studies by e-mail.

Data collection and analysis

We followed available guidelines as provided in *the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (DTA Handbook 2021)*.

Selection of studies

Two review authors (VG and TN) independently scrutinised titles and abstracts identified by electronic literature searching to identify potentially eligible studies. We selected any citation, identified by either of the two review authors, as potentially eligible for full-text review. The same review authors independently assessed full-text papers for study eligibility, using predefined inclusion and exclusion criteria. We resolved any discrepancies by discussion. We identified and excluded duplicates and collated multiple reports of the same study, so that each study rather than each report, was the unit of interest in the review. We recorded all studies after full-text assessment, and their reasons for exclusion, in the [Characteristics of excluded studies](#) table and illustrated the study selection process using a PRISMA diagram ([Moher 2009](#)).

Data extraction and management

We developed a standardised data extraction form and piloted the form on five included studies before finalising it. Then, two review authors (VG and TN) completed a piloted data extraction form for each included study. Each review author independently retrieved study data. In cases of disagreement, we reached consensus through discussion with a third review author (GC). We extracted the following data and completed a [Characteristics of included studies](#) table:

- general information: title, journal, year, publication type, and study design (prospective versus retrospective), surveillance programme, or clinical cohort;
- sample size: number of participants meeting the criteria and total number of participants included and tested;
- baseline characteristics: baseline diagnosis, age, sex, presence of cirrhosis or advanced chronic liver disease, and mean diameter of HCC;
- index test with predefined positivity criteria;
- type of contrast media used;
- reference standard tests;
- numbers of true positive, true negative, false positive, and false negative findings. We extracted these data for the two target conditions (HCC of any size and stage and resectable HCC);
- number of uninterpretable results;

- number of examinations not performed due to contraindications to MRI;
- possible conflict of interest of study authors.

We summarised the data from each study in 2×2 tables (true positive; false positive; false negative; true negative), according to the index test considered, and we entered the data into Review Manager 5 ([Review Manager 2020](#)).

Missing data

In the process of full-text study retrieval, we used available sources to retrieve the relevant studies. When full-text studies were not available, we contacted the primary authors directly by email to request the studies or data in question.

We contacted primary authors by email to request missing data that were needed to design the 2×2 tables. If we received no reply, we sent a second e-mail after two weeks. If no reply was received, we excluded the study in question. We reported on how many studies we had excluded for this reason.

Assessment of methodological quality

Two review authors (VG and TN) independently assessed the risk of bias of included studies and applicability of their results using QUADAS-2 (revised tool for quality assessment of diagnostic accuracy studies) ([Appendix 2](#); [Whiting 2011](#)). In cases of disagreement, we reached consensus through discussion. We addressed aspects of study quality involving the participant spectrum, index tests, target conditions, reference standards, and flow and timing. If some participants were not included in the analyses, we considered the study at high risk of bias. We classified a study at high risk of bias if at least one of the QUADAS-2 domains was judged at high risk.

We defined a time interval between the index test and the reference standard of three months as appropriate. According to a recent systematic review, the approximate HCC volume doubling time is four months to five months with significant range of 2.2 months to 11.3 months ([Nathani 2021](#)). In accordance with suggestions from a previous systematic review, which noted the acceptable time interval being from one month to three months ([Kim 2008](#)), we assumed 90 days to be the most acceptable threshold.

Statistical analysis and data synthesis

We carried out statistical analyses according to recommendations provided in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (DTA Handbook 2021)*. We performed a graphical descriptive analysis of the included studies. We reported forest plots (sensitivity and specificity separately, with their 95% confidence intervals (CIs)) and we provided a graphical representation of studies in the receiver operating characteristic (ROC) space (sensitivity against $1 - \text{specificity}$) and then, we performed a meta-analysis using the bivariate model and provided estimates of summary sensitivity and specificity. We used the pooled estimates obtained from the fitted models to calculate summary estimates of positive (LR+) and negative (LR-) likelihood ratios. We performed all statistical analyses using SAS statistical software, version 9.4 (SAS Institute Inc, Cary, NC, USA) and macro METADAS ([DTA Handbook 2021](#)).

Investigations of heterogeneity

We investigated the effects of the following sources of heterogeneity and rationale for our choice:

- study date (studies before compared to after the year 2011);
- inclusion of participants without cirrhosis: studies including 10% or more participants without cirrhosis compared to studies including less than 10% participants without cirrhosis;
- differences in prior tests: studies including participants who underwent US with or without AFP compared to studies with participants who underwent CT or CEUS;
- study location (population differences): studies conducted in the North and South America compared to Europe compared to Asia compared to Africa;
- participant selection: participants recruited from planned screening programmes compared to clinical cohorts;
- different HCC stage: studies with 20% or greater of resectable HCC compared to studies with less than 20% of resectable HCC;
- different reference standard: histology of the explanted liver compared to liver biopsy compared to another reference standard;
- different liver cirrhosis aetiology: studies including more than 80% of participants with hepatitis C or hepatitis B virus-associated cirrhosis compared to studies including more than 20% of participants with non-viral cirrhosis;
- studies with clear predefined MR positivity criteria compared to studies without predefined MR positivity criteria;
- studies using LI-RADS as MR positivity criteria compared to studies using other definitions of positivity criteria;
- studies using LI-RADS 5 only as MR positivity criteria compared to studies using LI-RADS 4 and 5 as positivity criteria;
- type of contrast media used;
- studies with radiologists (defined experts in MRI technique) compared to studies without any definition of operator's expertise.

We chose the above listed variables for the following reasons. Due to advancements in technology and change in diagnostic criteria, we considered the date of study publication. Searching the relevant literature, the earliest study on the accuracy of MRI for the diagnosis of HCC was published in 1998 (Hori 1998), and since then a significant number of studies have been published, reporting technological improvements. The LI-RADS diagnostic criteria were first presented in 2011, with several updated versions published during the following years. The latest was published in 2018 (LI-RADS 2018; Elsayes 2019). We chose 2011 as a cut-off value, separating studies published before and after the first LI-RADS criteria. The proportion of participants without cirrhosis is relevant because HCC in the absence of cirrhosis has different MRI characteristics. In epidemiological studies, this proportion is usually less than 10% (Lok 2009; Forner 2018a). Inclusion of participants who underwent US as the only prior test as opposed to those who underwent CT or CEUS, which might produce differences in MRI accuracy estimates secondary to this different selection. There are differences in epidemiology, and clinical and radiological characteristics of HCC in Asia when compared to Western countries and also Africa. Differences in clinical and radiological characteristics are also expected according to the selection of study participants: surveillance programme or

clinical setting. The proportion of resectable HCC found in the studies reflect different epidemiology and participant selection. The accuracy of MRI may vary according to the different reference standard, the type of contrast used, and the definition of positivity criteria. Different type of contrast media and operator's expertise may explain differences in interpretation of images.

We estimated effects by adding covariates to the bivariate models. We assessed the statistical significance of the covariate effect on sensitivity and specificity using the log-LR test for comparison of models with and without the covariate term. We considered P less than 0.05 as two-sided and statistically significant.

Sensitivity analyses

We assessed effects of risk of bias of included studies on diagnostic accuracy by performing a sensitivity analysis in which we excluded studies classified at high risk of bias in at least one of the domains of QUADAS-2 (Appendix 2). In addition, we defined the following signalling questions as most relevant, and we conducted a sensitivity analyses in which we excluded studies with answers of 'no' or 'unclear'.

- Were the positivity criteria defined?
- Were the reference standard results interpreted without the knowledge of the results of the index test?

We planned to conduct sensitivity analyses in which we excluded studies published only in abstract or letter form.

Assessment of reporting bias

We did not plan to test for publication bias due to the lack of validated methods for diagnostic test accuracy reviews.

Summary of findings

We prepared summary of findings tables to present the main results and key information regarding the certainty of evidence assessed using the GRADE approach (Balslem 2011; Schünemann 2020a; Schünemann 2020b). As recommended, we rated the certainty of evidence as high (not downgraded), moderate (downgraded by one level), low (downgraded by two levels), or very low (downgraded by more than two levels) based on five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias. For each outcome, the certainty of evidence starts as high when there are high-quality observational studies (cross-sectional or cohort studies) that enrolled participants with diagnostic uncertainty. When we found a reason for downgrading, we used our judgement to classify the reason as either serious (downgraded by one level) or very serious (downgraded by two levels) and recorded them in the footnotes.

We applied the GRADE judgements for the GRADE domains as following.

- Risk of bias: we used QUADAS-2 to assess risk of bias.
- Indirectness: we used QUADAS-2 for concerns of applicability and looked for important differences between the populations studied (e.g. the spectrum of disease), the setting, and the index test.
- Inconsistency: we carried out prespecified analyses to investigate potential sources of heterogeneity and downgraded

when we could not explain inconsistency in the accuracy estimates.

- Imprecision: we looked at the CIs of sensitivity and specificity estimates and at the unexplained heterogeneity of the results.
- Publication bias: we did not evaluate publication bias due to the lack of validated methods for diagnostic test accuracy reviews.

RESULTS

Results of the search

We ran the searches on 24 February 2021, and reran them on 9 November 2021. We identified 14,423 records by searching the Cochrane Hepato-Biliary Group Controlled Trials Register (46 records), the Cochrane Hepato-Biliary Group Diagnostic Test of Accuracy Studies Register (8), the Cochrane Library (351), MEDLINE Ovid (3074), Embase Ovid (6749), LILACS (60), and Science Citation Index – Expanded with Conference Proceedings Citation Index – Science (4135). We retrieved three additional records through handsearching other sources. We identified 4765 duplicates and excluded them from further analysis. After reading the title and abstract, we excluded 9544 records, as they did not meet the

inclusion criteria. We retrieved full texts of the remaining 117 records, and after reading the full texts, we excluded 83 studies for various reasons (see [Characteristics of excluded studies](#) table). Finally, we included in our review 34 records reporting data on 34 studies ([Figure 2](#)), including 4841 participants ([Born 1998](#); [de Lédinghen 2002](#); [Libbrecht 2002](#); [Bhartia 2003](#); [Teefey 2003](#); [Giorgio 2007](#); [Lauenstein 2007](#); [Hanna 2008](#); [Seçil 2008](#); [Golfieri 2009](#); [Sangiovanni 2010](#); [Yu 2011](#); [Di Carlo 2012](#); [Sersté 2012](#); [Dumitrescu 2013](#); [Hwang 2014](#); [Maiwald 2014](#); [Marks 2015](#); [Marrero 2005](#); [Lin 2016](#); [Villacastin Ruiz 2016](#); [Besa 2017](#); [Kim 2017](#); [Shin 2017](#); [Sutherland 2017](#); [McNamara 2018](#); [Min 2018a](#); [Brunsing 2019](#); [Demirtas 2020](#); [Khatri 2020](#); [Kim 2020](#); [Vietti Violi 2020](#); [Wu 2020](#); [Darnell 2021](#)). The three additional studies that were retrieved through handsearching were all included in the analysis ([Besa 2017](#); [Brunsing 2019](#); [Khatri 2020](#)). We applied no language restrictions in the inclusion criteria, which resulted in retrieving full-text articles of 17 studies published in non-English languages, of which two were included in the final analysis ([Born 1998](#); [Golfieri 2009](#)) after translation by a member of the review team (AC). Further information was requested by e-mail regarding three studies, and replies were not received ([Ueda 1995](#); [Puig 1997](#); [Simon 2005](#)). No data were obtained through correspondence with investigators.

Figure 2. Study flow diagram. Date of search 9 November 2021

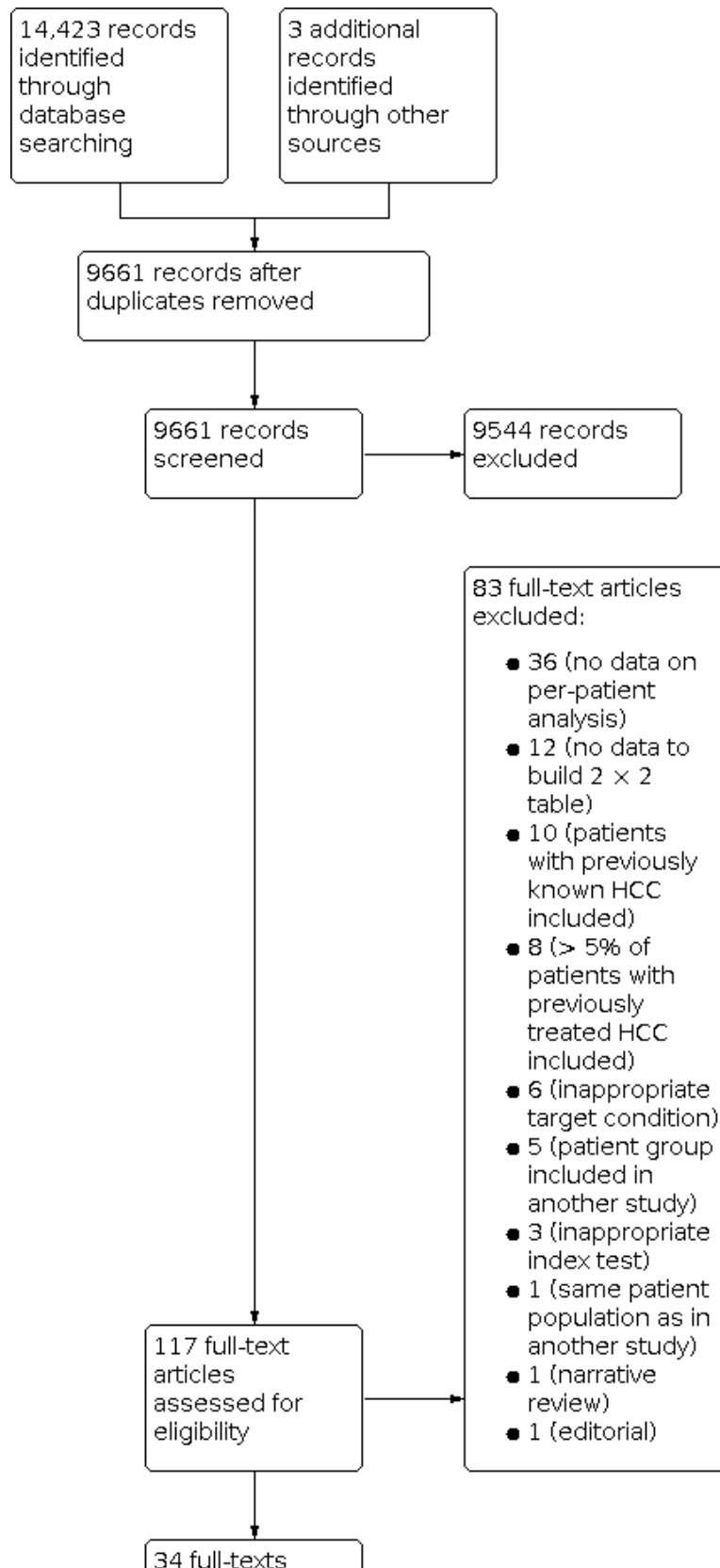
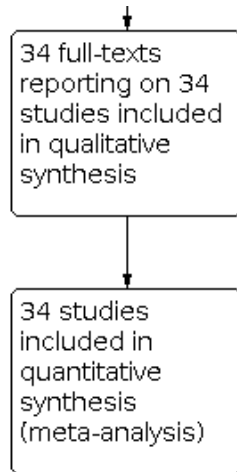


Figure 2. (Continued)



We reported the main characteristics of the 34 references in the [Characteristics of included studies](#) table. All references are reported as full-text publications, except one, which was published in abstract form only (Di Carlo 2012). The studies were conducted from 1998 to 2021.

Methodological quality of included studies

We reported in detail results of the quality assessment of included studies in the [Characteristics of included studies](#) tables, and we summarised this information in [Figure 3](#) and [Figure 4](#).

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

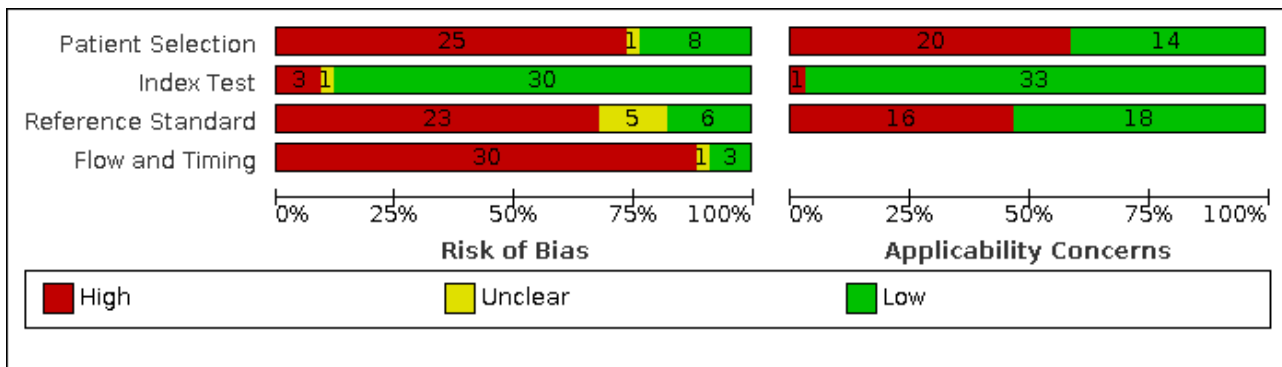
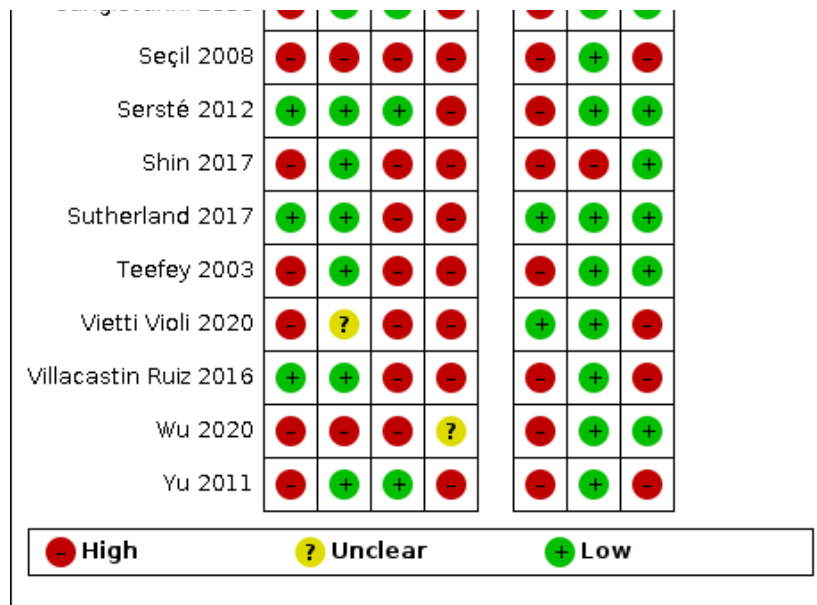


Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

	<u>Risk of Bias</u>				<u>Applicability Concerns</u>		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Besa 2017	-	+	-	-	+	+	+
Bhartia 2003	+	+	?	-	-	+	-
Born 1998	-	+	?	-	-	+	-
Brunsing 2019	-	+	-	-	+	+	-
Darnell 2021	-	+	-	-	+	+	+
de Lédinghen 2002	-	+	-	-	-	+	-
Demirtas 2020	-	+	-	-	+	+	-
Di Carlo 2012	?	+	-	-	+	+	-
Dumitrescu 2013	-	-	-	-	-	+	-
Giorgio 2007	-	+	?	+	-	+	+
Golfieri 2009	-	+	?	-	-	+	+
Hanna 2008	-	+	-	-	-	+	-
Hwang 2014	-	+	-	+	-	+	-
Khatri 2020	-	+	-	-	+	+	+
Kim 2017	-	+	-	-	+	+	+
Kim 2020	-	+	-	-	+	+	+
Lauenstein 2007	-	+	-	+	-	+	-
Libbrecht 2002	+	+	+	-	-	+	-
Lin 2016	-	+	+	-	+	+	+
Maiwald 2014	+	+	-	-	+	+	+
Marks 2015	+	+	-	-	+	+	+
Marrero 2005	-	+	-	-	-	+	+
McNamara 2018	+	+	+	-	-	+	-
Min 2018a	-	+	?	-	+	+	+
Sangiovanni 2010	-	+	+	-	-	+	+
Secil 2008	-	-	-	-	-	+	-

Figure 4. (Continued)



Participant selection

Risk of bias

Eight studies were at low risk of bias regarding patient selection (Libbrecht 2002; Bhartia 2003; Sersté 2012; Maiwald 2014; Marks 2015; Villacastin Ruiz 2016; Sutherland 2017; McNamara 2018). One study was judged unclear for this domain, since there were no data on the presence of exclusion criteria (Di Carlo 2012). Twenty-five were at high risk of bias due to exclusion criteria we considered inappropriate in the domain of population characteristics (de Lédinghen 2002; Marrero 2005; Lauenstein 2007; Seçil 2008; Golfieri 2009; Sangiovanni 2010; Lin 2016; Besa 2017; Kim 2017; Wu 2020), unavailable data (Born 1998; Teefey 2003; Hanna 2008; Yu 2011; Dumitrescu 2013; Hwang 2014; Min 2018a; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Vietti Violi 2020), or HCC features (Giorgio 2007; Shin 2017; Darnell 2021).

Applicability

We judged 14 studies at low concern regarding applicability (Di Carlo 2012; Maiwald 2014; Marks 2015; Lin 2016; Besa 2017; Kim 2017; Sutherland 2017; Min 2018a; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Vietti Violi 2020; Darnell 2021). The other 20 studies were judged at high concern because they included only participants with decompensated liver disease stage in waiting list for OLT (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Teefey 2003; Lauenstein 2007; Hanna 2008; Yu 2011; Hwang 2014; Villacastin Ruiz 2016; McNamara 2018), participants with a defined HCC diameter (Giorgio 2007; Golfieri 2009; Sangiovanni 2010; Sersté 2012; Shin 2017; Wu 2020), participants with suspected enhancing mass (Marrero 2005), participants with available MRI only (Seçil 2008), or not all participants having chronic liver disease (Dumitrescu 2013).

Index test

Risk of bias

We judged 30 studies regarding the index test at low risk of bias as they clearly predefined the MRI positivity criteria (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Teefey 2003; Marrero 2005; Giorgio 2007; Lauenstein 2007; Hanna 2008; Golfieri 2009; Sangiovanni 2010; Yu 2011; Di Carlo 2012; Sersté 2012; Hwang 2014; Maiwald 2014; Marks 2015; Lin 2016; Villacastin Ruiz 2016; Besa 2017; Kim 2017; Shin 2017; Sutherland 2017; McNamara 2018; Min 2018a; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Darnell 2021). One study was unclear for this domain due to unclear blinding to reference standard results (Vietti Violi 2020). Three studies were at high risk of bias due to undefined MRI positivity criteria (Seçil 2008; Dumitrescu 2013; Wu 2020).

Applicability

We judged 33 studies regarding the index test at low concern (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Teefey 2003; Marrero 2005; Giorgio 2007; Lauenstein 2007; Hanna 2008; Seçil 2008; Golfieri 2009; Sangiovanni 2010; Yu 2011; Di Carlo 2012; Sersté 2012; Dumitrescu 2013; Hwang 2014; Maiwald 2014; Marks 2015; Lin 2016; Villacastin Ruiz 2016; Besa 2017; Kim 2017; Sutherland 2017; McNamara 2018; Min 2018a; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Vietti Violi 2020; Wu 2020; Darnell 2021). We judged one study at high concern due to positivity criteria not being used in routine clinical practice (Shin 2017).

Reference standard

In 10 studies, the reference standard was the pathology of the explanted liver (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Lauenstein 2007; Hanna 2008; Yu 2011; Hwang 2014; Villacastin Ruiz 2016; McNamara 2018). In four studies the reference standard was histology in all participants (Giorgio 2007; Sangiovanni 2010; Sersté 2012; Shin 2017), in six studies it was OLT in some participants and histology in others (Teefey 2003; Golfieri

2009; Lin 2016; Besa 2017; Min 2018a; Wu 2020). In 14 studies, the reference standard was the combination of following options: OLT, histology (either biopsy or resection), or follow-up using US, CT, MRI, AFP, laboratory, and clinical data (Marrero 2005; Seçil 2008; Di Carlo 2012; Dumitrescu 2013; Maiwald 2014; Marks 2015; Kim 2017; Sutherland 2017; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Vietti Violi 2020; Darnell 2021).

Risk of bias

We judged six studies regarding the reference standard at low risk of bias (Libbrecht 2002; Sangiovanni 2010; Yu 2011; Sersté 2012; Lin 2016; McNamara 2018), five studies at uncertain risk (Born 1998; Bhartia 2003; Giorgio 2007; Golfieri 2009; Min 2018a), and 23 at high risk of bias (de Lédinghen 2002; Teefey 2003; Marrero 2005; Lauenstein 2007; Hanna 2008; Seçil 2008; Di Carlo 2012; Dumitrescu 2013; Hwang 2014; Maiwald 2014; Marks 2015; Villacastin Ruiz 2016; Besa 2017; Kim 2017; Shin 2017; Sutherland 2017; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Vietti Violi 2020; Wu 2020; Darnell 2021). Main reasons for judging studies at high risk of bias included statements explaining that the reference standard results were interpreted with the knowledge of the results of the index test, and in cases of biopsy, the procedure is usually performed after reviewing all available preprocedural imaging data. Uncertain risk of bias was judged due to lack of detailed information regarding the reference standard.

Applicability

We judged 18 studies regarding the reference standard at low concern (Teefey 2003; Marrero 2005; Giorgio 2007; Golfieri 2009; Sangiovanni 2010; Sersté 2012; Maiwald 2014; Marks 2015; Lin 2016; Besa 2017; Kim 2017; Shin 2017; Sutherland 2017; Min 2018a; Khatri 2020; Kim 2020; Wu 2020; Darnell 2021). Sixteen studies were at high concern due to OLT being the only reference standard (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Lauenstein 2007; Hanna 2008; Yu 2011; Hwang 2014; Villacastin Ruiz 2016; McNamara 2018), and use of other inappropriate reference standards (clinical and laboratory data, US, CEUS) (Seçil 2008; Di Carlo 2012; Dumitrescu 2013; Brunsing 2019; Demirtas 2020; Vietti Violi 2020).

Flow and timing

Risk of bias

We judged three studies at low risk of bias regarding flow and timing (Giorgio 2007; Lauenstein 2007; Hwang 2014). Thirty studies were at high risk due to: inappropriate time between index test and reference standard (greater than 90 days) (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Teefey 2003; Hanna 2008; Golfieri 2009; Yu 2011; Marks 2015; Villacastin Ruiz 2016; Besa 2017; McNamara 2018; Brunsing 2019; Darnell 2021), not all participants underwent the same reference standard (Teefey 2003; Marrero 2005; Giorgio 2007; Seçil 2008; Di Carlo 2012; Dumitrescu 2013; Maiwald 2014; Marks 2015; Lin 2016; Besa 2017; Shin 2017; Sutherland 2017; Min 2018a; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Vietti Violi 2020; Darnell 2021), or participants missing in the final analysis with no explanation (Marrero 2005; Sangiovanni 2010; Sersté 2012; Villacastin Ruiz 2016; Brunsing 2019; Khatri 2020). One study was at uncertain risk of bias due to lack of information on time interval between index test and reference standard (Wu 2020). Nine studies reported non-evaluable results

(Lauenstein 2007; Hwang 2014; Besa 2017; McNamara 2018; Min 2018a; Brunsing 2019; Khatri 2020; Kim 2020; Darnell 2021).

Overall assessment

All included studies were at overall high risk of bias. We judged five studies at low concern for applicability (Maiwald 2014; Lin 2016; Besa 2017; Kim 2017; Min 2018a).

Findings

Thirty-four studies with 4841 participants provided data assessing MRI for the diagnosis of HCC. The median prevalence of the target disease was 56% (interquartile range 36% to 66%).

Thirty-two studies reported the prevalence of participants with hepatic cirrhosis, and in 25 of them the reported prevalence was 100%. Fourteen studies reported the Child-Pugh classification with a median of 76% (interquartile range 57% to 79%) classified as Child-Pugh class A. Twenty-nine studies reported information on liver disease aetiology and a median of 68% (interquartile range 48% to 79%) had viral aetiology. Twenty-six studies reported the proportion of participants with resectable HCC, among which 16 reported all participants to have resectable HCC. Twenty-three studies reported the mean diameter of the lesions with a median of 23 mm (interquartile range 18 mm to 32 mm). The studies were conducted from 1998 to 2021.

Regarding study location, 15 studies were conducted in Europe, 11 in North America, and eight in Asia. Twenty studies were conducted in people with clinical suspicion of having HCC, six were conducted in the context of a surveillance programme, and eight performed MRI as a confirmatory test after a surveillance programme. Fourteen studies reported the number of uninterpretable index test results (Born 1998; de Lédinghen 2002; Bhartia 2003; Lauenstein 2007; Hwang 2014; Besa 2017; Kim 2017; McNamara 2018; Min 2018a; Brunsing 2019; Demirtas 2020; Khatri 2020; Kim 2020; Darnell 2021), ranging from 0/407 to 19/141. Two studies reported the number of examinations not performed due to contraindications. Demirtas 2020 reported 23/294 participants and Sangiovanni 2010 reported 2/55 participants.

Seventeen studies reported no information about authors' possible conflict of interest (Born 1998; de Lédinghen 2002; Libbrecht 2002; Bhartia 2003; Teefey 2003; Marrero 2005; Giorgio 2007; Lauenstein 2007; Hanna 2008; Seçil 2008; Golfieri 2009; Di Carlo 2012; Sersté 2012; Lin 2016; McNamara 2018; Kim 2020; Khatri 2020), 10 reported possible conflict of interest (Yu 2011; Maiwald 2014; Marks 2015; Besa 2017; Kim 2017; Shin 2017; Sutherland 2017; Brunsing 2019; Vietti Violi 2020; Darnell 2021), and seven reported no possible conflict of interest (Sangiovanni 2010; Dumitrescu 2013; Hwang 2014; Villacastin Ruiz 2016; Min 2018a; Demirtas 2020; Wu 2020).

Among the 10 studies with the pathology of explanted liver as the reference standard, four studies reported no alternative diagnosis in participants without HCC (Born 1998; Bhartia 2003; Lauenstein 2007; McNamara 2018). de Lédinghen 2002 reported eight dysplastic nodules and six macroregenerative nodules in 13 participants without HCC; Libbrecht 2002 reported one haemangioma and one focal nodular hyperplasia in 14 participants without HCC; Hanna 2008 reported three focal areas of fibrosis, three vessels and three benign regenerative nodules in 23 participants without HCC; Yu 2011 reported six dysplastic or

regenerative macronodules, two haemangiomas and one focal infarct in 99 participants without HCC; [Hwang 2014](#) reported 15 dysplastic nodules and three large regenerative nodules in 11 participants without HCC; and [Villacastin Ruiz 2016](#) reported six cholangiocarcinomas, two haemangiomas, and six dysplastic nodules in 164 participants without HCC.

In the four studies with histology of biopsied focal lesions in all participants as the reference standard [Giorgio 2007](#) reported eight regenerative nodules, four dysplastic nodules, six areas of focal steatosis, four haemangiomas, one metastasis, one non-Hodgkin's lymphoma, and one focal nodular hyperplasia in 25 participants without HCC; [Sangiovanni 2010](#) reported

two cholangiocarcinomas, three low-grade dysplastic nodules, and 18 macroregenerative nodules in 22 participants without HCC; [Sersté 2012](#) reported seven dysplastic nodules, nine macroregenerative nodules, one cholangiocarcinoma, one epithelioid haemangioendothelioma, and nine areas of chronic liver disease in 27 participants without HCC; and [Shin 2017](#) reported one dysplastic nodule in 18 participants without HCC.

[Figure 5](#) shows a graphical representation of studies in the ROC space (sensitivity against 1 – specificity) and [Figure 6](#) shows a forest plot of sensitivity and specificity with their 95% CIs. For the 34 studies, the reported sensitivity ranged from 44% to 96% and the specificity ranged from 33% to 100%.

Figure 5. Summary receiver operating characteristic (ROC) comparing magnetic resonance imaging (MRI) and different reference standards in 34 studies. Reference standards were: the pathology of the explanted liver in case of transplantation; the histology of resected focal liver lesions, or the histology of biopsied focal liver lesion(s) with a follow-up period of at least six months, using periodic testing with ultrasound, alpha-fetoprotein, computed tomography or MRI.

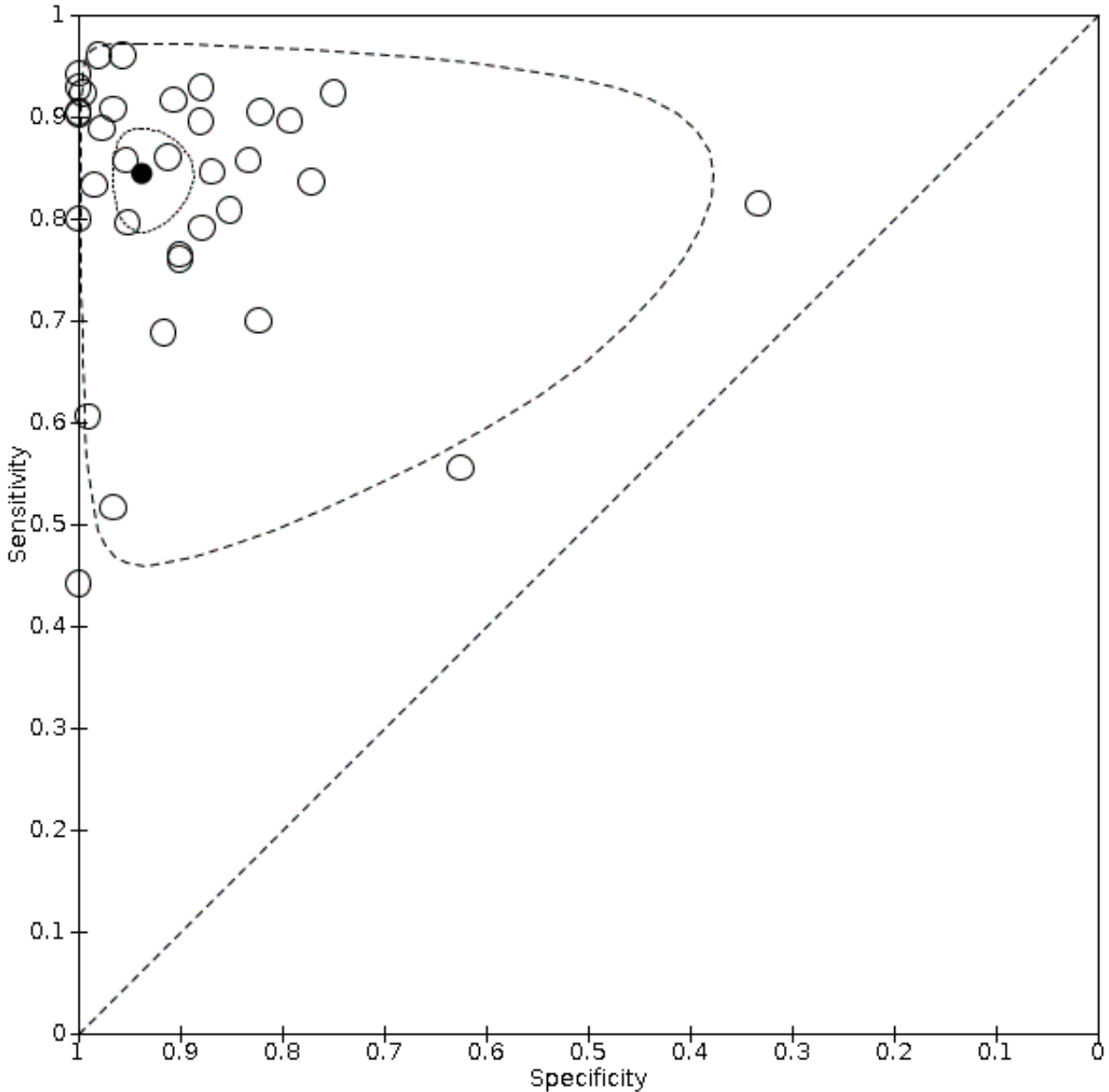
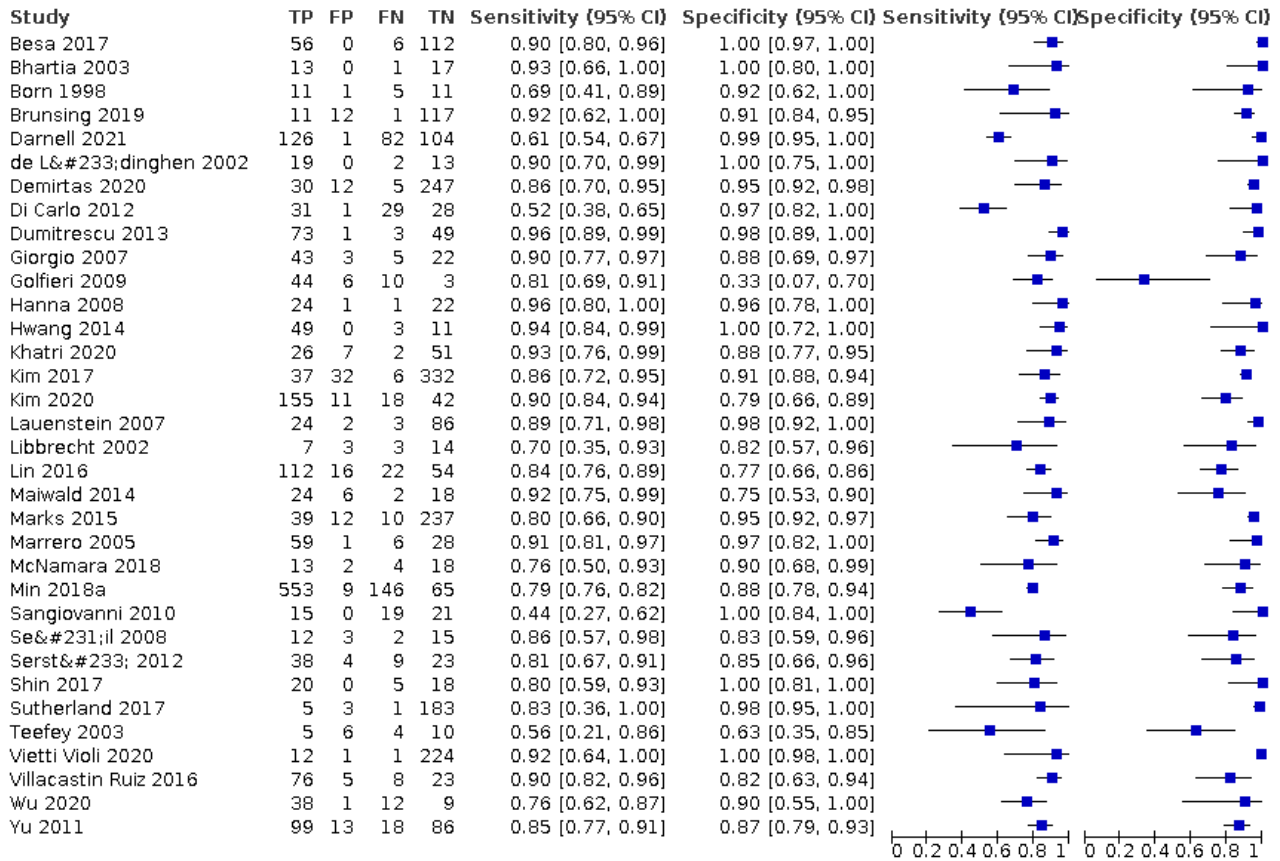


Figure 6. Forest plots of sensitivity and specificity of magnetic resonance imaging for detection of hepatocellular carcinoma of any size and stage against different reference standards in 34 studies in alphabetical order. Reference standards were: the pathology of the explanted liver in case of transplantation, the histology of resected focal liver lesions, or the histology of biopsied focal liver lesions with a follow-up period of at least six months. Values between square brackets are the 95% confidence intervals (CIs) of sensitivity and specificity. The figure shows the estimated sensitivity and specificity of the study (blue square) and its 95% CI (black horizontal line). CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.

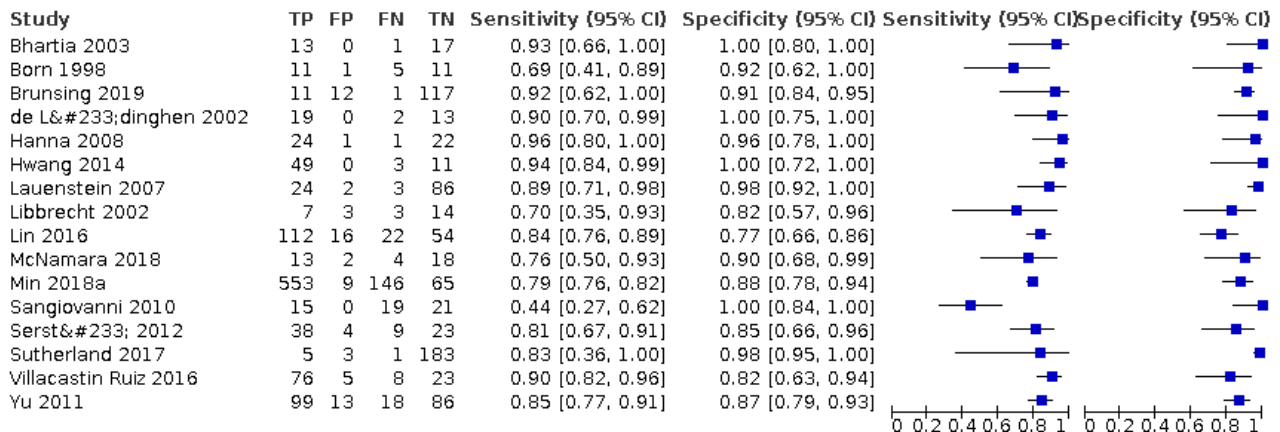


We performed a meta-analysis of all 34 included studies using the bivariate model, and we obtained the following pooled estimates: sensitivity 84.4% (95% CI 80.1% to 87.9%); specificity 93.8% (95% CI 90.1% to 96.1%); LR+ 13.5 (95% CI 8.5 to 21.7); and LR- 0.17 (95% CI 0.13 to 0.21).

Table 1 shows post-test probabilities calculated using pooled LRs, according to three different pre-test probabilities, the median and interquartile range of HCC prevalence derived from our study analysis.

We assessed the diagnostic accuracy for resectable HCC as a secondary objective. We found 16 studies including all participants with resectable HCC (Born 1998; de Lédizinghen 2002; Libbrecht 2002; Bhartia 2003; Lauenstein 2007; Hanna 2008; Sangiovanni 2010; Yu 2011; Sersté 2012; Hwang 2014; Lin 2016; Villacastin Ruiz 2016; Sutherland 2017; McNamara 2018; Min 2018a; Brunsing 2019). We performed a meta-analysis and obtained the following estimates: sensitivity 84.3% (95% CI 77.6% to 89.3%); specificity 92.9% (95% CI 88.3% to 95.9%); LR+ 11.9 (95% CI 7.0 to 20.2); and LR- 0.17 (95% CI 0.12 to 0.25). Figure 7 shows the forest plot of sensitivity and specificity with their 95% CIs.

Figure 7. Forest plots of sensitivity and specificity of magnetic resonance imaging for detection of resectable hepatocellular carcinoma against different reference standards in 16 studies in alphabetical order. Reference standards were: the pathology of the explanted liver in case of transplantation, the histology of resected focal liver lesions, or the histology of biopsied focal liver lesions with a follow-up period of at least six months. Values between brackets are the 95% confidence intervals (CIs) of sensitivity and specificity. The figure shows the estimated sensitivity and specificity of the study (blue square) and its 95% CI (black horizontal line). CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.



Heterogeneity analysis

We investigated heterogeneity for all the predefined potential sources (Secondary objectives). Table 2 shows the comparisons of the different predefined subgroups. The prevalence of viral aetiology may in part explain the inconsistency of the overall results. In fact, studies which included less than 80% of participants with viral aetiology showed a higher sensitivity (87.6%, 95% CI 83.8% to 90.7% compared to 74.9%, 95% CI 64.4% to 83.1%) and a lower specificity (94.5%, 95% CI 90.7% to 96.8% compared to 96.5%, 95% CI 68.6% to 99.7%) than studies which included greater than 80% of participants with viral aetiology. Another possible source of heterogeneity was the study setting (clinical setting, confirmatory test after screening, or initial screening test). The sensitivity was lowest in the setting of confirmatory test after screening (77.1%, 95% CI 63.7% to 86.6% compared to 85.0%, 95% CI 78.0% to 90.1% in clinical setting and 86.8%, 95% CI 82.9% to 90.0% in initial screening), while specificity was the highest in initial screening test setting (96.2%, 95% CI 92.2% to 98.2% compared to 91.7%, 95% CI 75.1% to 97.6% in clinical setting and 93.3% 95% CI 88.1% to 96.3% in the setting of confirmatory test after screening). The comparison of the other subgroups assessing the possible role of study date and location, inclusion of participants without cirrhosis, the proportion of included participants with resectable HCC, the use of different contrast media, and the use of different reference standard did not show any differences.

Sensitivity analysis

When considering the 31 studies that clearly prespecified the positivity criteria, we obtained a pooled sensitivity of 83.9% (95% CI 79.3% to 87.5%) and a specificity of 93.8% (95% CI 89.9% to 96.3%) (Table 2).

When considering only the seven studies in which the reference standard results were interpreted without the knowledge of the results of the index test, we obtained a pooled sensitivity of 76.8%

(95% CI 66.4% to 84.7%) and a specificity of 89.3% (95% CI 81.5% to 94.0%) (Table 2).

We did not perform the planned sensitivity analysis in which studies published only in abstract or letter form were excluded because only one study was published in abstract form (Di Carlo 2012). All other included studies were published as full-text articles.

We did not perform the planned sensitivity analysis in which studies at high risk of bias were excluded as all the included studies were judged at high risk of bias.

Summary of findings tables

The main results are shown in the Summary of findings 1 and Summary of findings 2.

DISCUSSION

Summary of main results

The aim of this review was to assess the diagnostic accuracy of MRI for the diagnosis of HCC of any size and at any stage in adults with chronic liver disease. We included 34 studies that assessed 4841 participants, 20 were conducted in people with clinical suspicion of having HCC, six were conducted in the context of a surveillance programme, and eight performed MRI as a confirmatory test after a surveillance programme. The main results are presented in Summary of findings 1 and Summary of findings 2.

For the 34 included studies, we performed a meta-analysis using the bivariate model, and we obtained the following pooled estimates: sensitivity 84.4% (95% CI 80.1% to 87.9%) and specificity 93.8% (95% CI 90.1% to 96.1%) for the diagnosis of HCC at any size and stage (primary outcome). In Table 1, we showed the post-test probability of having HCC in the case of positive or negative result of the index test, assuming different values of pretest probability.

Sixteen studies included only participants with HCC amenable for surgical resection, and the pooled estimate of sensitivity was 84.3% (95% CI 77.6% to 89.3%) and specificity 92.9% (95% CI 88.3% to 95.9%) for the diagnosis of resectable HCC (secondary outcome).

We judged all included studies at high risk of bias in at least one domain, and we assessed the results of 29/34 studies to be at high concern for applicability.

We summarised these main results of analyses in [Summary of findings 1](#) and [Summary of findings 2](#), assuming three different prevalence values (36%, 56%, and 66%). The prevalence of HCC varied widely in all included studies, from 3% to 90%, according to the study design and different settings. For exemplification, we considered three values of HCC prevalence: 36% for a population with low clinical suspicion, 56% as a median derived from our study analysis, and 66% for a population with high clinical suspicion. These values represent the median and interquartile range of HCC prevalence derived from our study analysis.

For participants with HCC at any size and stage, we assumed the following consequences of test results: people with true-positive results, that is, those with HCC and positive test results, will receive the appropriate treatment (surgery, local ablative therapy, or systemic chemotherapy); people with true-negative results, that is, those without HCC and negative test results, will not undergo inappropriate treatment or unnecessary further testing; people with false-negative results, that is, those with HCC and negative test results, will be misdiagnosed, not receive the appropriate treatment, and might be detected later with more severe HCC; people with false-positive results, that is, those without HCC and positive test results, will undergo further testing and possibly an inappropriate treatment.

Considering a hypothetical cohort of 1000 people with HCC prevalence of 56% (the median value in the included studies), we can expect 87 false-negative and 27 false-positive results; with a lower prevalence of 36%, we can expect 56 false-negative and 40 false-positive results, and with a higher prevalence of 66%, we can expect 103 false-negative and 21 false-positive results. We judged the certainty of evidence to be low, downgrading two levels due to high risk of bias and indirectness.

For participants with resectable HCC, considering a hypothetical cohort of 1000 people with HCC prevalence of 56%, we can expect 88 false-negative and 31 false-positive results; with a prevalence of 36%, we can expect 57 false-negative and 45 false-positive results; with a prevalence of 66%, we can expect 104 false-negative and 24 false-positive results. We judged the certainty of evidence to be low, downgrading two levels due to high risk of bias and indirectness.

Strengths and weaknesses of the review

This review included 34 studies, covering a time span of 23 years, from 1998 to 2021 and wide geographical areas, including areas with high and low prevalence of chronic liver disease and HCC.

Fifteen studies were conducted in Europe, 11 in the USA, and eight in Asia. In terms of number of participants, studies performed in Asia included 1969 participants, the USA 1471 participants, and Europe 1401. We found no studies from Africa, where HCC is highly prevalent ([Ferlay 2019](#)).

An overall quality assessment of the studies showed their methodological weaknesses. We judged all studies at high risk of bias mainly due to inappropriate exclusion criteria, unavailable data, reference standard results interpreted with knowledge of the index test (unavoidable in cases of biopsy), fact that not all participants underwent the same reference standard, and time interval between index test and reference standard being more than 90 days. The choice of reference standard represents a major concern for all studies, and we recognise none is perfect. Fourteen of 34 studies used as the reference standard the combination of OLT, histology (either biopsy or resection), or follow-up using US, CT, MRI, AFP, laboratory, and clinical data. This choice, even if it reflects usual clinical practice, introduces an unavoidable different-verification bias as the results of the index test influence the decision on which reference standard is used. In contrast, the most common single reference standard used in 10/34 studies, was the pathology of the explanted liver. It is the most accurate reference standard, allowing the histological evaluation of the whole liver in all participants. However, this almost perfect reference standard is feasible only in studies conducted on participants with advanced and decompensated liver diseases on a waiting list for transplantation, that do not represent the intended spectrum of liver disease severity. In fact, the aim of the present review was to assess MRI accuracy in participants with the whole spectrum of liver disease severity without any exclusion for severity of liver disease or HCC volume. Accordingly, correct estimates of MRI accuracy can be obtained only at the expense of their applicability. In 4/34 studies, the reference standard was histology in all participants avoiding a different-verification bias but preventing blinding to the results of the index test as biopsy is usually performed after reviewing all the available preprocedural imaging data with a consequent high risk of bias.

We judged 14 studies in which the time interval between the index test and reference standard was longer than 90 days to be at high risk of bias. In fact, in diagnostic test accuracy assessment, it is necessary to have the time interval between index test and reference standard as short as possible ([Colli 2014](#)). Longer time intervals impair accurate assessment due to possible changes in lesion size and morphological features during certain periods of time. According to the latest systematic review, the approximate HCC volume doubling time is four to five months, with significant range of 2.2 months to 11.3 months ([Nathani 2021](#)). In accordance with suggestions from a previous systematic review that noted the acceptable time interval being from one to three months ([Kim 2008](#)), we assumed 90 days to be the most acceptable threshold.

We found two studies reporting the number of examinations not performed due to contraindications (7.8% and 3.6%), and 14 studies reporting the number of uninterpretable index test results (ranging from 0% to 13%). In the process of visual interpretation of MRI examinations, sometimes it is impossible for the radiologist to make a definite diagnosis of HCC. This is primarily due to unclear visual representation and absence of morphological criteria needed for a definite diagnosis (non-rim-like hyperenhancement, non-peripheral washout in portal-venous and subsequent phases, enhancing capsule, etc.) ([LI-RADS 2018](#)). Technical aspects of an MRI examination such as participant movement and breath-hold, scanning protocol, application of adequate type and amount of contrast, and acquisition of correct phases and sequences (arterial, portal-venous, late phase) can impair liver imaging and its correct interpretation. All the studies

that reported some uninterpretable results excluded these results from analyses preventing an assessment of their effect on the accuracy estimates.

Using QUADAS-2, we judged 29/34 studies at high concern for applicability mainly due to the selective inclusion of participants with decompensated advanced liver disease or a definite HCC diameter, and the use of pathological examination of the whole liver as the reference standard.

Not all studies reported on all covariates that we planned to assess as a possible source of heterogeneity, and this might have impaired the analyses. Most information on MELD (Model for End-Stage Liver Disease) and Child-Pugh class A stage in the studies was missing.

Strengths and weaknesses of the review process

Search strategy

Our search strategy provided a significant number of studies performed in various geographical areas with high and low prevalence of chronic liver disease and HCC. Manually searching the references of the included studies and previous narrative and systematic reviews identified three additional studies, which were included in the final analysis. We applied no language restrictions in the inclusion criteria, which resulted in retrieving full-text articles of 17 studies published in non-English languages, two of which we included in the final analysis. We requested further information from study authors regarding five studies, but they provided no information. We are confident that the search strategy resulted in the detection of most eligible studies, with a low probability of undetected relevant studies.

Quality assessment and data extraction

We consider our attempts to reduce subjectivity in our judgements to minimise errors and miscalculations in data extraction to be the strength of this review. Two review authors independently assessed the risk of bias of the included studies and applicability of their results using the QUADAS-2 tool. We extracted data using a standardised and piloted data extraction form. In case of disagreement, we reached consensus through discussion. Disagreements were most frequent for the two QUADAS-2 domains participant selection (10 studies), and reference standard (six studies). All agreements were reached through discussion between two review authors, and the conclusions were discussed and approved by a third review author. For data extraction, most of the discordances were due to miscalculations and typographical errors, which were easily resolved. The same review authors assessed the certainty of evidence using the GRADE approach and the level of agreement was high.

Review analysis

We performed a meta-analysis using the bivariate model, as the results of the index test were reported as dichotomous (positive or negative) with no explicit threshold. We recognise that implicit thresholds cannot be excluded. The pooled estimates of sensitivity ranged from 44% to 96% and those of specificity from 33% to 100%.

Three studies included fewer than 30 participants and their results were quite imprecise with very wide CIs (Born 1998; Libbrecht 2002; Teeffey 2003; Figure 6).

Inconsistency of the overall results may in part be explained by the inclusion of participants with viral aetiology of chronic liver disease and study setting. Studies which included less than 80% of participants with viral aetiology showed a higher sensitivity and a lower specificity than studies which included more than 80% of participants with viral aetiology. Another possible source of heterogeneity was the study setting (clinical setting, confirmatory test after screening, or initial screening test). The sensitivity was lowest in the setting of confirmatory test after screening, while specificity was the highest in initial screening test setting. Moreover, different geographic areas, advancements in technology (studies published before and after the year 2011), severity of the underlying disease (prevalence of cirrhosis), difference in the choice of the reference standards, use of LI-RADS positivity criteria, use of different contrast types, clear definition of positivity criteria, and operator expertise seem unable to explain the observed inconsistencies. Some of our planned investigations were not possible due to lack of data (MELD score, Child-Pugh classification of severity of cirrhosis), and lack of published studies (comparison of studies using LI-RADS 5 only as MR positivity criteria compared to studies using LI-RADS 4 and 5 as positivity criteria, and comparison of studies including participants who underwent US with or without AFP compared to studies with participants who underwent CT or CEUS).

Furthermore, we were able to investigate only characteristics that could be assessed at study level whereas participants' factors or HCC characteristics can only be assessed by aggregate statistics with the inherent risk of ecological bias (Robinson 1950; Reade 2008). Therefore, some important relationships, such as the one with HCC volume, could have been missed. In addition, many of the included studies did not report data on the covariates of interest.

We excluded studies that reported only per-lesion analyses and included only studies with per-patient analyses. Per-patient and per-lesion analyses represent two different approaches to diagnostic accuracy assessment and their choice depends on the type of clinical or scientific question, and requires different and appropriate statistical methodology. In the present review, we aimed to assess the accuracy of MRI for the diagnosis HCC. Consequently, we chose to include studies that evaluated how MRI is able to detect people with HCC at any size and any stage, therefore we applied a per-patient approach. Otherwise, per-lesion analysis is properly used to assess accuracy in detecting multiple lesions on a single image, providing information that is relevant for HCC staging. Studies planning per-lesion analysis require a different methodological approach and cannot be pooled with studies using a per-patient approach (Chang 2006; Zwinderman 2008). Furthermore, the inclusion criteria of studies planning a per-lesion analysis are quite different and do not match our review question. In fact, they usually do not include people with chronic liver disease and suspected HCC, but people with known focal liver lesions, encompassing HCCs, cholangiocarcinomas, benign liver tumours, and even metastases from abdominal or extra-abdominal primary cancers.

Most studies (31/34) reported a clear definition of diagnostic criteria and we tried to explore the effect on the diagnostic accuracy estimates of different criteria, traditional perfusion compared to LI-RADS criteria. However, only 5/34 studies used LI-RADS positivity criteria and we were unable to find a difference.

We were also unable to estimate the effect of uninterpretable results as only 14 studies reported the frequency of technical failures that excluded these results from analysis. The exclusion of uninterpretable results could have produced an overestimation of the accuracy estimates (Cohen 2016).

The sensitivity analysis shows that the obtained results are arguably robust with no variation, after including only studies that clearly prespecified the positivity criteria, and including only those in which the reference standard results were interpreted without the knowledge of the results of the index test.

Comparison with previous research

We found 21 non-Cochrane systematic reviews or reviews that assessed the accuracy of MRI for detection of HCCs (Colli 2006; Xie 2011; Chen 2013; Liu 2013; Wu 2013; Chen 2014; Junqiang 2014; Chou 2015; Lee 2015; Ye 2015; Li 2015a; Li 2015b; Guo 2016; Hanna 2016; Kierans 2016; Liu 2017; Roberts 2018; Li 2019; Chan 2021; Feng 2021; Gupta 2021).

Ten reviews assessed the accuracy of MRI alone (Liu 2013; Wu 2013; Chen 2014; Junqiang 2014; Li 2015a; Li 2015b; Kierans 2016; Chan 2021; Feng 2021; Gupta 2021); seven assessed the accuracy of CT and MRI (Chen 2013; Lee 2015; Ye 2015; Guo 2016; Liu 2017; Roberts 2018; Li 2019); one assessed US, CT, and MRI (Hanna 2016); one assessed US, CEUS, CT, and MRI (Chou 2015); one assessed CEUS, CT, and MRI (Xie 2011); and one assessed AFP, US, CT, and MRI (Colli 2006). Due to differences in methodological approaches, inclusion and exclusion criteria, and in statistical analyses, these results are not comparable to each other neither to our present results. Six reviews performed per-patient analysis, and the pooled sensitivity of MRI for the diagnosis of HCC in these reviews ranged from 80.6% to 91% and the specificity from 84.8% to 94% (Table 3) (Colli 2006; Chen 2014; Chou 2015; Lee 2015; Chan 2021; Gupta 2021). These results are in accordance with our present results, despite the methodological differences and the number of included studies. We additionally evaluated all the primary studies included in these systematic reviews and assessed them for inclusion in our analysis.

Fifteen reviews performed per-lesion analysis and the pooled sensitivity of MRI for detection of HCC ranged from 74% to 95%, and specificity from 78% to 96% (Table 3) (Xie 2011; Chen 2013; Liu 2013; Wu 2013; Junqiang 2014; Ye 2015; Li 2015a; Li 2015b; Guo 2016; Hanna 2016; Kierans 2016; Liu 2017; Roberts 2018; Li 2019; Feng 2021).

Applicability of findings to the review question

Using the QUADAS-2 tool, we assessed the applicability of the results of the included studies. We judged only five studies to be at low concern for applicability, and we downgraded one level the certainty of evidence because of indirectness the other 29 studies. Twenty studies selected participants based on the diameter of the focal lesions or included only people from the waiting list for OLT excluding participants with less severe disease. In 16 studies, the choice of the pathology of the explanted liver as the reference standard also impaired the applicability of the results as this reference standard is applied exclusively to participants who had received a transplant.

AUTHORS' CONCLUSIONS

Implications for practice

In the clinical pathway for the diagnosis of hepatocellular carcinoma (HCC) in people with chronic liver disease, magnetic resonance imaging (MRI), as an alternative to computed tomography (CT), is currently the second step after ultrasound and alpha-fetoprotein, or the combination of the two, and its main role is to confirm the presence of the disease. As an ideal diagnostic test, MRI should ensure a low proportion of false-negative results because people with undetected HCC cannot receive proper treatment. Meanwhile, people with false-positive results are exposed to unnecessary further diagnostic workup and possible invasive treatment. We meta-analysed the results of 34 studies and estimated a sensitivity of about 84% and a specificity of about 94% suggesting that 16% of people with HCC would be missed, and 6% of people would receive unnecessary additional tests or even treatments. Considering the 16 studies that assessed the MRI diagnostic accuracy for resectable HCC, we obtained similar estimates than for any size any stage HCC; 16% of people with HCC would be incorrectly classified as without any HCC, while 7% of people without carcinoma will undergo inappropriate further testing or surgery. For people on a waiting list for orthotopic liver transplantation for an indication not related to a HCC, the consequences of false-negative results of preoperative MRI are not completely known and might be less severe: studies report no significant difference in terms of overall survival and tumour recurrence compared to people with previously diagnosed HCCs (Castillo 2009; Senkerikova 2014; Madaleno 2015; El Moghazy 2016).

The main hallmarks of HCC on an MRI study are non-rim-like hyperenhancement in the arterial phase, and washout in the portal-venous and delayed phases. However, around 40% of HCCs present with atypical morphological features, which pose a significant diagnostic challenge for radiologists. This significant number of atypical HCCs may influence the sensitivity, and the radiologist should be acquainted to these atypical appearances to correctly interpret MRI findings. Another issue is the presence of HCC mimickers, such as intrahepatic cholangiocarcinoma, combined HCC-cholangiocarcinoma, arterioportal shunt, or haemangioma in cirrhotic liver (Lee 2012b; Shirki 2015).

The MRI accuracy estimates might only indirectly be compared with the results of CT obtained in a recent systematic review (Nadarevic 2021a) showing a sensitivity of 77.5% (95% CI 70.9% to 82.9%) and a specificity of 91.3% (95% CI 86.5% to 94.5%). However, a direct comparison of the diagnostic modalities in the same participants is needed to support the choice between these techniques which depends also on their availability, costs, and risks.

Overall, caution is needed in interpreting our review results as we judged all the studies at high risk of bias, and most of them with high concern regarding their applicability, mainly due to patient selection and reference standard domain.

Implications for research

Currently, available evidence on the diagnostic accuracy of MRI for diagnosis of HCC is inconclusive. Therefore, more high-quality primary studies are needed. With the introduction of LI-RADS criteria, there is no longer necessary to dichotomise results of MRI studies as these criteria also allow assessment of inconclusive and

probable results. Apart from typical HCC appearances, atypical features of HCC need to be taken into consideration, so we hypothesise that further studies using Liver Imaging Reporting and Data System (LI-RADS) positivity criteria may improve sensitivity. Also, it may be possible that including additional major features such as threshold growth, along with arterial hyperenhancement and subsequent washout, may improve sensitivity. Therefore, we welcome future cross-sectional studies using score systems of positivity criteria. A direct comparison of MRI and CT as a second step in the clinical pathway for the diagnosis of HCC after ultrasound and alpha-fetoprotein is also needed.

ACKNOWLEDGEMENTS

Cochrane Review Group funding acknowledgement: the Danish State is the largest single funder of the Cochrane Hepato-Biliary Group through its investment in the Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital-Rigshospitalet, Copenhagen, Denmark. Disclaimer: the views and opinions expressed in this review are those of the authors and do not necessarily reflect those of the Danish State or the Copenhagen Trial Unit.

Cochrane Hepato-Biliary and Cochrane Diagnostic Test Accuracy Editorial Teams supported the authors in the development of this review.

The following people from the Cochrane Hepato-Biliary (CHB) Editorial Team conducted the editorial process for this article:

- Sign-off Editor (final editorial decision): Christian Gluud, Coordinating Editor CHBG, Denmark;
- Contact Editor (provided editorial decision): Chavdar Pavlov, Russia;
- Managing Editor (selected peer reviewers, provided editorial guidance to authors, edited the article): Dimitrinka Nikolova, Denmark;
- Information Specialist (database searches): Sarah Louise Klingenberg, Denmark;
- Peer-reviewer (provided clinical and content review comments): Rita Golfieri, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Italy.

Cochrane Diagnostic Test Accuracy Editorial Team

Sign-off Editor (DTAR methods): Ingrid Arevalo-Rodriguez, Clinical Biostatistics Unit, Spain, and Screening and Diagnostic Tests Methods Group, UK

Peer-reviewers (general methods): Keele University, UK; (statistical comments) Enzo Cerullo, University of Leicester, UK; (peer review of search strategies): Kate Misso, Information Retrieval Methods Group, UK.

Copy Editor (copy editing and production): Anne Lawson, Copy Edit Support, Cochrane.

REFERENCES

References to studies included in this review

Besa 2017 {published data only}

Besa C, Lewis S, Pandharipande PV, Chhatwal J, Kamath A, Cooper N, et al. Hepatocellular carcinoma detection: diagnostic performance of a simulated abbreviated MRI protocol combining diffusion-weighted and T1-weighted imaging at the delayed phase post gadoteric acid. *Abdominal Radiology* 2017;**42**(1):179-90.

Bhartia 2003 {published data only}

Bhartia B, Ward J, Guthrie JA, Robinson PJ. Hepatocellular carcinoma in cirrhotic livers: double-contrast thin-section MR imaging with pathologic correlation of explanted tissue. *American Journal of Roentgenology* 2003;**180**(3):577-84.

Born 1998 {published data only}

Born M, Layer G, Kreft B, Schwarz N, Schild H. MRI, CT and CT arterial portography in the diagnosis of malignant liver tumors in liver cirrhosis [MRT, CT und CTAP in der Diagnostik maligner Lebertumoren bei Leberzirrhose]. *RöFo: Fortschritte auf dem Gebiete der Röntgenstrahlen und der Nuklearmedizin* 1998;**168**(6):567-72.

Brunsing 2019 {published data only}

Brunsing RL, Chen DH, Schlein A, Wolfson T, Gamst A, Mamidipalli A, et al. Gadoteric acid-enhanced abbreviated MRI for hepatocellular carcinoma surveillance: preliminary experience. *Radiology. Imaging Cancer* 2019;**1**(2):e190010.

Darnell 2021 {published data only}

Darnell A, Rimola J, Belmonte E, Ripoll E, Garcia-Criado Á, Caparroz C, et al. Evaluation of LI-RADS 3 category by magnetic resonance in US-detected nodules \leq 2 cm in cirrhotic patients. *European Radiology* 2021;**31**(7):4794-803.

de Lédínghen 2002 {published data only}

de Lédínghen V, Laharie D, Lecesne R, Le Bail B, Winnock M, Bernard PH, et al. Detection of nodules in liver cirrhosis: spiral computed tomography or magnetic resonance imaging? A prospective study of 88 nodules in 34 patients. *European Journal of Gastroenterology and Hepatology* 2002;**14**(2):159-65.

Demirtas 2020 {published data only}

Demirtas CO, Gunduz F, Tuney D, Baltacioglu F, Kani HT, Bugdayci O, et al. Annual contrast-enhanced magnetic resonance imaging is highly effective in the surveillance of hepatocellular carcinoma among cirrhotic patients. *European Journal of Gastroenterology and Hepatology* 2020;**32**(4):517-23.

Di Carlo 2012 {published data only}

Di Carlo I, Loria F, Loria G, Basile S, Randazzo D, Frosina L, et al. Diagnosis of small hepatic nodules in cirrhotic liver: role of contrast-enhanced ultrasound and contrast enhanced magnetic resonance imaging in the validation of the non invasive diagnostic criteria for hepato-cellular carcinoma. *HPB: the Official Journal of the International Hepato Pancreato Biliary Association* 2012;**14**(Suppl 2):435.

Dumitrescu 2013 {published data only}

Dumitrescu CI, Gheonea IA, Săndulescu L, Surlin V, Săftoiu A, Dumitrescu D. Contrast enhanced ultrasound and magnetic resonance imaging in hepatocellular carcinoma diagnosis. *Medical Ultrasonography* 2013;**15**(4):261-7.

Giorgio 2007 {published data only}

Giorgio A, De Stefano G, Coppola C, Ferraioli G, Esposito V, Di Sarno A, et al. Contrast-enhanced sonography in the characterization of small hepatocellular carcinomas in cirrhotic patients: comparison with contrast-enhanced ultrafast magnetic resonance imaging. *Anticancer Research* 2007;**27**(6C):4263-9.

Golfieri 2009 {published data only}

Golfieri R, Marini E, Bazzocchi A, Fusco F, Trevisani F, Sama C, et al. Small (\leq 3 cm) hepatocellular carcinoma in cirrhosis: the role of double contrast agents in MR imaging vs. multidetector-row CT [Ruolo della RM a doppio contrasto rispetto alla TC multidetettore nelladiagnosi del piccolo epatocarcinoma (\leq 3 cm) su cirrosi]. *Radiologia Medica* 2009;**14**(8):1239-66.

Hanna 2008 {published data only}

Hanna RF, Kased N, Kwan SW, Gamst AC, Santosa AC, Hassanein T, et al. Double-contrast MRI for accurate staging of hepatocellular carcinoma in patients with cirrhosis. *American Journal of Roentgenology* 2008;**190**(1):47-57.

Hwang 2014 {published data only}

Hwang J, Kim YK, Kim JM, Lee WJ, Choi D, Hong SS. Pretransplant diagnosis of hepatocellular carcinoma by gadoteric acid-enhanced and diffusion-weighted magnetic resonance imaging. *Liver Transplantation* 2014;**20**(12):1436-46.

Khatri 2020 {published data only}

Khatri G, Pedrosa I, Ananthkrishnan L, de Leon AD, Fetzer DT, Leyendecker J, et al. Abbreviated-protocol screening MRI vs. complete-protocol diagnostic MRI for detection of hepatocellular carcinoma in patients with cirrhosis: an equivalence study using LI-RADS v2018. *Journal of Magnetic Resonance Imaging: JMIR* 2020;**51**(2):415-25.

Kim 2017 {published data only}

Kim SY, An J, Lim YS, Han S, Lee JY, Byun JH, et al. MRI with liver-specific contrast for surveillance of patients with cirrhosis at high risk of hepatocellular carcinoma. *JAMA Oncology* 2017;**3**(4):456-63.

Kim 2020 {published data only}

Kim JS, Lee JK, Baek SY, Yun HI. Diagnostic performance of a minimized protocol of non-contrast MRI for hepatocellular carcinoma surveillance. *Abdominal Radiology* 2020;**45**(1):211-9.

Lauenstein 2007 {published data only}

Lauenstein TC, Salman K, Morreira R, Heffron T, Spivey JR, Martinez E, et al. Gadolinium-enhanced MRI for tumor surveillance before liver transplantation: center-based experience. *American Journal of Roentgenology* 2007;**189**(3):663-70.

Libbrecht 2002 {published data only}

Libbrecht L, Bielen D, Verslype C, Vanbeckevoort D, Pirenne J, Nevens F, et al. Focal lesions in cirrhotic explant livers: pathological evaluation and accuracy of pretransplantation imaging examinations. *Liver Transplantation* 2002;**8**(9):749-61.

Lin 2016 {published data only}

Lin MT, Wang CC, Cheng YF, Eng HL, Yen YH, Tsai MC, et al. Comprehensive comparison of multiple-detector computed tomography and dynamic magnetic resonance imaging in the diagnosis of hepatocellular carcinoma with varying degrees of fibrosis. *PLoS One* 2016;**11**(11):e0166157.

Maiwald 2014 {published data only}

Maiwald B, Lobsien D, Kahn T, Stumpp P. Is 3-Tesla Gd-EOB-DTPA-enhanced MRI with diffusion-weighted imaging superior to 64-slice contrast-enhanced CT for the diagnosis of hepatocellular carcinoma? *PLoS One* 2014;**9**(11):e111935.

Marks 2015 {published data only}

Marks RM, Ryan A, Heba ER, Tang A, Wolfson TJ, Gamst AC, et al. Diagnostic per-patient accuracy of an abbreviated hepatobiliary phase gadoteric acid-enhanced MRI for hepatocellular carcinoma surveillance. *American Journal of Roentgenology* 2015;**204**(3):527-35.

Marrero 2005 {published data only}

Marrero JA, Hussain HK, Nghiem HV, Umar R, Fontana RJ, Lok AS. Improving the prediction of hepatocellular carcinoma in cirrhotic patients with an arterially-enhancing liver mass. *Liver Transplantation* 2005;**11**(3):281-9.

McNamara 2018 {published data only}

McNamara MM, Thomas JV, Alexander LF, Little MD, Bolus DN, Li YE, et al. Diffusion-weighted MRI as a screening tool for hepatocellular carcinoma in cirrhotic livers: correlation with explant data – a pilot study. *Abdominal Radiology* 2018;**43**(10):2686-92.

Min 2018a {published data only}

Min JH, Kim YK, Sinn DH, Choi SY, Jeong WK, Lee WJ, et al. Adding ancillary features to enhancement patterns of hepatocellular carcinoma on gadoteric acid-enhanced magnetic resonance imaging improves diagnostic performance. *Abdominal Radiology* 2018;**43**(9):2309-20.

Sangiovanni 2010 {published data only}

Sangiovanni A, Manini MA, Iavarone M, Romeo R, Forzenigo LV, Fraquelli M, et al. The diagnostic and economic impact of contrast imaging techniques in the diagnosis of small hepatocellular carcinoma in cirrhosis. *Gut* 2010;**59**(5):638-44.

Seçil 2008 {published data only}

Seçil M, Obuz F, Altay C, Gencel O, Iğci E, Sağol O, et al. The role of dynamic subtraction MRI in detection of hepatocellular carcinoma. *Diagnostic and Interventional Radiology (Ankara, Turkey)* 2008;**14**(4):200-4.

Sersté 2012 {published data only}

Sersté T, Barrau V, Ozenne V, Vullierme MP, Bedossa P, Farges O, et al. Accuracy and disagreement of computed tomography

and magnetic resonance imaging for the diagnosis of small hepatocellular carcinoma and dysplastic nodules: role of biopsy. *Hepatology* 2012;**55**(3):800-6.

Shin 2017 {published data only}

Shin SK, Kim YS, Choi SJ, Shim YS, Jung DH, Kwon OS, et al. Characterization of small (≤ 3 cm) hepatic lesions with atypical enhancement feature and hypointensity in hepatobiliary phase of gadoteric acid-enhanced MRI in cirrhosis: a STARD-compliant article. *Medicine* 2017;**96**(29):e7278.

Sutherland 2017 {published data only}

Sutherland T, Watts J, Ryan M, Galvin A, Temple F, Vuong J, et al. Diffusion-weighted MRI for hepatocellular carcinoma screening in chronic liver disease: direct comparison with ultrasound screening. *Journal of Medical Imaging and Radiation Oncology* 2017;**61**(1):34-9.

Teefey 2003 {published data only}

Teefey SA, Hildeboldt CC, Dehdashti F, Siegel BA, Peters MG, Heiken JP, et al. Detection of primary hepatic malignancy in liver transplant candidates: prospective comparison of CT, MR imaging, US, and PET. *Radiology* 2003;**226**(2):533-42.

Vietti Violi 2020 {published data only}

Vietti Violi N, Lewis S, Liao J, Hulkower M, Hernandez-Meza G, Smith K, et al. Gadoteric acid-enhanced abbreviated MRI is highly accurate for hepatocellular carcinoma screening. *European Radiology* 2020;**30**(11):6003-13.

Villacastin Ruiz 2016 {published data only}

Villacastin Ruiz E, Caro-Patón Gómez A, Calero Aguilar H, Pérez Saborido B, García Pajares F, Sánchez Antolín G, et al. Review of imaging techniques in the diagnosis of hepatocellular carcinoma in patients who require a liver transplant. *European Journal of Gastroenterology and Hepatology* 2016;**28**(4):412-20.

Wu 2020 {published data only}

Wu Y, Huang L, Li B, Li H. The diagnostic value of Gd-EOB-DTPA-enhanced MRI scans in small hepatocellular carcinoma in patients with liver cirrhosis. *International Journal of Clinical and Experimental Medicine* 2020;**13**(9):6909-15.

Yu 2011 {published data only}

Yu NC, Chaudhari V, Raman SS, Lassman C, Tong MJ, Busuttill RW, et al. CT and MRI improve detection of hepatocellular carcinoma, compared with ultrasound alone, in patients with cirrhosis. *Clinical Gastroenterology and Hepatology* 2011;**9**(2):161-7.

References to studies excluded from this review
Abdel Kader 2017 {published data only}

Abdel Kader M, Abdel Ghani HS, Saad ZM, Abdalla NH. To what extent the DW-MRI and ADC value can be used in assessment of hepatic focal lesions in cirrhotic patients. *Egyptian Journal of Radiology and Nuclear Medicine* 2017;**48**(4):825-37.

Akhtar 2020 {published data only}

Akhtar S, Hussain M, Ali S, Maqsood S, Akram S, Abba N. Comparison of positive predictive value of multiphase dynamic contrast enhanced MRI with dynamic contrast enhanced CT for the detection of hepatocellular carcinoma. *Pakistan Journal of Medical and Health Sciences* 2020;**14**(3):562-4.

Araki 2000 {published data only}

Araki T. SPIO-MRI in the detection of hepatocellular carcinoma. *Journal of Gastroenterology* 2000;**35**:874-6.

Aubé 2017 {published data only}

Aubé C, Oberti F, Lonjon J, Pageaux G, Seror O, N'Kontchou G, et al. EASL and AASLD recommendations for the diagnosis of HCC to the test of daily practice. *Liver International* 2017;**37**(10):1515-25.

Ayuso 2019 {published data only}

Ayuso C, Forner A, Darnell A, Rimola J, García-Criado Á, Bianchi L, et al. Prospective evaluation of gadoxetic acid magnetic resonance for the diagnosis of hepatocellular carcinoma in newly detected nodules ≤ 2 cm in cirrhosis. *Liver International* 2019;**39**(7):1281-91.

Baird 2013 {published data only}

Baird AJ, Amos GJ, Saad NF, Benson MD. Retrospective audit to determine the diagnostic accuracy of Primovist-enhanced MRI in the detection of hepatocellular carcinoma in cirrhosis with explant histopathology correlation. *Journal of Medical Imaging and Radiation Oncology* 2013;**57**(3):314-20.

Barabino 2021 {published data only}

Barabino M, Gurgitano M, Fochesato C, Angileri SA, Franceschelli G, Santambrogio R, et al. LI-RADS to categorize liver nodules in patients at risk of HCC: tool or a gadget in daily practice? *La Radiologia Medica* 2021;**126**(1):5-13.

Bartolozzi 2013 {published data only}

Bartolozzi C, Battaglia V, Bargellini I, Bozzi E, Campani D, Pollina LE, et al. Contrast-enhanced magnetic resonance imaging of 102 nodules in cirrhosis: correlation with histological findings on explanted livers. *Abdominal Imaging* 2013;**38**(2):290-6.

Basha 2018 {published data only}

Basha MA, Al Azzazy MZ, Ahmed AF, Yousef HY, Shehata SM, El Sammak DA, et al. Does a combined CT and MRI protocol enhance the diagnostic efficacy of LI-RADS in the categorization of hepatic observations? A prospective comparative study. *European Radiology* 2018;**28**(6):2592-603.

Becker-Weidman 2011 {published data only}

Becker-Weidman DJ, Kalb B, Sharma P, Kitajima HD, Lurie CR, Chen Z, et al. Hepatocellular carcinoma lesion characterization: single-institution clinical performance review of multiphase gadolinium-enhanced MR imaging – comparison to prior same-center results after MR systems improvements. *Radiology* 2011;**261**(3):824-33.

Blondin 2011 {published data only}

Blondin D, Erhardt A, Crynen K, Sagir A, Scherer A, Kröpil P, et al. Diagnosis of focal liver lesions in cirrhotic patients: comparison of contrast-enhanced ultrasound using sulphur hexafluoride (SF6) microbubbles and MRI using Gd-EOB-DTPA [Vergleich der kontrastverstärkten Sonografie und der MRT mit Gd-EOB-DTPA zur Diagnostik fokaler Leberläsionen bei Patienten mit Leberzirrhose]. *Zeitschrift für Gastroenterologie* 2011;**49**(1):23-9.

Burrel 2003 {published data only}

Burrel M, Llovet JM, Ayuso C, Iglesias C, Sala M, Miquel R, et al. MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation. *Hepatology* 2003;**38**(4):1034-42.

Choi 2008 {published data only}

Choi SH, Lee JM, Yu NC, Suh KS, Jang JJ, Kim SH, et al. Hepatocellular carcinoma in liver transplantation candidates: detection with gadobenate dimeglumine-enhanced MRI. *American Journal of Roentgenology* 2008;**191**(2):529-36.

Chou 2014 {published data only}

Chou CT, Wu WP, Chen CB, Su WW, Chen RC, Chen YL. The utility of gadoxetic acid-enhanced MR imaging to characterize atypical cirrhotic nodules detected on dynamic CT images. *PLoS One* 2014;**9**(10):e107869.

Clarke 2021 {published data only}

Clarke CG, Albazaz R, Smith CR, Rowe I, Treanor D, Wyatt JI, et al. Comparison of LI-RADS with other non-invasive liver MRI criteria and radiological opinion for diagnosing hepatocellular carcinoma in cirrhotic livers using gadoxetic acid with histopathological explant correlation. *Clinical Radiology* 2021;**76**(5):333-41.

Compagnon 2008 {published data only}

Compagnon P, Grandadam S, Lorho R, Turlin B, Camus C, Jianrong Y, et al. Liver transplantation for hepatocellular carcinoma without preoperative tumor biopsy. *Transplantation* 2008;**86**(8):1068-76.

Debees 2016 {published data only}

Debees NL, Sherif MF, Yones SG, Ahmad AH. Assessment of hepatic focal lesions on top of cirrhotic liver using dynamic and diffusion weighted magnetic resonance imaging. *Egyptian Journal of Radiology and Nuclear Medicine* 2016;**47**:1221-30.

Ercalli 2013 {published data only}

Ercalli U, Tokgoz O, Erdem Z, Erdem O. The role of magnetic resonance imaging with dynamic subtraction in determining liver lesions with cirrhosis [Sirozlu hastalarda karaciger lezyonlarinin degerlendirilmesinde dinamik cikarmali manyetik rezonans goruntulemenin rolu]. *Turkish Journal of Geriatrics* 2013;**16**(1):31-7.

Faletti 2015 {published data only}

Faletti R, Cassinis MC, Fonio P, Bergamasco L, Pavan LJ, Rapellino A, et al. Multiparametric Gd-EOB-DTPA magnetic resonance in diagnosis of HCC: dynamic study, hepatobiliary phase, and diffusion-weighted imaging compared to histology

after orthotopic liver transplantation. *Abdominal Imaging* 2015;**40**(1):46-55.

Fischer 2015 {published data only}

Fischer MA, Raptis DA, Donati OF, Hunziker R, Schade E, Sotiropoulos GC, et al. MR imaging features for improved diagnosis of hepatocellular carcinoma in the non-cirrhotic liver: multi-center evaluation. *European Journal of Radiology* 2015;**84**(10):1879-87.

Fornier 2008 {published data only}

Fornier A, Vilana R, Ayuso C, Bianchi L, Solé M, Ayuso JR, et al. Diagnosis of hepatic nodules 20 mm or smaller in cirrhosis: prospective validation of the noninvasive diagnostic criteria for hepatocellular carcinoma. *Hepatology* 2008;**47**(1):97-104.

Fornier 2018b {published data only}

Fornier A, Darnell A, Ayuso C, Rimola J, García-Criado MA, Vilana R, et al. Prospective evaluation of gadoteric-acid MR for the non-invasive diagnosis of hepatocellular carcinoma in newly detected nodules ≤ 2 cm in cirrhotic patients. *Journal of Hepatology* 2018;**68**:S424.

Fornier 2019 {published data only}

Fornier A, Darnell A, Caparroz C, Rimola J, Garcia-Criado MA, Belmonte E. Evaluation of LI-RADS v2018 by magnetic resonance in US-detected nodules < 2 cm in cirrhotics. *Journal of Hepatology* 2019;**70**(Suppl 1):e73.

Giorgio 2005 {published data only}

Giorgio A, Ferraioli G, de Stefano G, Coppola C, Di Sarno A, Del Visco L. Contrast-enhanced sonography in the characterization of small hepatocellular carcinomas in cirrhotic patients: comparison with contrast-enhanced ultrafast magnetic resonance imaging. *Journal of Hepatology* 2005;**42**:93-4.

Guiu 2008 {published data only}

Guiu B, Loffroy R, Ben Salem D, Lepage C, Guiu S, Aho S, et al. Combined SPIO-gadolinium magnetic resonance imaging in cirrhotic patients: negative predictive value and role in screening for hepatocellular carcinoma. *Abdominal Imaging* 2008;**33**(5):520-8.

Guo 2012 {published data only}

Guo L, Liang C, Yu T, Wang G, Li N, Sun H, et al. 3T MRI of hepatocellular carcinomas in patients with cirrhosis: does T2-weighted imaging provide added value? *Clinical Radiology* 2012;**67**(4):319-28.

Hanna 2014 {published data only}

Hanna RF, Ward TJ, Chow DS, Lagana SM, Moreira RK, Emond JC, et al. An evaluation of the sensitivity of MRI at detecting hepatocellular carcinoma in cirrhotic patients utilizing an explant reference standard. *Clinical Imaging* 2014;**38**(5):693-7.

Hardie 2010 {published data only}

Hardie AD, Romano PB. The use of T2*-weighted multi-echo GRE imaging as a novel method to diagnose hepatocellular

carcinoma compared with gadolinium-enhanced MRI: a feasibility study. *Magnetic Resonance Imaging* 2010;**28**(2):281-5.

Hardie 2011a {published data only}

Hardie AD, Kizziah MK, Rissing MS. Can the patient with cirrhosis be imaged for hepatocellular carcinoma without gadolinium?: comparison of combined T2-weighted, T2*-weighted, and diffusion-weighted MRI with gadolinium-enhanced MRI using liver explantation standard. *Journal of Computer Assisted Tomography* 2011;**35**(6):711-5.

Hardie 2011b {published data only}

Hardie AD, Nance JW, Boulter DJ, Kizziah MK. Assessment of the diagnostic accuracy of T2*-weighted MR imaging for identifying hepatocellular carcinoma with liver explant correlation. *European Journal of Radiology* 2011;**80**(3):e249-52.

Hecht 2006 {published data only}

Hecht EM, Holland AE, Israel GM, Hahn WY, Kim DC, West AB, et al. Hepatocellular carcinoma in the cirrhotic liver: gadolinium-enhanced 3D T1-weighted MR imaging as a stand-alone sequence for diagnosis. *Radiology* 2006;**239**(2):438-47.

Heilmaier 2008 {published data only}

Heilmaier C, Sutter R, Lutz AM, Seifert B, Willmann JK. Dynamic MRI of the liver with parallel acquisition technique: characterization of focal liver lesions and analysis of the hepatic vasculature in a single MRI session [Dynamische MRT der Leber mit paralleler Akquisitionstechnik: Charakterisierung fokaler Leberläsionen und Analyse des Gefäßstatus in einem Untersuchungsgang]. *RöFo: Fortschritte auf dem Gebiete der Röntgenstrahlen und der Nuklearmedizin* 2008;**180**(5):440-8.

Heilmaier 2009 {published data only}

Heilmaier C, Lutz AM, Bolog N, Weishaupt D, Seifert B, Willmann JK. Focal liver lesions: detection and characterization at double-contrast liver MR imaging with ferucarbotran and gadobutrol versus single-contrast liver MR imaging. *Radiology* 2009;**253**(3):724-33.

Ichikawa 2021 {published data only}

Ichikawa S, Motosugi U, Morisaka H, Kozaka K, Goshima S, Ichikawa T. Optimal combination of features on gadoterate disodium-enhanced MR imaging for non-invasive differential diagnosis of hepatocellular carcinoma: the JAMP-HCC study. *Magnetic Resonance in Medical Sciences* 2021;**20**(1):47-59.

Kang 2012 {published data only}

Kang HT, Shin HD, Kim SB, Song IH. Diagnostic validity for sequential approach of dynamic image modalities in high-risk patients for hepatocellular carcinoma. *Journal of Hepatology* 2012;**56**:S295-6.

Kim 2011 {published data only}

Kim TK, Lee KH, Jang HJ, Haider MA, Jacks LM, Menezes RJ, et al. Analysis of gadobenate dimeglumine-enhanced MR findings for characterizing small (1-2-cm) hepatic nodules in patients at high risk for hepatocellular carcinoma. *Radiology* 2011;**259**(3):730-8.

Kim 2012 {published data only}

Kim DJ, Yu JS, Kim JH, Chung JJ, Kim KW. Small hypervascular hepatocellular carcinomas: value of diffusion-weighted imaging compared with "washout" appearance on dynamic MRI. *British Journal of Radiology* 2012;**85**(1018):e879-86.

Kim 2021 {published data only}

Kim YY, Kim YK, Min JH, Cha DI, Kim JM, Choi GS, et al. Intraindividual comparison of hepatocellular carcinoma washout between MRIs with hepatobiliary and extracellular contrast agents. *Korean Journal of Radiology* 2021;**22**(5):725-34.

Kondo 2000 {published data only}

Kondo H, Kanematsu M, Hoshi H, Murakami T, Kim T, Hori M, et al. Preoperative detection of malignant hepatic tumors: comparison of combined methods of MR imaging with combined methods of CT. *American Journal of Roentgenology* 2000;**174**(4):947-54.

Lee 2021 {published data only}

Lee SM, Lee JM, Ahn SJ, Kang HJ, Yang HK, Yoon JH. Diagnostic performance of 2018 KLCA-NCC practice guideline for hepatocellular carcinoma on gadoxetic acid-enhanced MRI in patients with chronic hepatitis B or cirrhosis: comparison with LI-RADS version 2018. *Korean Journal of Radiology* 2021;**22**(7):1066-76.

Li 2015 {published data only}

Li R, Zeng M, Qiang J, Rao S, Chen C, Chen L, et al. Value of MR susceptibility weighted imaging for characterization of small hepatocellular carcinoma in cirrhotic livers. *Chinese Journal of Radiology* 2015;**49**(7):520-4.

Li 2018 {published data only}

Li J, Li X, Weng J, Lei L, Gong J, Wang J, et al. Gd-EOB-DTPA dynamic contrast-enhanced magnetic resonance imaging is more effective than enhanced 64-slice CT for the detection of small lesions in patients with hepatocellular carcinoma. *Medicine* 2018;**97**(52):e13964.

Li 2021 {published data only}

Li S, Zhou L, Chen R, Chen Y, Niu Z, Qian L, et al. Diagnostic efficacy of contrast-enhanced ultrasound versus MRI Liver Imaging Reporting and Data System (LI-RADS) for categorising hepatic observations in patients at risk of hepatocellular carcinoma. *Clinical Radiology* 2021;**76**(2):161.e1-e10.

Macarini 2006 {published data only}

Macarini L, Marini S, Milillo P, Vinci R, Ettorre GC. Double-contrast MRI (DC-MRI) in the study of the cirrhotic liver: utility of administering Gd-DTPA as a complement to examinations in which SPIO liver uptake and distribution alterations (SPIO-LUDA) are present and in the identification and characterisation of focal lesions. [RM epatica a doppio contrasto (RM-DC) nello studio del fegato cirrotico: utilità della somministrazione di Gd-DTPA come completamento di esami in cui siano presenti alterazioni di captazione e distribuzione epatiche (ACDE) del MdC SPIOe nell'identificazione e caratterizzazione delle lesioni focali]. *La Radiologia Medica* 2006;**111**(8):1087-102.

Macarini 2009 {published data only}

Macarini L, Milillo P, Cascavilla A, Scalzo G, Stoppino L, Vinci R, et al. MR characterisation of dysplastic nodules and hepatocarcinoma in the cirrhotic liver with hepatospecific superparamagnetic contrast agents: pathological correlation in explanted livers [Caratterizzazione RM dei noduli displasici e dell'epatocarcinoma nel fegato cirrotico con mezzi di contrasto superparamagnetici epatospecifici: correlazione anatomico-patologica nei fegati espiantati]. *La Radiologia Medica* 2009;**114**(8):1267-82.

Mann 2001 {published data only}

Mann GN, Marx HF, Lai LL, Wagman LD. Clinical and cost effectiveness of a new hepatocellular MRI contrast agent, mangafodipir trisodium, in the preoperative assessment of liver resectability. *Annals of Surgical Oncology* 2001;**8**(7):573-9.

Matsuo 2001 {published data only}

Matsuo M, Kanematsu M, Itoh K, Ito K, Maetani Y, Kondo H, et al. Detection of malignant hepatic tumors: comparison of gadolinium- and ferumoxide-enhanced MR imaging. *American Journal of Roentgenology* 2001;**177**(3):637-43.

Maurea 2014 {published data only}

Maurea S, Mainenti PP, Tambasco A, Imbriaco M, Mollica C, Laccetti E, et al. Diagnostic accuracy of MR imaging to identify and characterize focal liver lesions: comparison between gadolinium and superparamagnetic iron oxide contrast media. *Quantitative Imaging in Medicine and Surgery* 2014;**4**(3):181-9.

Mehra 2012 {published data only}

Mehra S. Hepatic imaging – role of contrast enhanced MR in focal hepatic masses. *Journal International Medical Sciences Academy* 2012;**25**(4):287.

Miller 2008 {published data only}

Miller JC, Hahn PF, Chung RT, Thrall JH, Lee SI. Screening for hepatocellular carcinoma in cirrhotic patients. *Journal of the American College of Radiology* 2008;**5**(9):1012-4.

Min 2018b {published data only}

Min JH, Kim JM, Kim YK, Kang TW, Lee SJ, Choi GS, et al. Prospective intraindividual comparison of magnetic resonance imaging with gadoxetic acid and extracellular contrast for diagnosis of hepatocellular carcinomas using the Liver Imaging Reporting and Data System. *Hepatology* 2018;**68**(6):2254-66.

Mita 2010 {published data only}

Mita K, Kim SR, Kudo M, Imoto S, Nakajima T, Ando K, et al. Diagnostic sensitivity of imaging modalities for hepatocellular carcinoma smaller than 2 cm. *World Journal of Gastroenterology* 2010;**16**(33):4187-92.

Moon 2018 {published data only}

Moon JY, Kim SH, Choi SY, Hwang JA, Lee JE, Lee J. Differentiating malignant from benign hyperintense nodules on unenhanced T1-weighted images in patients with chronic liver disease: using gadoxetic acid-enhanced and diffusion-weighted MR imaging. *Japanese Journal of Radiology* 2018;**36**(8):489-99.

Mori 2002 {published data only}

Mori K, Scheidler J, Helmberger T, Holzknrecht N, Schauer R, Schirren CA, et al. Detection of malignant hepatic lesions before orthotopic liver transplantation: accuracy of ferumoxides-enhanced MR imaging. *American Journal of Roentgenology* 2002;**179**(4):1045-51.

Obuz 2006 {published data only}

Obuz F, Oksüzler M, Seçil M, Sağol O, Karademir S, Astarcioglu H. Efficiency of MR imaging in the detection of malignant liver lesions. *Diagnostic and Interventional Radiology (Ankara, Turkey)* 2006;**12**(1):17-21.

Pahade 2016 {published data only}

Pahade JK, Juice D, Staib L, Israel G, Cornfeld D, Mitchell K, et al. Is there an added value of a hepatobiliary phase with gadoxetate disodium following conventional MRI with an extracellular gadolinium agent in a single imaging session for detection of primary hepatic malignancies? *Abdominal Radiology* 2016;**41**(7):1270-84.

Paisant 2020 {published data only}

Paisant A, Vilgrain V, Riou J, Oberti F, Sutter O, Laurent V, et al. Comparison of extracellular and hepatobiliary MR contrast agents for the diagnosis of small HCCs. *Journal of Hepatology* 2020;**72**(5):937-45.

Park 2012 {published data only}

Park MS, Kim S, Patel J, Hajdu CH, Do RK, Mannelli L, et al. Hepatocellular carcinoma: detection with diffusion-weighted versus contrast-enhanced magnetic resonance imaging in pretransplant patients. *Hepatology* 2012;**56**(1):140-8.

Park 2020 {published data only}

Park HJ, Jang HY, Kim SY, Lee SJ, Won HJ, Byun JH, et al. Non-enhanced magnetic resonance imaging as a surveillance tool for hepatocellular carcinoma: comparison with ultrasound. *Journal of Hepatology* 2020;**72**(4):718-24.

Phongkitkarun 2013 {published data only}

Phongkitkarun S, Limsamutpetch K, Tannaphai P, Jatchavala J. Added value of hepatobiliary phase gadoxetic acid-enhanced MRI for diagnosing hepatocellular carcinoma in high-risk patients. *World Journal of Gastroenterology* 2013;**19**(45):8357-65.

Puig 1997 {published data only}

Puig J, Martin J, Donoso L, Falco J, Rue M. Comparison between biphasic helical CT and dynamic gadolinium-enhanced MR in the detection and characterization of focal hepatic lesions in cirrhotic patients. *Radiologia* 1997;**39**(9):617-23.

Qayyum 2006 {published data only}

Qayyum A, Thoeni RF, Coakley FV, Lu Y, Guay JP, Ferrell LD. Detection of hepatocellular carcinoma by ferumoxides-enhanced MR imaging in cirrhosis: incremental value of dynamic gadolinium-enhancement. *Journal of Magnetic Resonance Imaging* 2006;**23**(1):17-22.

Rimola 2012 {published data only}

Rimola J, Forner A, Tremosini S, Reig M, Vilana R, Bianchi L, et al. Non-invasive diagnosis of hepatocellular carcinoma \leq 2 cm in cirrhosis. Diagnostic accuracy assessing fat, capsule and signal intensity at dynamic MRI. *Journal of Hepatology* 2012;**56**(6):1317-23.

Rofsky 1993 {published data only}

Rofsky NM, Weinreb JC, Bernardino ME, Young SW, Lee JK, Noz ME. Hepatocellular tumors: characterization with Mn-DPDP-enhanced MR imaging. *Radiology* 1993;**188**(1):53-9.

Samoylova 2018 {published data only}

Samoylova ML, Mehta N, Roberts JP, Yao FY. Predictors of ultrasound failure to detect hepatocellular carcinoma. *Liver Transplantation* 2018;**24**(9):1171-7.

Serste 2010 {published data only}

Serste T, Barrau V, Ozenne V, Vullierme MP, Goutte N, Valla D, et al. Diagnosis of hepatic nodules < 20 mm in chronic liver disease. Histologic validation of CT and MRI criteria. *Hepatology* 2010;**52**(Suppl 1):1181A.

Shankar 2016 {published data only}

Shankar S, Kalra N, Bhatia A, Srinivasan R, Singh P, Dhiman RK, et al. Role of diffusion weighted imaging (DWI) for hepatocellular carcinoma (HCC) detection and its grading on 3T MRI: a prospective study. *Journal of Clinical and Experimental Hepatology* 2016;**6**(4):303-10.

Simon 2005 {published data only}

Simon G, Link TM, Wörtler K, Doebereiner F, Schulte-Frohlinde E, Daldrup-Link H, et al. Detection of hepatocellular carcinoma: comparison of Gd-DTPA- and ferumoxides-enhanced MR imaging. *European Radiology* 2005;**15**(5):895-903.

Snowberger 2007 {published data only}

Snowberger N, Chinnakotla S, Lepe RM, Peattie J, Goldstein R, Klintmalm GB, et al. Alpha fetoprotein, ultrasound, computerized tomography and magnetic resonance imaging for detection of hepatocellular carcinoma in patients with advanced cirrhosis. *Alimentary Pharmacology & Therapeutics* 2007;**26**(9):1187-94.

Stocker 2018 {published data only}

Stocker D, Marquez HP, Wagner MW, Raptis DA, Clavien PA, Boss A, et al. MRI texture analysis for differentiation of malignant and benign hepatocellular tumors in the non-cirrhotic liver. *Heliyon* 2018;**4**(11):e00987.

Suárez-Muñoz 2006 {published data only}

Suárez-Muñoz MA, Leiva-Vera MC, Santoyo-Santoyo J, Fernández-Aguilar JL, Pérez-Daga JA, Sánchez-Pérez B, et al. Detection of neoplastic lesions in cirrhotic patients waiting for liver transplantation [Detección de lesiones neoplásicas en pacientes cirróticos candidatos a trasplante hepático]. *Cirugía Española* 2006;**80**(3):157-61.

Sugimoto 2015 {published data only}

Sugimoto K, Kim SR, Imoto S, Tohyama M, Kim SK, Matsuoka T, et al. Characteristics of hypovascular versus hypervascular

well-differentiated hepatocellular carcinoma smaller than 2 cm – focus on tumor size, markers and imaging detectability. *Digestive Diseases (Basel, Switzerland)* 2015;**33**(6):721-7.

Teerasamit 2017 {published data only}

Teerasamit W, Tongdee R, Yodying J. Diagnostic performance of gadoteric acid-enhanced MR imaging in the diagnosis of hepatocellular carcinoma in cirrhotic liver. *Journal of the Medical Association in Thailand* 2017;**100**(8):918-26.

Timofte 2016 {published data only}

Timofte O, Stefanescu G, Gologan E, Balan G. Liver nodules in cirrhosis: do we need contrast-enhanced ultrasonography? *Journal of Gastrointestinal and Liver Diseases* 2016;**25**(Suppl 2):409-14.

Tsang 2011 {published data only}

Tsang LL, Chen CL, Huang TL, Chen TY, Ou HY, Eng HL, et al. Superparamagnetic iron oxide-enhanced magnetic resonance for tumor surveillance in cirrhotic liver before liver transplantation with explanted liver correlation. *Transplantation Proceedings* 2011;**43**(5):1674-7.

Ueda 1995 {published data only}

Ueda K, Kitagawa K, Kadoya M, Matsui O, Takashima T, Yamahana T. Detection of hypervascular hepatocellular carcinoma by using spiral volumetric CT: comparison of US and MR imaging. *Abdominal Imaging* 1995;**20**(6):547-54.

Vandecaveye 2009 {published data only}

Vandecaveye V, De Keyser F, Verslype C, Op de Beeck K, Komuta M, Topal B, et al. Diffusion-weighted MRI provides additional value to conventional dynamic contrast-enhanced MRI for detection of hepatocellular carcinoma. *European Radiology* 2009;**19**(10):2456-66.

van Wettre 2019 {published data only}

van Wettre M, Purcell Y, Bruno O, Payancé A, Plessier A, Rautou PE, et al. Low specificity of washout to diagnose hepatocellular carcinoma in nodules showing arterial hyperenhancement in patients with Budd-Chiari syndrome. *Journal of Hepatology* 2019;**70**(6):1123-32.

Xu 2009 {published data only}

Xu PJ, Yan FH, Wang JH, Lin J, Ji Y. Added value of breathhold diffusion-weighted MRI in detection of small hepatocellular carcinoma lesions compared with dynamic contrast-enhanced MRI alone using receiver operating characteristic curve analysis. *Journal of Magnetic Resonance Imaging* 2009;**29**(2):341-9.

Xu 2013 {published data only}

Xu Y, Lu X, Mao Y, Sang X, Zhao H, Du S, et al. Clinical diagnosis and treatment of alpha-fetoprotein-negative small hepatic lesions. *Chinese Journal of Cancer Research* 2013;**25**(4):382-8.

Yoo 2008 {published data only}

Yoo HJ, Lee JM, Lee MW, Kim SJ, Lee JY, Han JK, et al. Hepatocellular carcinoma in cirrhotic liver: double-contrast-enhanced, high-resolution 3.0T-MR imaging with pathologic correlation. *Investigative Radiology* 2008;**43**(7):538-46.

Zhao 2014 {published data only}

Zhao XT, Li WX, Chai WM, Chen KM. Detection of small hepatocellular carcinoma using gadoteric acid-enhanced MRI: Is the addition of diffusion-weighted MRI at 3.0T beneficial? *Journal of Digestive Diseases* 2014;**15**(3):137-45.

Zhong 2020 {published data only}

Zhong X, Li JS, Chen ZJ, Yin JX, Gui S, Sun Z, et al. Texture analysis of diffusion-weighted magnetic resonance imaging to identify atypically enhanced small hepatocellular carcinoma and dysplastic nodules under the background of cirrhosis. *Chinese Journal of Hepatology* 2020;**28**(1):37-42.

Additional references

ACR 2021

ACR Committee on Drugs and Contrast Media. ACR manual on contrast media. www.acr.org/-/media/ACR/Files/Clinical-Resources/Contrast_Media.pdf (accessed 26 February 2021).

Balshem 2011

Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2011;**64**(4):401-6.

Beckett 2015

Beckett KR, Moriarity AK, Langer JM. Safe use of contrast media: what the radiologist needs to know. *Radiographics* 2015;**35**(6):1738-50.

Bertuccio 2017

Bertuccio P, Turati F, Carioli G, Rodriguez T, La Vecchia C, Malvezzi M, et al. Global trends and predictions in hepatocellular carcinoma mortality. *Journal of Hepatology* 2017;**67**(2):302-9.

Bosetti 2013

Bosetti C, Bertuccio P, Malvezzi M, Levi F, Chatenoud L, Negri E, et al. Cancer mortality in Europe, 2005–2009, and an overview of trends since 1980. *Annals of Oncology* 2013;**24**:2657-71.

Bosetti 2014

Bosetti C, Turati F, La Vecchia C. Hepatocellular carcinoma epidemiology. *Best Practice & Research. Clinical Gastroenterology* 2014;**28**:753-70.

Bottinor 2013

Bottinor W, Polkampally P, Jovin I. Adverse reactions to iodinated contrast media. *International Journal of Angiology* 2013;**22**:149-54.

Bralet 2000

Bralet MP, Regimbeau JM, Pineau P, Dubois S, Loas G, Degos F, et al. Hepatocellular carcinoma occurring in nonfibrotic liver: epidemiologic and histopathologic analysis of 80 French cases. *Hepatology* 2000;**32**(2):200-4.

Bruix 2011

Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology* 2011;**53**(3):1020.

Castillo 2009

Castillo E, Pelletier S, Kumer S, Abouljoud M, Divine G, Moonka D. Incidental hepatocellular carcinoma after liver transplantation: population characteristics and outcomes. *Transplant Proceedings* 2009;**41**(1):219-21.

Chan 2021

Chan MV, Huo YR, Trieu N, Mitchell A, George J, He E, et al. Noncontrast MRI for hepatocellular carcinoma detection: a systematic review and meta-analysis – a potential surveillance tool? *Clinical Gastroenterology and Hepatology* 2021;**S1542-3565**(21):00215-9.

Chang 2006

Chang KC, Leung CC, Tam CM. Per lesion analysis is misleading. *Thorax* 2006;**61**(4):364.

Chen 2013

Chen L, Zhang L, Bao J, Zhang J, Li C, Xia Y, et al. Comparison of MRI with liver-specific contrast agents and multidetector row CT for the detection of hepatocellular carcinoma: a meta-analysis of 15 direct comparative studies. *Gut* 2013;**62**:1520-1.

Chen 2014

Chen L, Zhang L, Liang M, Bao J, Zhang J, Xia Y, et al. Magnetic resonance imaging with gadoxetic acid disodium for the detection of hepatocellular carcinoma: a meta-analysis of 18 studies. *Academic Radiology* 2014;**21**(12):1603-13.

Chernyak 2018

Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, et al. Liver Imaging Reporting and Data System (LI-RADS) version 2018: imaging of hepatocellular carcinoma in at-risk patients. *Radiology* 2018;**289**(3):816-30.

Choi 2014a

Choi JY, Lee JM, Sirlin CB. CT and MR imaging diagnosis and staging of hepatocellular carcinoma: part I. Development, growth, and spread: key pathologic and imaging aspects. *Radiology* 2014;**272**:635-54.

Choi 2014b

Choi JY, Lee JM, Sirlin CB. CT and MR imaging diagnosis and staging of hepatocellular carcinoma: part II. Extracellular agents, hepatobiliary agents, and ancillary imaging features. *Radiology* 2014;**273**(1):30-50.

Chou 2015

Chou R, Cuevas C, Fu R, Devine B, Wasson N, Ginburg A, et al. Imaging techniques for the diagnosis of hepatocellular carcinoma: a systematic review and meta-analysis. *Annals of Internal Medicine* 2015;**162**:697-711.

Chung 2015

Chung YE, Kim KW. Contrast-enhanced ultrasonography: advance and current status in abdominal imaging. *Ultrasonography* 2015;**34**(1):3-18.

Cohen 2016

Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hoo PL, et al. STARD 2015 guidelines for reporting diagnostic

accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**:e012799.

Colli 2006

Colli A, Fraquelli M, Casazza G, Massironi S, Colucci A, Conte D, et al. Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. *American Journal of Gastroenterology* 2006;**101**:513-23.

Colli 2014

Colli A, Fraquelli M, Casazza G, Conte D, Nikolova D, Duca P, et al. The architecture of diagnostic research: from bench to bedside – research guidelines using liver stiffness as an example. *Hepatology* 2014;**60**(1):408-18.

Colli 2021

Colli A, Nadarevic T, Miletic D, Giljaca V, Fraquelli M, Štimac D, et al. Abdominal ultrasound and alpha-fetoprotein for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease. *Cochrane Database of Systematic Reviews* 2021, Issue 4. Art. No: CD013346. [DOI: [10.1002/14651858.CD013346.pub2](https://doi.org/10.1002/14651858.CD013346.pub2)]

Covidence 2020 [Computer program]

Veritas Health Innovation Covidence. Version accessed 10 June 2020. Melbourne, Australia: Veritas Health Innovation. Available at [covidence.org](https://www.covidence.org).

Davila 2012

Davila JA, Duan Z, McGlynn KA, El-Serag HB. Utilization and outcomes of palliative therapy for hepatocellular carcinoma: a population-based study in the United States. *Journal of Clinical Gastroenterology* 2012;**46**:71-7.

de Franchis 2015

de Franchis R. Expanding consensus in portal hypertension. Report of the Baveno VI Consensus Workshop: stratifying risk and individualizing care for portal hypertension. *Journal of Hepatology* 2015;**63**:743-52.

DTA Handbook 2021

Deeks JJ, Bossuyt PM, Leeflang M, Takwoingi Y, Flemyng E, Mellor L. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 2. London: Cochrane. Available from training.cochrane.org/handbook-diagnostic-test-accuracy/pdf/v2.

EASL 2016

European Association for the Study of the Liver. EASL Clinical Practice Guidelines: liver transplantation. *Journal of Hepatology* 2016;**64**(2):433-85.

EASL 2018

Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul JL, et al. EASL clinical practice guidelines: management of hepatocellular carcinoma. *Journal of Hepatology* 2018;**69**(1):182-236.

Edelman 2014

Edelman RR. The history of MR imaging as seen through the pages of radiology. *Radiology* 2014;**273**(2 Suppl):S181-200.

El Moghazy 2016

El Moghazy W, Kashkoush S, Meeberg G, Kneteman N. Incidence, characteristics, and prognosis of incidentally discovered hepatocellular carcinoma after liver transplantation. *Journal of Transplantation* 2016;**2016**:1916387.

Elsayes 2019

Elsayes KM, Kielar AZ, Chernyak V, Morshid A, Furlan A, Masch WR, et al. LI-RADS: a conceptual and historical review from its beginning to its recent integration into AASLD clinical practice guidance. *Journal of Hepatocellular Carcinoma* 2019;**6**:49-69.

Federle 2001

Federle MP, Blachar A. CT evaluation of the liver: principles and techniques. *Seminars in Liver Disease* 2001;**21**(2):135-45.

Feng 2021

Feng Z, Zhao H, Guan S, Wang W, Rong P. Diagnostic performance of MRI using extracellular contrast agents versus gadoxetic acid for hepatocellular carcinoma: a systematic review and meta-analysis. *Liver International* 2021;**41**(5):1117-28.

Ferlay 2015

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide: IARC Cancer Base No. 11. publications.iarc.fr/Databases/larc-Cancerbases/GLOBOCAN-2012-Estimated-Cancer-Incidence-Mortality-And-Prevalence-Worldwide-In-2012-V1.0-2012 (accessed 26 February 2021).

Ferlay 2019

Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *International Journal of Cancer* 2019;**144**(8):1941-53.

Floriani 2013

Floriani I, D'Onofrio M, Rulli E, Chen MH, Li R, Musicco L. Performance of imaging modalities in the diagnosis of hepatocellular carcinoma: a systematic review and meta-analysis. *Ultraschall in der Medizin* 2013;**34**:454-62.

Fornier 2018a

Fornier A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet* 2018;**391**(10127):1301-14.

Fraquelli 2019

Fraquelli M, Nadarevic T, Giljaca V, Colli A, Miletic D, Štimac D, et al. Contrast-enhanced ultrasound for the diagnosis of hepatocellular carcinoma in advanced chronic liver disease. *Cochrane Database of Systematic Reviews* 2019, Issue 11. Art. No: CD013483. [DOI: [10.1002/14651858.CD013483](https://doi.org/10.1002/14651858.CD013483)]

Grover 2015

Grover VP, Tognarelli JM, Crossey MM, Cox IJ, Taylor-Robinson SD, McPhail MJ. Magnetic resonance imaging: principles and techniques: lessons for clinicians. *Journal of Experimental Hepatology* 2015;**5**(3):246-55.

Guo 2016

Guo J, Seo Y, Ren S, Hong S, Lee D, Kim S, et al. Diagnostic performance of contrast-enhanced multidetector computed tomography and gadoxetic acid disodium-enhanced magnetic resonance imaging in detecting hepatocellular carcinoma: direct comparison and a meta-analysis. *Abdominal Radiology* 2016;**41**(10):1960-72.

Gupta 2021

Gupta P, Soundararajan R, Patel A, Kumar MP, Sharma V, Kalra N. Abbreviated MRI for hepatocellular carcinoma screening: a systematic review and meta-analysis. *Journal of Hepatology* 2021;**75**(1):108-19.

Hanna 2016

Hanna RF, Miloushev VZ, Tang A, Finklestone LA, Brejt SZ, Sandhu RS, et al. Comparative 13-year meta-analysis of the sensitivity and positive predictive value of ultrasound, CT, and MRI for detecting hepatocellular carcinoma. *Abdominal Radiology* 2016;**41**(1):71-90.

Hartwig 2009

Hartwig V, Giovannetti G, Vanello N, Lombardi M, Landini L, Simi S. Biological effects and safety in magnetic resonance imaging: a review. *International Journal of Environmental Research and Public Health* 2009;**6**(6):1778-98.

Hashim 2016

Hashim D, Boffetta P, La Vecchia C, Rota M, Bertuccio P, Malvezzi M, et al. The global decrease in cancer mortality: trends and disparities. *Annals of Oncology* 2016;**27**(5):926-33.

Heimbach 2018

Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* 2018;**67**(1):358-80.

Hennedige 2013

Hennedige T, Venkatesh SK. Imaging of hepatocellular carcinoma: diagnosis, staging and treatment monitoring. *Cancer Imaging* 2013;**12**(3):530-47.

Hori 1998

Hori M, Murakami T, Oi H, Kim T, Takahashi S, Matsushita M, et al. Sensitivity in detection of hypervascular hepatocellular carcinoma by helical CT with intra-arterial injection of contrast medium, and by helical CT and MR imaging with intravenous injection of contrast medium. *Acta Radiologica* 1998;**39**(2):144-51.

Hussain 2002

Hussain SM, Zondervan PE, IJzermans JN, Schalm SW, de Man RA, Krestin GP. Benign versus malignant hepatic nodules: MR imaging findings with pathologic correlation. *Radiographics* 2002;**22**(5):1026-39.

Junqiang 2014

Junqiang J, Yinzhong W, Li Z, Shulin G, Xiaohui W, Yanan Z, et al. Gadoteric acid disodium (Gd-EOB-DTA)-enhanced magnetic resonance imaging for the detection of hepatocellular carcinoma: a meta-analysis. *Journal of Magnetic Resonance Imaging* 2014;**39**(5):1079-87.

Kanwal 2019

Kanwal F, Singal AG. Surveillance for hepatocellular carcinoma: current best practice and future direction. *Gastroenterology* 2019;**157**(1):54-64.

Kierans 2016

Kierans AS, Kang SK, Rosenkrantz AB. The diagnostic performance of dynamic contrast-enhanced MR imaging for detection of small hepatocellular carcinoma measuring up to 2 cm: a meta-analysis. *Radiology* 2016;**278**(1):82-94.

Kim 2008

Kim SH, Choi BI, Lee JY, Kim SJ, So YH, Eun HW, et al. Diagnostic accuracy of multi-/single-detector row CT and contrast-enhanced MRI in the detection of hepatocellular carcinomas meeting the Milan criteria before liver transplantation. *Intervirology* 2008;**51**(Suppl 1):52-60.

Kinoshita 2015

Kinoshita A, Onoda H, Fushiya N, Koike K, Nishino H, Tajiri H. Staging systems for hepatocellular carcinoma: current status and future perspectives. *World Journal of Hepatology* 2015;**7**(3):406-24.

Lee 2012a

Lee JM, Yoon JH, Kim KW. Diagnosis of hepatocellular carcinoma: newer radiological tools. *Seminars in Oncology* 2012;**39**:399-409.

Lee 2012b

Lee JH, Lee JM, Kim SJ, Baek JH, Yun SH, Kim KW, et al. Enhancement patterns of hepatocellular carcinomas on multiphase multidetector row CT: comparison with pathological differentiation. *British Journal of Radiology* 2012;**85**(101):e573-83.

Lee 2015

Lee YJ, Lee JM, Lee JS, Lee HY, Park BH, Kim YH, et al. Hepatocellular carcinoma: diagnostic performance of multidetector CT and MR imaging – a systematic review and meta-analysis. *Radiology* 2015;**275**(1):97-109.

Li 2015a

Li X, Li C, Wang R, Ren J, Yang J, Zhang Y. Combined application of gadoteric acid disodium-enhanced magnetic resonance imaging (MRI) and diffusion-weighted imaging (DWI) in the diagnosis of chronic liver disease-induced hepatocellular carcinoma: a meta-analysis. *PLoS One* 2015;**10**(12):e0144247.

Li 2015b

Li YW, Chen ZG, Wang JC, Zhang ZM. Superparamagnetic iron oxide-enhanced magnetic resonance imaging for focal hepatic lesions: systematic review and meta-analysis. *World Journal of Gastroenterology* 2015;**21**(14):4334-44.

Li 2019

Li J, Wang J, Lei L, Yuan G, He S. The diagnostic performance of gadoteric acid disodium-enhanced magnetic resonance imaging and contrast-enhanced multi-detector computed tomography in detecting hepatocellular carcinoma: a meta-analysis of eight prospective studies. *European Radiology* 2019;**29**(12):6519-28.

Lijmer 1999

Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *Journal of the American Medical Association* 1999;**282**(11):1061-6.

LI-RADS 2017

American College of Radiology. Liver imaging reporting and data system. www.acr.org/-/media/ACR/Files/RADS/LI-RADS/CEUS-LI-RADS-2017-Core.pdf (accessed 7 July 2020).

LI-RADS 2018

American College of Radiology. Liver imaging reporting and data system, 2018. www.acr.org/-/media/ACR/Files/RADS/LI-RADS/LI-RADS-2018-Core.pdf (accessed 10 July 2020).

Liu 2013

Liu X, Zou L, Liu F, Zhou Y, Song B. Gadoteric acid disodium-enhanced magnetic resonance imaging for the detection of hepatocellular carcinoma: a meta-analysis. *PLoS One* 2013;**8**(8):e70896.

Liu 2017

Liu X, Jiang H, Chen J, Zhou Y, Huang Z, Song B. Gadoteric acid disodium-enhanced magnetic resonance imaging outperformed multidetector computed tomography in diagnosing small hepatocellular carcinoma: a meta-analysis. *Liver Transplantation* 2017;**23**(12):1505-18.

Llovet 1999

Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Seminars in Liver Disease* 1999;**19**(3):329-38.

Llovet 2003

Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. *Lancet* 2003;**362**(9399):1907-17.

Llovet 2008

Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *Journal of the National Cancer Institute* 2008;**100**(10):698-711.

Lok 2009

Lok AS, Seeff LB, Morgan TR, di Bisceglie AM, Sterling RK, Curto TM, et al. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. *Gastroenterology* 2009;**136**(1):138-48.

Madaleno 2015

Madaleno J, Alves R, Silva N, Calretas S, Tome L, Ferrao J, et al. Incidentally discovered hepatocellular carcinoma in

explanted liver: clinical, histopathologic features and outcome. *Transplantation Proceedings* 2015;**47**(4):1051-4.

Manini 2014

Manini MA, Sangiovanni A, Fornari F, Piscaglia F, Biolato M, Fanigliulo L, et al. Clinical and economical impact of 2012 AASLD guidelines for the diagnosis of hepatocellular carcinoma. *Journal of Hepatology* 2014;**60**(5):995-1001.

Mann 1990

Mann JR, Kasthuri N, Raafat F, Pincott JR, Parkes SE, Muir KR, et al. Malignant hepatic tumours in children: incidence, clinical features and aetiology. *Paediatric and Perinatal Epidemiology* 1990;**4**(3):276-89.

Mazzaferro 1996

Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *New England Journal of Medicine* 1996;**334**(11):693-9.

Mazzaferro 2011

Mazzaferro V, Bhoori S, Sposito C, Bongini M, Langer M, Miceli R, et al. Milan criteria in liver transplantation for hepatocellular carcinoma: an evidence based analysis of 15 years of experience. *Liver Transplantation* 2011;**17**(Suppl 2):S44-57.

McDonald 2019

McDonald JS, Hunt CH, Kolbe AB, Schmitz JJ, Hartman RP, Maddox DE, et al. Acute adverse events following gadolinium-based contrast agent administration: a single-center retrospective study of 281945 injections. *Radiology* 2019;**292**(3):620-7.

Mitelman 2018

Mitelman F, Johansson B, Mertens F, editor(s). Mitelman database of chromosome aberrations and gene fusions in cancer. mitelmandatabase.isb-cgc.org (accessed 14 October 2020).

Mogul 2018

Mogul DB, Ling SC, Murray KF, Schwarzenberg SJ, Rudzinski ER, Schwarz KB. Characteristics of hepatitis B virus-associated hepatocellular carcinoma in children: a multi-center study. *Journal of Pediatric Gastroenterology and Nutrition* 2018;**67**(4):437-40.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* 2009;**6**(7):e1000097.

Nadarevic 2021a

Nadarevic T, Giljaca V, Colli A, Fraquelli M, Casazza G, Miletic D, et al. Computed tomography for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease. *Cochrane Database of Systematic Reviews* 2021, Issue 10. Art. No: CD013362. [DOI: [10.1002/14651858.CD013362.pub2](https://doi.org/10.1002/14651858.CD013362.pub2)]

Nathani 2021

Nathani P, Gopal P, Rich N, Yopp A, Yokoo T, John B, et al. Hepatocellular carcinoma tumour volume doubling time: a systematic review and meta-analysis. *Gut* 2021;**70**(2):401-7.

Navin 2019

Navin PJ, Venkatesh SK. Hepatocellular carcinoma: state of the art imaging and recent advances. *Journal of Clinical and Translational Hepatology* 2019;**7**(1):72-85.

Ni 2004

Ni YH, Chang MH, Wang KJ, Hsu HY, Chen HL, Kao JH, et al. Clinical relevance of hepatitis B virus genotype in children with chronic infection and hepatocellular carcinoma. *Gastroenterology* 2004;**127**(6):1733-8.

O'Neill 2015

O'Neill EK, Cogley JR, Miller FH. The ins and outs of liver imaging. *Clinics in Liver Disease* 2015;**19**(1):99-121.

Omata 2017

Omata M, Cheng AL, Kokudo N, Kudo M, Lee JM, Jia J, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. *Hepatology International* 2017;**11**(4):317-70.

Pang 2018

Pang EH, Chan A, Ho SG, Harris AC. Contrast-enhanced ultrasound of the liver: optimizing technique and clinical applications. *American Journal of Radiology* 2018;**210**(2):320-32.

Park 2017

Park HJ, Choi BI, Lee ES, Park SB, Lee JB. How to differentiate borderline hepatic nodules in hepatocarcinogenesis: emphasis on imaging diagnosis. *Liver Cancer* 2017;**6**(3):189-203.

Piscaglia 2006

Piscaglia F, Bolondi L. The safety of Sonovue in abdominal applications: retrospective analysis of 23188 investigations. *Ultrasound in Medicine and Biology* 2006;**32**(9):1369-75.

Pomfret 2010

Pomfret EA, Washburn K, Wald C, Nalesnik MA, Douglas D, Russo M, et al. Report of a national conference on liver allocation in patients with hepatocellular carcinoma in the United States. *Liver Transplantation* 2010;**16**(3):262-78.

Reade 2008

Reade MC, Delaney A, Bailey MJ, Angus DC. Bench-to bedside review: avoiding pitfalls in critical care meta-analysis – funnel plots, risk estimates, types of heterogeneity, baseline risk and the ecologic fallacy. *Critical Care* 2008;**12**(4):220.

Review Manager 2020 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager (RevMan). Version 5.4. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

Roberts 2018

Roberts LR, Sirlin CB, Zaiem F, Almasri J, Prokop LJ, Heimbach JK, et al. Imaging for the diagnosis of hepatocellular

carcinoma: a systematic review and meta-analysis. *Hepatology* 2018;**67**(1):401-21.

Robinson 1950

Robinson WS. Ecological correlations and the behavior of individuals. *American Sociological Review* 1950;**15**:351-7.

Ryerson 2016

Ryerson AB, Eheman CR, Altekruze SF, Ward JW, Jemal A, Sherman RL, et al. Annual report to the nation on the status of cancer, 1975–2012, featuring the increasing incidence of liver cancer. *Cancer* 2016;**122**(9):1312-37.

Schirner 2004

Schirner M, Menrad A, Stephens A, Frenzel T, Hauff P, Licha K. Molecular imaging of tumor angiogenesis. *Annals of the New York Academy of Sciences* 2004;**1014**:67-75.

Schünemann 2020a

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* 2020;**122**:129-41.

Schünemann 2020b

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2020;**122**:142-52.

Senkerikova 2014

Senkerikova R, Frankova S, Sperl J, Oliverius M, Kieslichova E, Filipova H, et al. Incidental hepatocellular carcinoma: risk factors and long-term outcome after liver transplantation. *Transplantation Proceedings* 2014;**46**(5):1426-9.

Shirki 2015

Shirki JE, Seyal AR, Dighe MK, Yeh MM, Jalikis FG, Andeen NK, et al. CT of atypical and uncommon presentations of hepatocellular carcinoma. *American Journal of Roentgenology* 2015;**205**(4):W411-23.

Silva 2008

Silva MA, Hegab B, Hyde C, Guo B, Buckels JA, Mirza DF. Needle track seeding following biopsy of liver lesions in the diagnosis of hepatocellular cancer: a systematic review and meta-analysis. *Gut* 2008;**57**(11):1592-6.

Sirli 2010

Sirli R, Sporea I, Martie A, Popescu A, Danila M. Contrast enhanced ultrasound in focal liver lesions – a cost-efficiency study. *Medical Ultrasonography* 2010;**12**(4):280-5.

Smajerova 2016

Smajerova M, Petrasova H, Little J, Ovesna P, Andrasina T, Valek V, et al. Contrast-enhanced ultrasonography in the evaluation of incidental focal liver lesions: a cost-

effectiveness analysis. *World Journal of Gastroenterology* 2016;**22**(38):8605-14.

Stanaway 2016

Stanaway JD, Flaxman AD, Naghavi M, Fitzmaurice C, Vos T, Abubakar I, et al. The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. *Lancet* 2016;**388**(10049):1081-8.

Tang 2017

Tang C, Fang K, Guo Y, Li R, Fan X, Chen P, et al. Safety of sulfur hexafluoride microbubbles in sonography of abdominal and superficial organs: retrospective analysis of 30,222 cases. *Journal of Ultrasound in Medicine* 2017;**36**(3):531-8.

Tao 1984

Tao LC, Ho CS, McLoughlin MJ, Evans WH, Donat EE. Cytologic diagnosis of hepatocellular carcinoma by fine-needle aspiration biopsy. *Cancer* 1984;**53**(3):547-52.

Westwood 2013

Westwood M, Joore M, Grutters J, Redekop K, Armstrong N, Lee K, et al. Contrast-enhanced ultrasound using SonoVue® (sulphur hexafluoride microbubbles) compared with contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging for the characterisation of focal liver lesions and detection of liver metastases: a systematic review and cost-effectiveness analysis. *Health Technology Assessment* 2013;**17**(16):1-243.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Wu 2013

Wu LM, Xu JR, Gu HY, Hua J, Chen J, Zhu J, et al. Is liver-specific gadoxetic acid-enhanced magnetic resonance imaging a reliable tool for detection of hepatocellular carcinoma in patients with chronic liver disease? *Digestive Diseases and Sciences* 2013;**58**(11):3313-25.

Xian 2015

Xian JF, Chen M, Jin ZY. Magnetic resonance imaging in clinical medicine: current status and potential future developments in China. *Chinese Medical Journal* 2015;**128**(5):569-70.

Xiao 2016

Xiao YD, Paudel R, Liu J, Ma C, Zhang ZS, Zhou SK. MRI contrast agents: classification and application (review). *International Journal of Molecular Medicine* 2016;**38**(5):1319-26.

Xie 2011

Xie L, Guang Y, Ding H, Cai A, Huang Y. Diagnostic value of contrast-enhanced ultrasound, computed tomography and magnetic resonance imaging for focal liver lesions: a meta-analysis. *Ultrasound in Medicine and Biology* 2011;**37**(6):854-61.

Yang 2011

Yang JD, Harmsen WS, Slettedahl SW, Chaiteerakij R, Enders FT, Therneau TM, et al. Factors that affect the risk for hepatocellular carcinoma and effects of surveillance. *Clinical Gastroenterology and Hepatology* 2011;**9**(7):617-23.

Ye 2015

Ye F, Liu J, Ouyang H. Gadolinium ethoxybenzyl diethylenetriamine pentaacetic acid (Gd-EOB-DTPA)-enhanced magnetic resonance imaging and multidetector-row computed tomography for the diagnosis of hepatocellular carcinoma: a systematic review and meta-analysis. *Medicine (Baltimore)* 2015;**94**(32):e11157.

Yokoyama 1990

Yokoyama I, Todo S, Iwatsuki S, Starzl TE. Liver transplantation in the treatment of primary liver cancer. *Hepato-gastroenterology* 1990;**37**(2):188-93.

Young 2012

Young AL, Adair R, Prasad KR, Toogood GJ, Lodge JP. Hepatocellular carcinoma within a noncirrhotic, nonfibrotic, seronegative liver: surgical approaches and outcomes. *Journal of the American College of Surgeons* 2012;**214**(2):174-83.

Zwinderman 2008

Zwinderman AH, Glas AS, Bossuyt PM, Florie J, Bipat S, Stoker J. Statistical models for quantifying diagnostic accuracy with multiple lesions per patient. *Biostatistics* 2008;**9**(3):513-22.

References to other published versions of this review
Nadarevic 2021b

Nadarevic T, Colli A, Giljaca V, Fraquelli M, Casazza G, Manzotti C, Štimac D, Miletic D. Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease. *Cochrane Database of Systematic Reviews* 2021, Issue 5. Art. No: CD014798. [DOI: [10.1002/14651858.CD014798](https://doi.org/10.1002/14651858.CD014798)]

CHARACTERISTICS OF STUDIES
Characteristics of included studies [ordered by study ID]
Besa 2017
Study characteristics

Patient Sampling	<p>Included: people who underwent gadoxetic acid-enhanced liver MRI between 1 January 2011 to 31 December 2011 for HCC screening/surveillance or diagnosis/follow-up post-therapy. Final cohort consisted of 174 consecutive participants, 62 had HCC, 112 were HCC free.</p> <p>Excluded: people with HCC sized < 1 cm.</p>
Patient characteristics and setting	Consecutive participants at risk of HCC underwent MRI.
Index tests	<p>MRI positivity criteria: HCC was diagnosed on CE-set when a nodule showed arterial hyperenhancement followed by washout or capsule/pseudocapsule (or both) on PVP images, according to the AASLD 2011 criteria (Bruix 2011).</p> <p>MRI observers were blinded to clinical MRI reports and pathological results.</p>
Target condition and reference standard(s)	<p>Aim: to assess the diagnostic performance of a simulated AMRI protocol using DWI and T1W-HBP after gadoxetic acid injection alone and in combination for HCC detection in comparison with dynamic CE-T1W imaging, using histopathology (resection, OLT, biopsy) and follow-up as the reference standard.</p> <p>Pathology reports were matched with imaging findings to ensure observers analysed the correct lesions.</p>
Flow and timing	Time between index test and reference standard 12–270 days.
Comparative	

Besa 2017 (Continued)

Notes

Study reported COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Bhartia 2003
Study characteristics

Patient Sampling	Between August 2000 and October 2001, 93 adults with advanced hepatic cirrhosis who were potential liver transplantation candidates underwent double-contrast MRI. 31 participants (25 men and 6 women) aged 31–66 years (mean 52 years) later underwent transplantation with correlation of the pathological findings in the explanted liver.
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>Index test was MRI. All MR images were analysed by 1 of 2 experienced observers before OLT was performed.</p> <p>Positivity criteria: lesions that were visible on T2W images after superparamagnetic iron oxide administration and were also hypervascular on the T1W arterial phase images acquired after gadolinium administration were considered to be displaying characteristics typical of HCC. Lesions identified after superparamagnetic iron oxide administration that were not hypervascular and those that were hypervascular but were not identified after superparamagnetic iron oxide administration were reported as highly suspicious and were considered to be HCCs.</p>
Target condition and reference standard(s)	<p>Aim: to assess the clinical sensitivity of a modified double-contrast MRI technique in the detection of HCC in people with a cirrhotic liver by correlating the prospective interpretation of MRI with pathological findings in the explanted liver.</p> <p>Reference standard: OLT.</p>
Flow and timing	Time between index test and reference standard > 90 days.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Bhartia 2003 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Were positivity criteria clearly defined?	Yes	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		High
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Born 1998
Study characteristics

Patient Sampling	The radiological examinations (CT, MRI, and CTAP) of 47 people with cirrhosis having undergone an OLT were analysed retrospectively. In 57 consecutive participants who had undergone liver transplantation, cirrhosis was histologically confirmed in 47 cases in the explanted liver. Examinations conducted > 6 months ago at the time of the OLT were not included in the study. Of 47 includible patients, 28 were included.
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	Aim: to assess the value of MRI in the detection of malignant liver lesions (HCC) in the presence of cirrhosis of the liver. Examinations were retrospectively evaluated by 2 radiologists experienced in all

Born 1998 (Continued)

examination modalities without knowledge of the histological results.

Positivity criteria: all lesions that could not be assigned to a benign disease entity were rated as malignancy.

Target condition and reference standard(s)	Reference standard: OLT.
Flow and timing	Time between index test and reference standard > 90 days.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High

Born 1998 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Brunsing 2019
Study characteristics

Patient Sampling	<p>Included: consecutive participants aged ≥ 18 years, gadoxetate-enhanced AMRI radiology report completed, and cirrhosis of any aetiology and non-cirrhotic CHBV.</p> <p>Excluded: 187 participants were lost at follow-up and excluded, furthermore, known primary or secondary liver cancer; vascular cause of liver disease (Budd-Chiari syndrome) for which the LI-RADS did not apply.</p>
Patient characteristics and setting	Consecutive participants at risk of HCC underwent MRI screening.
Index tests	<p>Aim: to describe experience with gadoxetate disodium-enhanced AMRI for HCC screening and surveillance in people with cirrhosis or CHBV.</p> <p>Positivity criteria: HCC diagnosed when ≥ 1 observations not definitely benign and demonstrating HBP hypointensity, restricted diffusion, other suspicious features such as nodule-in-nodule appearance, or a combination of these with ≥ 1 measuring ≥ 10 mm.</p> <p>Prospectively rendered interpretation reports of all imaging studies were reviewed retrospectively.</p>
Target condition and reference standard(s)	<p>Reference standards: histology (biopsy), and US/CT/MRI follow-up.</p> <p>Blinded to the gadoxetate-enhanced AMRI results, the senior resident reviewed each participant's results of the reference standard tests.</p>
Flow and timing	Time between index test and reference standard 4–248 days. 187 participants were lost at follow-up and excluded.
Comparative	
Notes	Authors reported COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

Brunsing 2019 (Continued)

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	Yes
--	-----

Did the study avoid inappropriate exclusions?	No
---	----

Could the selection of patients have introduced bias?	High risk
--	-----------

Are there concerns that the included patients and setting do not match the review question?	Low concern
--	-------------

DOMAIN 2: Index Test (All tests)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
---	-----

Were positivity criteria clearly defined?	Yes
---	-----

Could the conduct or interpretation of the index test have introduced bias?	Low risk
--	----------

Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
--	-------------

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	No
---	----

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
--	-----

Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
---	-----------

Are there concerns that the target condition as defined by the reference standard does not match the question?	High
---	------

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	No
--	----

Did all patients receive the same reference standard?	No
---	----

Were all patients included in the analysis?	No
---	----

Could the patient flow have introduced bias?	High risk
---	-----------

Darnell 2021
Study characteristics

Patient Sampling	Aim: to evaluate the diagnostic accuracy of each LI-RADS category by MRI according to LI-RADS v2018 in a cohort of people with
------------------	--

Darnell 2021 (Continued)

	<p>cirrhosis in whom a solitary nodule ≤ 20 mm was detected during screening US. 315 people with cirrhosis with a single nodule ≤ 20 mm detected by surveillance US were included. 2 participants were excluded due to non-evaluable results.</p>
Patient characteristics and setting	<p>Patients at risk were screened with US, and suspected cases were referred to MRI.</p>
Index tests	<p>Used the positivity criteria from LI-RADS v2018, extracted data for the case of LI-RADS 5 category. Radiologists were unaware of the final diagnosis of the lesion.</p>
Target condition and reference standard(s)	<p>Reference standards: US-guided biopsy with histology, specific vascular profile (arterial phase hyperenhancement with washout) by MRI and follow-up. The index test was used as a reference standard and it represents incorporation bias.</p> <p>Further testing was decided based on the index test results.</p>
Flow and timing	<p>Time between index test and reference standard > 90 days. Used different reference standards.</p>
Comparative	
Notes	<p>Authors disclosed potential COI and funding was provided by a public institution.</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Darnell 2021 (Continued)

Is the reference standards likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

de Lédínghe 2002
Study characteristics

Patient Sampling	From February 1997 to July 1999, 34 participants were included on the basis that both MRI and spiral CT were performed before OLT. 20 participants were excluded because they did not have both MRI and CT.
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>Aim: to investigate the accuracy of MRI and spiral CT for hepatic nodule diagnosis in people with cirrhosis when compared with pathological findings of the whole explanted liver.</p> <p>Positivity criteria for MRI: all nodules that were hyperintense on T2W (excepted both liver cyst and haemangioma) or enhanced during arterial phase (or both) were considered as HCC. Radiologists were unaware of the final diagnosis of the lesion.</p>
Target condition and reference standard(s)	Liver histology of the whole explanted liver was considered the gold standard for HCC, dysplastic nodule and macroregenerative nodules diagnosis. In all cases, the pathologists were aware of the presence or absence of an HCC diagnosed at radiology and, most of the time, the gross location (right or left lobe) of the tumour was known.
Flow and timing	<p>Time between index test and reference standard for some participants > 90 days.</p> <p>All participants underwent the same reference standard (OLT).</p>

de Lédighen 2002 (Continued)

Comparative

Notes Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Demirtas 2020
Study characteristics

Patient Sampling	<p>Authors retrospectively reviewed a prospectively collected database of 1261 people with cirrhosis at a tertiary centre between December 2008 and February 2017.</p> <p>Excluded: people with insufficient data, those lost to follow-up, with contraindications to MRI, ineligible MRI surveillance periods (< 10 and > 14 months), AFP > 20 ng/mL at entry, and HCC diagnosis within 1 year. Patients with an AFP level 10–20 ng/mL at enrolment were only included if previous imaging confirmed the absence of any suspicious masses within 3 months from enrolment.</p>
Patient characteristics and setting	People with cirrhosis were screened for HCC.
Index tests	<p>Aim: to determine the performance of a surveillance strategy with annual MRI scans to detect HCCs at earlier curable stages.</p> <p>HCC was diagnosed radiologically or histologically (or both) using the EASL guidelines.</p> <p>Radiologists were unaware of the final diagnosis of the lesion.</p>
Target condition and reference standard(s)	Reference standards: biopsy or AFP/CT/MRI follow-up. Further testing was decided on the basis of the index test results.
Flow and timing	Time between index test and reference standard not reported. Used multiple reference standards.
Comparative	
Notes	Authors declared no COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		

Demirtas 2020 (Continued)

Could the conduct or interpretation of the index test have introduced bias?	Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	High
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Di Carlo 2012
Study characteristics

Patient Sampling	Study prospectively evaluated the accuracy of CEUS and MRI for the diagnosis of small nodules detected during US surveillance in people with cirrhosis. It included prospective evaluation of 89 people with cirrhosis.
Patient characteristics and setting	People with cirrhosis were screened by US and suspected cases were referred to CEUS and MRI.
Index tests	Index test positivity criteria: intense arterial uptake followed by washout in the venous/delayed phase was registered as conclusive for HCC. Radiologists were unaware of the final diagnosis of the lesion.
Target condition and reference standard(s)	Reference standards: used AASLD criteria and concordance of 2 contrast tests (CEUS and MRI). Used CEUS and MRI as confirmatory tests.
Flow and timing	Time between index test and reference standard not reported. Not all participants received the same reference standard.

Di Carlo 2012 (Continued)

Comparative

Notes Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Dumitrescu 2013
Study characteristics

Patient Sampling	Retrospective trial conducted between January 2010 and January 2013. Evaluated 126 people with focal liver lesions. Only patients with available data were included.
Patient characteristics and setting	< 30% of participants had evidence of CLD.
Index tests	Aim: to assess the role of transabdominal CEUS and MRI in characterisation and detection of HCC in people diagnosed with focal liver lesions. MRI positivity criteria were not clearly reported. 1 independent experienced radiologist analysed MR images.
Target condition and reference standard(s)	Reference standards: clinical data, blood analysis, imaging (CT/MRI), and histopathological information. Histopathology was used in only 15 cases with inconclusive imaging results, in benign cases MRI was considered the reference standard, other cases were unclear.
Flow and timing	Time between index test and reference standard unclear.
Comparative	
Notes	Authors declared no COI. Authors reported a public institution regarding source of funding.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Dumitrescu 2013 (Continued)

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	High

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Giorgio 2007
Study characteristics

Patient Sampling	Study designed to investigate the value of CE sonography in the characterisation of liver tumours and comparing the technique to dynamic gadobenate dimeglumine-enhanced MRI. A prospective evaluation of 73 consecutive patients with cirrhosis was conducted. Only patients with single nodules sized ≤ 30 mm were included.
Patient characteristics and setting	Only patients with single nodules sized ≤ 30 mm were included.
Index tests	<p>MRI positivity criteria: detection of nodules with a typical pattern of a round-shaped area of arterial hypervascularisation and a lack of portal supply was considered suggestive for HCC.</p> <p>Radiologists were unaware of the final diagnosis of the lesion.</p> <p>Aim: to investigate the characterisation of small HCCs in people with cirrhosis using CEUS and MRI.</p>
Target condition and reference standard(s)	Reference standards: US-guided biopsy and histology. Specimens for histological diagnosis of focal liver lesions were obtained at US-guided percutaneous needle biopsy. Biopsies were performed in all participants the following day, after both imaging studies. It was unclear whether the results of the MRI were known before performing the procedure.
Flow and timing	Time between index test and reference standard 1 day.

Giorgio 2007 (Continued)

Comparative

Notes Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Golfieri 2009
Study characteristics

Patient Sampling	283 consecutive patients with cirrhosis recruited between July 2003 and October 2004. The final study group included 63 participants who underwent MDCT and DC-MRI (SPIO-MRI plus dynamic MRI) within an interval of 15 days. People were excluded for having no nodules or benign regenerative nodule (122 people), and large (> 3 cm) HCC (4 people).
Patient characteristics and setting	People with HCC > 3 cm were excluded.
Index tests	<p>MR positivity criteria: dynamic MRI + SPIO-MRI (DC-MRI) – nodule was considered an HCC whenever it was typical in only 1 or in both MR studies.</p> <p>Radiologists were unaware of the final diagnosis of the lesion.</p> <p>Aim: to prospectively compare the diagnostic performances of ferucarbotran-enhanced MRI, and gadolinium-enhanced MRI in small (≤ 3 cm) nodules detected in people with cirrhosis during a US surveillance programme.</p>
Target condition and reference standard(s)	Final diagnosis was established at pathology on the explanted liver (10 participants), resection (6 participants), and biopsy (38 participants) specimens or at 2-year follow-up (9 participants). Participants underwent US-guided biopsy, it was unclear whether the results of the MRI were known before performing the procedure.
Flow and timing	In some cases of liver transplantation after MRI, the time interval was > 90 days.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Golfieri 2009 (Continued)

Were positivity criteria clearly defined?	Yes	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Hanna 2008
Study characteristics

Patient Sampling	From April 2001 to August 2004, double-contrast MRI examinations were performed on 434 patients for clinical care. Authors selected a study sample consisting of all patients with histologically confirmed cirrhosis and liver explant findings as the reference standard for HCC stage. Patients who underwent ablative therapy before MRI were excluded. 58 patients were excluded because MRI-explant time interval exceeded 12 months and the explant had positive results for HCC. The final study group consisted of 48 participants.
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	MRI positivity criteria: SPIO-MRI criteria – lesions were considered to have malignant features if on T1- and T2*-weighted SPGR images, the signal intensity of the nodules was increased relative to background liver parenchyma. Gadolinium MRI – features that suggested malignancy on gadolinium-enhanced images included arterial hyperenhancement with signal intensity greater than that of liver in the HAP; venous washout with hypointensity relative to surrounding liver in the portal venous or equilibrium phase; heteroge-

Hanna 2008 (Continued)

neous or mosaic enhancement; or presence of discrete capsule or pseudocapsule.

The radiologists were aware that participants had cirrhosis but were unaware of all other clinical, laboratory, pathological, and imaging findings.

Target condition and reference standard(s)	Participants had cirrhosis and all underwent liver transplantation after MRI. Pathologist also reviewed the clinical radiology reports and attempted to localise radiographic lesions with nodules found on tissue sections.
Flow and timing	Time between index test and reference standard > 90 days. All participants underwent OLT.
Comparative	
Notes	Information regarding COI not reported. Funding from public and private institutions.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		

Hanna 2008 (Continued)

Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	High
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Hwang 2014
Study characteristics

Patient Sampling	Authors searched retrospectively the institutional database for people aged > 18 years who underwent liver transplantation for CLD with or without HCC from April 2008 to October 2013. This search identified 699 patients, among whom 636 were excluded for the following reasons: no preoperative MRI (401), MRI obtained with other than gadoxetic acid as a contrast agent or a 3.0-T system (8), no MRI performed within 90 days of liver transplantation (147), a history of HCC treatment before MRI (79), and poor image quality (1). In all, 63 consecutive participants were included in the study.
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>Positivity criteria: the diagnostic criteria for the combined postcontrast and DWI images were a nodule showing enhancing foci during the arterial phase and hypointensity during the HBP, with or without washout during the PVP or the 3-minute late phase, with hyperintensity on DWI; a nodule showing either iso- or hypointensity during the arterial phase and hypointensity during the HBP with hyperintensity on DWI; a nodule showing enhancing foci during the arterial phase and hypointensity during the HBP, with or without washout during the PVP or the 3-minute late phase, without hyperintensity on DWI; and a nodule showing enhancing foci during the arterial phase and iso- or hyperintensity during the HBP with hyperintensity on DWI.</p> <p>1 on-site radiologist (with 12 years of experience in liver MRI interpretation) and 1 off-site gastrointestinal radiologist (with 5 years of experience) independently reviewed the MRI scans, and they were blinded to the initial MRI reports and pathology results.</p> <p>Aim: to investigate the diagnostic performance of gadoxetic acid-enhanced MRI with and without DWI in the detection of HCC in pretransplant livers.</p>
Target condition and reference standard(s)	Reference standard: liver transplantation.

Hwang 2014 (Continued)

A matched analysis was conducted between all liver lesions identified at imaging and pathology on a 1-by-1 basis with their location, size, and number.

Flow and timing

Time between index test and reference standard < 90 days. All participants underwent the same reference standard.

Comparative

Notes

Authors declared no COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High

Hwang 2014 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Khatri 2020
Study characteristics

Patient Sampling	<p>100 consecutive adults referred for diagnostic MRI of the liver in June–October 2016 at a large transplant centre.</p> <p>Inappropriate exclusion criteria: use of hepatobiliary contrast agent.</p>
Patient characteristics and setting	Consecutive patients with cirrhosis underwent MRI.
Index tests	<p>Aim: to compare dynamic screening AMRI to complete diagnostic cMRI for HCC detection in people with cirrhosis.</p> <p>For HCC screening, LR \geq 4 categories (i.e. LR-4, 5, M, and TIV) were considered test-positive.</p>
Target condition and reference standard(s)	<p>Reference standards: histopathology, follow-up imaging, consensus expert panel imaging review.</p> <p>The expert review panel formed by 2 senior fellowship-trained abdominal radiologists performed consensus unblinded review of any relevant imaging (including historical priors), clinical, or pathological data to determine the definitive HCC status.</p> <p>Positivity and negativity criteria: HCC positive – histopathological confirmation of HCC, or the presence of any LR 5 observation on CT or MRI at < 6 months confirmed by an expert review panel. HCC negative – the final HCC-negative status was determined by: the absence of any actionable observation on follow-up CT or MRI at < 12 months confirmed by expert review panel, or the absence of any imaging or clinical concerns for liver malignancy at a clinic visit > 12 months.</p> <p>Abdominal radiologists retrospectively performed MR analyses blinded to clinical data.</p>
Flow and timing	<p>Time between index test and reference standard not reported.</p> <p>7 participants excluded due to inability to determine final HCC status at 6 months.</p>
Comparative	
Notes	Potential COI not reported. Authors declared public funding.

Khatri 2020 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Kim 2017
Study characteristics

Patient Sampling	<p>Prospective study conducted at Asan Medical Center, an academic tertiary care centre in Korea (The PRIUS study, ClinicalTrials.gov ID: NCT01446666). Study participants were recruited between November 2011 and August 2012.</p> <p>Included: aged ≥ 20 years and presence of cirrhosis with an estimated annual HCC risk of $> 5\%$.</p> <p>Excluded: Child-Pugh class C liver function or an estimated glomerular filtration rate < 30 mL/minute/1.73 m².</p>
Patient characteristics and setting	Participants with cirrhosis underwent screening with MRI and US.
Index tests	Consecutive participants at risk of HCC underwent MRI. The positive screening criterion was a category 4 or 5 from modified LI-RADS.
Target condition and reference standard(s)	Reference standards: histology (biopsy) and follow-up with imaging. The confirmation of HCC was based on the results of histological examination or typical CT images (nodule > 1 cm with arterial hypervascularity and portal/delayed-phase washout) as recommended by practice guidelines (or both).
Flow and timing	Time between index test and reference standard < 90 days.
Comparative	
Notes	Authors disclosed COI and reported private and public funding.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Kim 2017 (Continued)

Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? No

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? No

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias? High risk

Kim 2020
Study characteristics

Patient Sampling	Study participants selected from 297 people with history of cirrhosis or CLD who underwent dynamic MRI of the liver for HCC surveillance between January 2010 and October 2017. Excluded people with previous HCC treatment (53 people), inadequate follow-up (< 6 months, 5 people), poor MRI quality or lacking DWI (6 people), and the presence of known malignancies other than HCC and intrahepatic cholangiocarcinoma (7 people). Remaining 226 patients were included in this study.
Patient characteristics and setting	Participants with cirrhosis or CLD underwent MRI.
Index tests	MR positivity criteria: the final imaging findings were scored as follows: (1) negative; (2) definitely benign; (3) probably benign; (4) indeterminate for HCC; (5) suspicious for HCC. A final imaging score of 1–3 was interpreted as negative for HCC while a score of 4 or 5 was interpreted as positive for HCC. Radiologists were blinded to the final outcome. Aim: to evaluate retrospectively the per-patient diagnostic performance of a minimised non-contrast MRI protocol for HCC surveillance and explored factors that might increase MRI sensitivity in this setting.
Target condition and reference standard(s)	Reference standards: liver explantation, resection, biopsy, radiological hallmarks of AASLD (modified version of LI-RADS, 2018 using complete MRI or multiphase CT), follow-up imaging with MRI or CT.

Kim 2020 (Continued)

Further testing was decided based on the index test results.

Flow and timing

Time between index test and reference standard not reported. Participants underwent diverse reference standards.

Comparative

Notes

Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Kim 2020 (Continued)

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Lauenstein 2007
Study characteristics

Patient Sampling	Between January 2004 and March 2006, 210 patients received a liver transplant at the authors' institution. 130/210 patients underwent MRI of the liver within 90 days before transplantation. The other 80 patients underwent MRI > 90 days before transplantation, had contraindications to MRI, or underwent CT. 15 patients were excluded because they had undergone liver chemoembolisation therapy for tumours (13 people) or were unable to complete MRI (2 people). Final study cohort had 115 participants.
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>Aim: to evaluate the accuracy of gadolinium-enhanced liver MRI in tumour surveillance.</p> <p>Positivity criteria: underlying criteria for the diagnosis of HCC included increased enhancement of the lesion compared with normal liver tissue in the arterial contrast phase; washout of lesions during the later contrast phases with isointensity or hypointensity compared with adjacent liver tissue; and development of peripheral rim enhancement, previously referred to as pseudocapsule enhancement, on delayed phase images. Lesions rated as HCC if ≥ 2 of the 3 features were present.</p> <p>Both reviewers were aware the participants were at high risk before liver transplantation, but they did not have access to the pathology reports.</p>
Target condition and reference standard(s)	<p>Reference standard: liver transplantation. Participants had cirrhosis and underwent liver transplantation after MRI.</p> <p>Lesion size and location were described in the manner used for the MRI evaluation and were compared with the findings in the detailed MRI report.</p>
Flow and timing	Time between index test and reference standard < 90 days.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Lauenstein 2007 (Continued)

Was a consecutive or random sample of patients enrolled?	Yes	
Did the study avoid inappropriate exclusions?	No	
Could the selection of patients have introduced bias?		High risk
Are there concerns that the included patients and setting do not match the review question?		High
DOMAIN 2: Index Test (All tests)		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Were positivity criteria clearly defined?	Yes	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		High
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Libbrecht 2002
Study characteristics
Magnetic resonance imaging for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease (Review)

Libbrecht 2002 (Continued)

Patient Sampling	<p>Between January 2000 and July 2001, 52 people with liver cirrhosis underwent liver transplantation.</p> <p>3 people without chronic hepatitis C virus infection for whom it was clear that their tumours exceeded the mentioned number and size limits received a donor liver from a person with positive serological markers for hepatitis C virus. These 3 people were excluded from the study.</p>
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>MRI examinations were interpreted in context of regular workup of people undergoing liver transplantation.</p> <p>Positivity criteria: nodular lesions that showed variable intensity on T1W images, hyperintensity on T2W images, enhancement during the arterial phase, or a combination of these were diagnosed as HCC.</p> <p>Aim: to correlate pathological results with pretransplantation clinical and imaging data, and to evaluate the accuracy of different imaging techniques performed during pretransplantation evaluation.</p>
Target condition and reference standard(s)	<p>Reference standard: liver transplantation.</p> <p>All cirrhotic explant livers were examined without knowledge of clinical or imaging data.</p>
Flow and timing	Time between index test and reference standard > 90 days.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		

Libbrecht 2002 (Continued)

Could the conduct or interpretation of the index test have introduced bias?	Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	High
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Lin 2016
Study characteristics

Patient Sampling	<p>Retrospective study. Between January 2006 and October 2010, 1016 people underwent liver tumour resections or liver transplantation. Of these, 841 people underwent liver CT or MRI examinations or had a pathological fibrosis score analysis, and were therefore enrolled in the study.</p> <p>Excluded: people who did not undergo liver CT or MRI examination before surgery, did not have a pathological fibrosis score analysis, or did not have liver tumours in the explanted liver.</p>
Patient characteristics and setting	841 participants with liver tumour who had liver CT or dynamic MRI examinations followed by surgical resection or OLT were included in the study. Patient population did not include only those with advanced stage of disease.
Index tests	<p>CT and MR were performed as preoperative evaluation, so no reference standard results were available.</p> <p>Positivity criteria: authors defined the typical HCC imaging characteristics as early enhancement in the arterial phase and early washout in the venous phase.</p>

Lin 2016 (Continued)

Aim: to comprehensively compare liver CT and dynamic MRI for HCC diagnosis before surgical resection.

Target condition and reference standard(s)	Reference standard: liver tumour resections or liver transplantation. Histological and surgical reports were reviewed to confirm HCC or other liver tumours. Pathological results read by pathologists with sufficient experience in the field and who were blinded to the clinical and radiological results.
Flow and timing	Time between index test and reference standard not reported.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Lin 2016 (Continued)

Are there concerns that the target condition as defined by the reference standard does not match the question?

Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Unclear

Did all patients receive the same reference standard? No

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias? High risk

Maiwald 2014
Study characteristics

Patient Sampling	<p>50 participants in prospective single-centre study to evaluate the diagnostic performance of CE CT and Gd-EOB-DTPA-enhanced MRI in terms of lesion detection.</p> <p>Included: suspicious findings in the US or increased laboratory parameters (e.g. alpha-fetoprotein) (or both).</p> <p>Excluded: renal failure, allergy to contrast agents, hyperthyroidism, pregnancy, and, especially for the MRI examination, pacemaker or other non-compatible implants and claustrophobia.</p>
Patient characteristics and setting	Participant at risk were screened with US and AFP, and suspected cases were referred to CT and MRI.
Index tests	<p>Aim: to compare the diagnostic power of CT with 3 Tesla MRI using Gd-EOB-DTPA for the diagnosis of HCC.</p> <p>Positivity criteria: diagnosis of HCC was based on hypervascularisation in the arterial phase and washout in the PVP or delayed phase, as suggested by the EASL and AASLD for MRI and CT.</p> <p>Radiologists were unaware of the final diagnosis of the lesion.</p>
Target condition and reference standard(s)	<p>Histopathological report after resection or biopsy of a lesion served as the gold standard for diagnosis, whereas a surrogate of follow-up (after 6 months) or complementary imaging technique (US, digital subtraction angiography) in combination with clinical (loss of weight, general state) and paraclinical parameters (especially alpha-fetoprotein) was used in unresected lesions.</p> <p>Reference standard for positive test: biopsy and liver resection. In cases of biopsy, the results of the test had to be known to plan the procedure.</p>
Flow and timing	Time between index test and reference standard not reported.
Comparative	
Notes	Authors declared no COI and reported private funding.

Maiwald 2014 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Marks 2015
Study characteristics

Patient Sampling	<p>580 participants enrolled in an MRI-based HCC surveillance programme who were imaged from 23 October 2008 to 31 January 2012 at the 2 participating institutions were eligible. Only the first available examination was included for each participant to avoid duplicates.</p> <p>Included: history of cirrhosis or other risk factors for HCC without having a known HCC or LI-RADS category 4 or 5 observation and without prior empirical treatment of an HCC.</p> <p>Excluded: no follow-up examinations (252 people), inadequate follow-up examinations or procedures to meet reference standard criteria such as non-multiphase or unenhanced follow-up studies (29), and a known malignancy other than HCC with liver metastases (1).</p>
Patient characteristics and setting	Participants at risk were screened with MRI.
Index tests	<p>Aim: to evaluate the per-patient diagnostic performance of an abbreviated gadoteric acid-enhanced MRI protocol for HCC surveillance that included gadoteric acid-enhanced HBP as a potentially lower-cost alternative to conventional MRI for HCC surveillance in people at risk for HCC.</p> <p>Radiologists were blinded to the clinical data, clinical reports, reference standard, and each other's interpretations.</p> <p>MR positivity criteria: imaging scores 4 and 5 based according to signal characteristics.</p>
Target condition and reference standard(s)	Reference standards: liver explant, liver resection, liver biopsy, and follow-up with imaging. Further testing was decided based on the index test results.
Flow and timing	Time between index test and reference standard > 90 days. Not all participants underwent the same reference standard.
Comparative	
Notes	Authors disclosed potential COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern

Marks 2015 *(Continued)*
DOMAIN 2: Index Test (All tests)

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Were positivity criteria clearly defined? Yes

Could the conduct or interpretation of the index test have introduced bias? Low risk

Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? No

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? No

Did all patients receive the same reference standard? No

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias? High risk

Marrero 2005
Study characteristics

Patient Sampling	<p>Between May 2002 and June 2003, consecutive patients with cirrhosis and a suspected liver mass who underwent MRI for further evaluation were included. Indications for MRI were an elevated (20 ng/mL) AFP level; suggestion of a mass on US or a CT scan without arterial phase; or unexplained symptoms.</p> <p>94 consecutive people with known cirrhosis and an enhancing liver mass were prospectively evaluated with dynamic gadolinium-enhanced MRI.</p> <p>12 people were excluded from the analysis; 8 had no visible mass on MRI and 4 had simple hepatic cysts.</p>
------------------	---

Marrero 2005 (Continued)

Patient characteristics and setting	Participants with a suspected arterially enhancing mass were included, which represents narrow inclusion criteria.
Index tests	<p>Radiologists were unaware of the final diagnosis of the lesion.</p> <p>Positivity criteria: HCC included all arterially enhancing lesions 2 cm regardless of their other imaging features, and all arterially enhancing lesions with transverse relaxation time–hyperintensity or delayed hypointensity (or both) regardless of their size.</p> <p>Consecutive patients with known cirrhosis and an enhancing liver mass were prospectively evaluated with dynamic gadolinium-enhanced MRI.</p> <p>Aim: to determine whether the combination of clinical, laboratory, radiological, or a combination of these data can improve the prediction of HCC.</p>
Target condition and reference standard(s)	<p>Reference standard: histology (biopsy, explantation), and follow-up with imaging.</p> <p>Further testing was decided on the basis of the index test results.</p>
Flow and timing	<p>Time between index test and reference standard not reported.</p> <p>12 participants were excluded from the analysis; 8 had no visible mass on MRI and 4 had simple hepatic cysts.</p>
Comparative	
Notes	Potential COI not reported. Authors reported public funding.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Marrero 2005 (Continued)

Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? No

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Unclear

Did all patients receive the same reference standard? No

Were all patients included in the analysis? No

Could the patient flow have introduced bias? High risk

McNamara 2018
Study characteristics

Patient Sampling	<p>364 consecutive patients who underwent liver transplant from August 2009 to April 2013 were screened for MRI within 6 months before transplantation.</p> <p>Excluded: patients with pretransplant TACE or other ablative therapy.</p> <p>Clinical reports generated at the time imaging were not reviewed as part of patient enrolment. 37 adults met criteria for enrolment. 17 participants had pathologically confirmed HCC, and 20 had no lesion by imaging, confirmed on explant specimen or biopsy. Finally, 37 participants were included in the analysis.</p>
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>3 abdominal imagers with 20, 3, and 5 years' experience, who were blinded to the pathology results and clinical interpretations reviewed the studies in 2 separate sessions.</p> <p>Positivity criteria: hyperintense lesion on DWI demonstrating arterial enhancement (HAP) and PVP washout with pseudo-capsule hypointense on ADC.</p>

McNamara 2018 (Continued)

Aim: to compare the sensitivity and specificity of diffusion-weighted liver MRI alone with complete, multiphasic gadoteridol-enhanced MRI for the detection of HCC in people with cirrhosis before liver transplant.

Target condition and reference standard(s)	Reference standard: OLT.
Flow and timing	Time between index test and reference standard > 90 days.
Comparative	
Notes	Information on potential COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			

McNamara 2018 (Continued)

Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Min 2018a
Study characteristics

Patient Sampling	<p>Retrospective, cross-sectional analysis of patients with available pathology information for primary hepatic tumour and had gadoxetic acid-enhanced liver MRI within 1 month prior to surgery. Authors screened 2592 consecutive, treatment-naïve patients who underwent gadoxetic acid-enhanced liver MRI for evaluation of primary hepatic tumours between August 2012 and November 2015 using a computerised search of the hospital system. Authors included 798 participants who underwent surgery (771 with liver resection and 27 with liver transplantation) as a first-line treatment.</p> <p>For the detailed radiological-pathological correlation for ancillary features, 1794 patients with only biopsy results or no pathological diagnosis were excluded.</p> <p>773 consecutive patients with surgically resected 773 primary hepatic tumours (699 HCCs, 63 intrahepatic cholangiocarcinomas, and 11 benign nodules) who underwent gadoxetic acid-enhanced MRI were retrospectively identified.</p>
Patient characteristics and setting	
Index tests	<p>Radiologists were blinded to the specific pathological diagnosis of tumours.</p> <p>Positivity criteria: based on enhancement pattern, the index study (MRI) was positive when a tumour showed arterial diffuse hyperenhancement with washout. The index study was negative when the tumour showed no arterial hyperenhancement without washout. The index study was inconclusive and considered negative when the tumour showed either arterial diffuse hyperenhancement alone or washout alone.</p> <p>Aim: to assess the accuracy of enhancement pattern and ancillary features for detection of HCC.</p>
Target condition and reference standard(s)	Reference standards: liver resection and liver transplantation. Authors included 798 participants who underwent surgery (771 with liver resection and 27 with liver transplantation).
Flow and timing	Time between index test and reference standard < 90 days.
Comparative	
Notes	Authors declared no potential COI. Information regarding funding not reported.

Methodological quality

Min 2018a (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Sangiovanni 2010

Study characteristics

Methodological quality

Sangiovanni 2010 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Were positivity criteria clearly defined?	Yes	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Seçil 2008
Study characteristics

Patient Sampling	<p>Study group of this retrospective investigation was composed of people with cirrhosis who had liver MRI in picture archiving and communication system archive.</p> <p>Included: histopathological diagnosis of cirrhosis, clinical and MRI follow-up > 1 year, and presence of a complete series of standard liver MR images. The MRIs were performed to evaluate the severity of cirrhosis or portal hypertension, screening for hepatic lesions suspected with other imaging modalities.</p>
Patient characteristics and setting	Only people with available MRI were included.
Index tests	Image analysis was performed independently by 2 other investigators who were experienced in liver MRI. The investigators were unaware of the clinical condition of the participants; all data were hidden during the image analysis.

Seçil 2008 (Continued)

	Positivity criteria not reported.
Target condition and reference standard(s)	Reference standards: OLT, resection, biopsy, chemoembolisation and lipiodol CT, follow-up with periodic CT Further testing was decided on the basis of the index test results.
Flow and timing	Time between index test and reference standard not reported.
Comparative	
Notes	Potential COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			

Seçil 2008 (Continued)

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Sersté 2012
Study characteristics

Patient Sampling	Between January 2005 and December 2010, all consecutive patients referred to the liver unit with cirrhosis or CLD and small nodules (diameter 1–2 cm) newly detected by US and without previous HCC were included in study. All participants underwent an initial evaluation, including systematic CE multiphasic CT, MRI, and a liver biopsy of the nodule, all performed within 1 month.
Patient characteristics and setting	Only people with solitary liver tumour sized 1–2 cm were included.
Index tests	<p>CT and MRI results were read by 2 radiologists in consensus who were blind to biopsy results.</p> <p>Positivity criteria: vascular pattern was qualified as "conclusive" for HCC if contrast washout occurred, defined as the presence of hypervascularity during the arterial phase followed up by a hypodense/hypointense appearance in later phases defining washout.</p> <p>Aim: to assess the accuracy of CT and MRI in US-detected nodules in 75 consecutive patients with CLD or cirrhosis (or both).</p>
Target condition and reference standard(s)	Reference standard: biopsy and follow-up with imaging. All biopsies were routinely read by 1 pathologist, then independently reviewed in a blinded manner by a second pathologist who was unaware of the previous pathological diagnosis and imaging results.
Flow and timing	<p>Time between index test and reference standard < 90 days.</p> <p>1 participant was withdrawn from the study because the studied nodule with washout on both examinations, without conclusive diagnosis on biopsy underwent radiofrequency ablation.</p>
Comparative	
Notes	Potential COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Sersté 2012 (Continued)

Was a consecutive or random sample of patients enrolled?	Yes	
Did the study avoid inappropriate exclusions?	Yes	
Could the selection of patients have introduced bias?		Low risk
Are there concerns that the included patients and setting do not match the review question?		High
DOMAIN 2: Index Test (All tests)		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Were positivity criteria clearly defined?	Yes	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Shin 2017
Study characteristics

Patient Sampling	Authors conducted a retrospective study of 90 people with cirrhosis with a single liver nodule ≤ 3 cm in diameter showing low signal in-
------------------	---

Shin 2017 (Continued)

tensity in the HBP of Gd-EOB-DTPA-enhanced MRI between December 2008 and December 2015. Gd-EOB-DTPA-enhanced MRI and DWI were performed, and pathological evaluation was performed in all patients. Atypical HCC was defined as not showing typical enhancement (arterial hyperenhancement and delayed washout) on dynamic MRI. Of these, 53 people with liver nodule showing typical enhancement pattern on dynamic MRI were excluded. The final group consisted of 43 participants.

Patient characteristics and setting	Participants with nodules \leq 3 cm in diameter with atypical features were included only.
Index tests	MR images were retrospectively analysed by 2 radiologists who were unaware of the pathological results. Positivity criteria: high signal intensity on T2W imaging plus high signal intensity on DWI. Aim: to assess the usefulness of Gd-EOB-DTPA-enhanced MRI including DWI for differentiation between atypical small HCCs and dysplastic nodules.
Target condition and reference standard(s)	Reference standard: US-guided biopsy or surgical resection and follow-up with imaging. Further testing was decided on the basis of the index test results.
Flow and timing	Time between index test and reference standard not reported.
Comparative	
Notes	Authors declared no potential COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Shin 2017 (Continued)

Are there concerns that the index test, its conduct, or interpretation differ from the review question? High

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? No

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Unclear

Did all patients receive the same reference standard? No

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias? High risk

Sutherland 2017
Study characteristics

Patient Sampling	<p>Prospective investigation of the role of diffusion-weighted hepatic MRI for HCC screening in the setting of CLD.</p> <p>Recruitment criteria: aged > 18 years, referred by the gastroenterology department with CLD for HCC screening liver US. Exclusion criteria: presence of a known mass as indicated on the US request form, non-English speaking (due to inability to gain informed consent), or contraindications to MRI such as pacemaker or metallic implant.</p>
Patient characteristics and setting	<p>People aged > 18 years with CLD were referred by the gastroenterology department for HCC screening.</p>
Index tests	<p>Positivity criteria: MRI lesions were considered suspicious if they had elevated signal on high b value DWI and were iso- or hypointense to background liver on the ADC map.</p> <p>Radiologist were unaware of the final diagnosis of the lesion.</p> <p>Aim: to prospectively investigate the role of diffusion-weighted hepatic MRI for HCC screening in the setting of CLD.</p>
Target condition and reference standard(s)	<p>Gold standard for the diagnosis of HCC histology (biopsy or resection) and follow-up with imaging.</p>

Sutherland 2017 (Continued)

Location and size of all confirmed HCC was correlated with the screening test result locations.

Flow and timing

Time between index test and reference standard not reported.

Comparative

Notes

Authors declared no potential COI and no funding.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Sutherland 2017 (Continued)

Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Teefey 2003
Study characteristics

Patient Sampling	<p>Between August 1996 and December 1998, authors examined 37 patients with end-stage liver disease who had been listed for hepatic transplantation. Only patients with an elevated serum alpha-fetoprotein level (30 g/L) or with primary sclerosing cholangitis were eligible. All patients underwent CT, MR imaging, US, and PET.</p> <p>2 patients whose names had been on the transplant list for > 2 years were not included in the study because of an inability to obtain follow-up images – inappropriate exclusion.</p>
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>Positivity criteria: a focal area of increased enhancement on arterial phase images was considered suggestive of HCC. An additional criterion for diagnosing HCC was mild lesion hyperintensity on T2W images.</p> <p>Radiologists were unaware of the final diagnosis.</p> <p>Aim: to determine and compare the diagnostic performance of CT, MRI, US, and PET against the standard of histological examination of the resected liver specimen to assess which single test or combination of tests was most accurate in the detection of HCC in listed liver transplant candidates.</p>
Target condition and reference standard(s)	<p>Reference standards: liver transplantation, biopsy, and follow-up with imaging.</p> <p>The presence or absence of all lesions identified with ≥ 1 of the imaging tests (CT, MRI, US, or PET) was determined histologically on a lesion-by-lesion basis.</p>
Flow and timing	Time between index test and reference standard > 90 days.
Comparative	
Notes	Information regarding COI and funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Teefey 2003 (Continued)

Was a consecutive or random sample of patients enrolled?	Yes	
Did the study avoid inappropriate exclusions?	No	
Could the selection of patients have introduced bias?		High risk
Are there concerns that the included patients and setting do not match the review question?		High
DOMAIN 2: Index Test (All tests)		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Were positivity criteria clearly defined?	Yes	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Vietti Violi 2020
Study characteristics

Vietti Violi 2020 (Continued)

Patient Sampling	<p>Retrospective single-centre study in which authors searched the institutional electronic imaging database in 2017 for consecutive patients who underwent in-house MRI for HCC screening/surveillance.</p> <p>Included: adult aged ≥ 18 years with cirrhosis of any aetiology and patients with chronic hepatitis B without cirrhosis who underwent a complete gadoxetate-enhanced MRI for HCC screening at authors' institution. According to the AASLD criteria for HCC screening, patients with Child-Pugh class C were included if listed for transplantation. Among 415 initial patients, 178 were excluded. The final group included 237 participants.</p> <p>Excluded: patients who underwent MRI with extracellular contrast.</p>		
Patient characteristics and setting	People at risk for HCC were screened with MRI.		
Index tests	<p>Positivity criteria: LI-RADS 5 – features include a 10–19 mm nodule with non-rim arterial phase hyperenhancement and non-peripheral washout appearance, regardless of capsule appearance, or a ≥ 20 mm nodule with non-rim arterial phase enhancement with either enhancing capsule appearance or non-peripheral washout appearance or both.</p> <p>Aim: to compare the performance of 3 different MRI protocol sets extracted from a complete gadoxetate-enhanced MRI obtained for HCC screening.</p>		
Target condition and reference standard(s)	<p>Reference standards: liver explant pathology, biopsy, MRI criteria, clinical data, and follow-up.</p> <p>Classification based on the review of all available participant data including imaging examinations, pathology, any subsequent treatment, and decision from the multidisciplinary tumour board.</p>		
Flow and timing	Time between index test and reference standard not reported.		
Comparative			
Notes	Authors disclosed potential COI and private and public funding.		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			

Vietti Violi 2020 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
Were positivity criteria clearly defined?	Yes
Could the conduct or interpretation of the index test have introduced bias?	Unclear risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	High
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Villacastin Ruiz 2016
Study characteristics

Patient Sampling	<p>From November 2001 to December 2011, 323 OLTs were performed on 313 patients. This study is based on the retrospective analysis of data from 273 patients who underwent scheduled transplants because of cirrhosis.</p> <p>Excluded: undergone urgent non-selective transplants, having undergone retransplantation, and absence of cirrhosis.</p> <p>273 consecutive patients with 218 HCC nodules, who underwent imaging and subsequent transplantation, were examined.</p>
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	Authors retrospectively revised all of the pretransplant reports carried out by experienced radiologists for each imaging study.

Villacastin Ruiz 2016 (Continued)

Positivity criteria: lesions suggesting HCC were typically characterised by hypervascularity, especially when accompanied by venous-phase washout. Moreover, signs of malignancy included new nodules, rapidly growing nodules, and nodules with an intermediate–high T2 signal.

Aim: to evaluate the accuracy of diverse imaging tests in the preoperative detection and correct tumour staging of HCC in patients being considered for liver transplantation by correlating the imaging findings with results of the pathological examination of the whole explant liver, which was the gold standard.

Target condition and reference standard(s)	Reference standard: pathology of the whole explanted liver. Correlation of nodules between the image and pathological results was based primarily on location and secondarily on size. All participants underwent liver transplantation only.
Flow and timing	Time between index test and reference standard > 90 days. Analysed 112/114 participants.
Comparative	
Notes	Authors declared no COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Were positivity criteria clearly defined?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Villacastin Ruiz 2016 (Continued)

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	High
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Wu 2020
Study characteristics

Patient Sampling	<p>Retrospective analysis of 120 people with liver cirrhosis who were surgically treated at a tertiary institution from December 2018 to November 2019. They were divided into 2 groups according to different checkups, with 60 cases in each group. Participants in group 1 underwent MRI plain scans and participants in group 2 underwent Gd-EOB-DTPA-enhanced MRI scans.</p> <p>Excluded: people with hepatic metastatic tumours, those who withdrew from study for no reason, and those who had had previous liver surgery.</p>
Patient characteristics and setting	Aim: to explore the diagnostic value of Gd-EOB-DTPA-enhanced MRI in small HCC in people with liver cirrhosis who underwent hepatic resection for focal lesion.
Index tests	MR positivity criteria not reported.
Target condition and reference standard(s)	<p>Reference standards: biopsy and surgery with histology.</p> <p>Compared diagnostic results and the final pathological results of both groups.</p>
Flow and timing	Time between index test and reference standard not reported.
Comparative	
Notes	Authors declared no COI. Reported public funding.

Wu 2020 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Were positivity criteria clearly defined?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Yu 2011

Study characteristics

Patient Sampling	<p>Retrospective study. Authors searched the institutional database, and found 1097 adults who received OLT from January 1999 to November 2006. Of these, 638 people with CLD underwent unenhanced US, CE single or multidetector helical CT, dynamic cMRI, or a combination of these within 6 months of the transplantation.</p> <p>Excluded: people with studies at outside imaging centres.</p> <p>The final participant group consisted of 638 consecutive adults with cirrhosis who received liver transplants within 6 months of imaging at a tertiary care institution.</p>
Patient characteristics and setting	All participants underwent OLT, therefore only people with advanced liver disease were included.
Index tests	<p>Positivity criteria: lesions suspicious for HCC were typically characterised by ≥ 1 of the following features: new or rapidly growing nodule; nodule with arterial hypervascularity, especially when accompanied by venous phase washout; dominant nodule containing fat; and nodule with intermediate–high T2 signal.</p> <p>Prospectively rendered interpretation reports of all imaging studies were reviewed retrospectively.</p> <p>This retrospective study provides a broad survey of the accuracy of US, CT, and MRI for HCC detection in a large population of people with cirrhosis undergoing liver transplantation. Authors' main goal was to evaluate the performance of the 3 cross-sectional imaging modalities in the context of routine clinical interpretations using explant pathology as the reference standard.</p>
Target condition and reference standard(s)	<p>Reference standard: liver transplantation for all participants.</p> <p>Pathologists were not provided with the imaging reports regarding number and locations of any suspected lesions.</p>
Flow and timing	Time between index test and reference standard > 90 days.
Comparative	
Notes	Authors disclosed potential COI. Funding not reported.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High

Yu 2011 (Continued)

DOMAIN 2: Index Test (All tests)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
---	-----

Were positivity criteria clearly defined?	Yes
---	-----

Could the conduct or interpretation of the index test have introduced bias?	Low risk
--	----------

Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
--	-------------

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
---	-----

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
--	-----

Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
---	----------

Are there concerns that the target condition as defined by the reference standard does not match the question?	High
---	------

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	No
--	----

Did all patients receive the same reference standard?	Yes
---	-----

Were all patients included in the analysis?	Yes
---	-----

Could the patient flow have introduced bias?	High risk
---	-----------

AASLD: American Association for the Study of Liver Diseases; ADC: apparent diffusion coefficient; AFP: alpha-fetoprotein; AMRI: abbreviated magnetic resonance imaging; CE: contrast-enhanced; CEUS: contrast-enhanced ultrasound; CHBV: chronic hepatitis B virus; CLD: chronic liver disease; cMRI: contrast-enhanced magnetic resonance imaging; COI: conflict of interest; CT: computed tomography; CTAP: computed tomography arterial portography; DC-MRI: double-contrast magnetic resonance imaging; DWI: diffusion weighted imaging; EASL: European Association for the Study of the Liver; FNB: fine-needle biopsy; Gd-EOB-DTPA: gadolinium ethoxybenzyl-diethylenetriaminepentaacetic acid; HAP: hepatic arterial phase; HBP: hepatobiliary phase; HCC: hepatocellular carcinoma; LI-RADS (criteria): Liver Imaging Reporting and Data System criteria; LR: likelihood ratio; MDCT: multidetector computed tomography; MR: magnetic resonance; MRI: magnetic resonance imaging; OLT: orthotopic liver transplantation; PET: positron emission tomography; PVP: portal venous phase; SPGR: spoiled gradient-echo; SPIO-MRI: superparamagnetic iron oxide magnetic resonance imaging; T1W: T1-weighted; T2W: T2-weighted; TACE: transarterial chemoembolisation; TIV: tumour in vein; US: ultrasound.

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Abdel Kader 2017	Target condition was hepatic focal lesions with differentiation between benign and malignant lesions. No data on per-patient analysis.
Akhtar 2020	2 × 2 table was not reported directly and could not be calculated/extracted based on the reported data.
Araki 2000	Editorial.
Aubé 2017	Study population included > 5% of participants previously treated for HCC. No data on per-patient analysis.
Ayuso 2019	No data on per-patient analysis.
Baird 2013	No data on per-patient analysis.
Barabino 2021	Study assessed the accuracy of LI-RADS features, no data on per-patient analysis reported, and no data on accuracy of MRI alone.
Bartolozzi 2013	No data on per-patient analysis.
Basha 2018	No data on per-patient analysis.
Becker-Weidman 2011	Study population included > 5% of participants previously treated for HCC.
Blondin 2011	No data on per-patient analysis.
Burrel 2003	Study included participants with previously known HCC.
Choi 2008	Study population included > 5% of participants previously treated for HCC.
Chou 2014	No data on per-patient analysis.
Clarke 2021	No data on per-patient analysis.
Compagnon 2008	Study population included > 5% of participants previously treated for HCC. The 2 × 2 table was not reported directly and could not be calculated/extracted based on the data that were reported.
Debees 2016	Target condition was differentiation of benign versus malignant lesions. No data on per-patient analysis.
Ercalli 2013	No data on per-patient analysis.
Faletti 2015	Study included participants with previously known HCC.
Fischer 2015	Study assessed the accuracy of MRI imaging features for the diagnosis of HCC.
Forner 2008	Participant data set was included in Darnell 2021 .
Forner 2018b	Participant data set was included in Darnell 2021 .
Forner 2019	Study assessed the accuracy of LI-RADS criteria.
Giorgio 2005	Study included the same participant population as Giorgio 2007 .

Study	Reason for exclusion
Guiu 2008	2 × 2 table was not reported directly and could not be calculated/extracted based on the reported data.
Guo 2012	2 × 2 table was not reported directly and could not be calculated/extracted based on the reported data. No data on per-patient analysis.
Hanna 2014	Study population included > 5% of participants previously treated for HCC.
Hardie 2010	Study included participants with previously known HCC.
Hardie 2011a	Participant data set was included in Hardie 2011b . No data on per-patient analysis.
Hardie 2011b	No data on per-patient analysis.
Hecht 2006	No data on per-patient analysis.
Heilmaier 2008	No data on per-patient analysis.
Heilmaier 2009	No data on per-patient analysis.
Ichikawa 2021	No data on per-patient analysis.
Kang 2012	2 × 2 table was not reported directly and could not be calculated/extracted based on the reported data.
Kim 2011	No data on per-patient analysis.
Kim 2012	No data on per-patient analysis.
Kim 2021	Study included participants with previously known HCC.
Kondo 2000	Study included participants with previously known HCC.
Lee 2021	Study included participants with previously known HCC.
Li 2015	No data on per-patient analysis.
Li 2018	Study included participants with previously known HCC.
Li 2021	No data on per-patient analysis.
Macarini 2006	No data on per-patient analysis. 2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Macarini 2009	No data on per-patient analysis. 2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Mann 2001	No data on per-patient analysis. 2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Matsuo 2001	No data on per-patient analysis.
Maurea 2014	Target condition was focal liver lesions, not HCC in particular.

Study	Reason for exclusion
Mehra 2012	Target condition was focal liver lesions (benign versus malignant lesions). No data on per-patient analysis.
Miller 2008	Narrative review.
Min 2018b	No data on per-patient analysis.
Mita 2010	No data on per-patient analysis.
Moon 2018	No data on per-patient analysis.
Mori 2002	No data on per-patient analysis for the accuracy of MRI for HCC.
Obuz 2006	Target condition was malignant liver lesions, not HCC in particular.
Pahade 2016	Study population included > 5% of participants previously treated for malignant lesions in general.
Paisant 2020	Study population included > 5% of participants previously treated for HCC. No data on per-patient analysis.
Park 2012	Study population included > 5% of participants previously treated for HCC.
Park 2020	No data on per-patient analysis.
Phongkitkarun 2013	No data on per-patient analysis.
Puig 1997	No data on per-patient analysis.
Qayyum 2006	No data on per-patient analysis.
Rimola 2012	Patient data set is included in Darnell 2021 .
Rofsky 1993	No data on per-patient analysis.
Samoylova 2018	Study included participants with previously known HCC.
Serste 2010	Participants included in Sersté 2012 .
Shankar 2016	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Simon 2005	No data on per-patient analysis.
Snowberger 2007	Study included participants with previously known HCC.
Stocker 2018	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Suárez-Muñoz 2006	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Sugimoto 2015	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Teerasamit 2017	No data on per-patient analysis.
Timofte 2016	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Tsang 2011	Target condition included malignant hepatic lesions, not HCC in particular.

Study	Reason for exclusion
Ueda 1995	No data on per-patient analysis.
Vandecaveye 2009	No data on per-patient analysis.
van Wettere 2019	No data on per-patient analysis.
Xu 2009	No data on per-patient analysis.
Xu 2013	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Yoo 2008	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Zhao 2014	2 × 2 table was not reported directly and could not be calculated/extracted based on reported data.
Zhong 2020	Study included participants with previously known HCC.

HCC: hepatocellular carcinoma; LI-RADS (criteria): Liver Imaging Reporting and Data System criteria; MRI: magnetic resonance imaging.

DATA

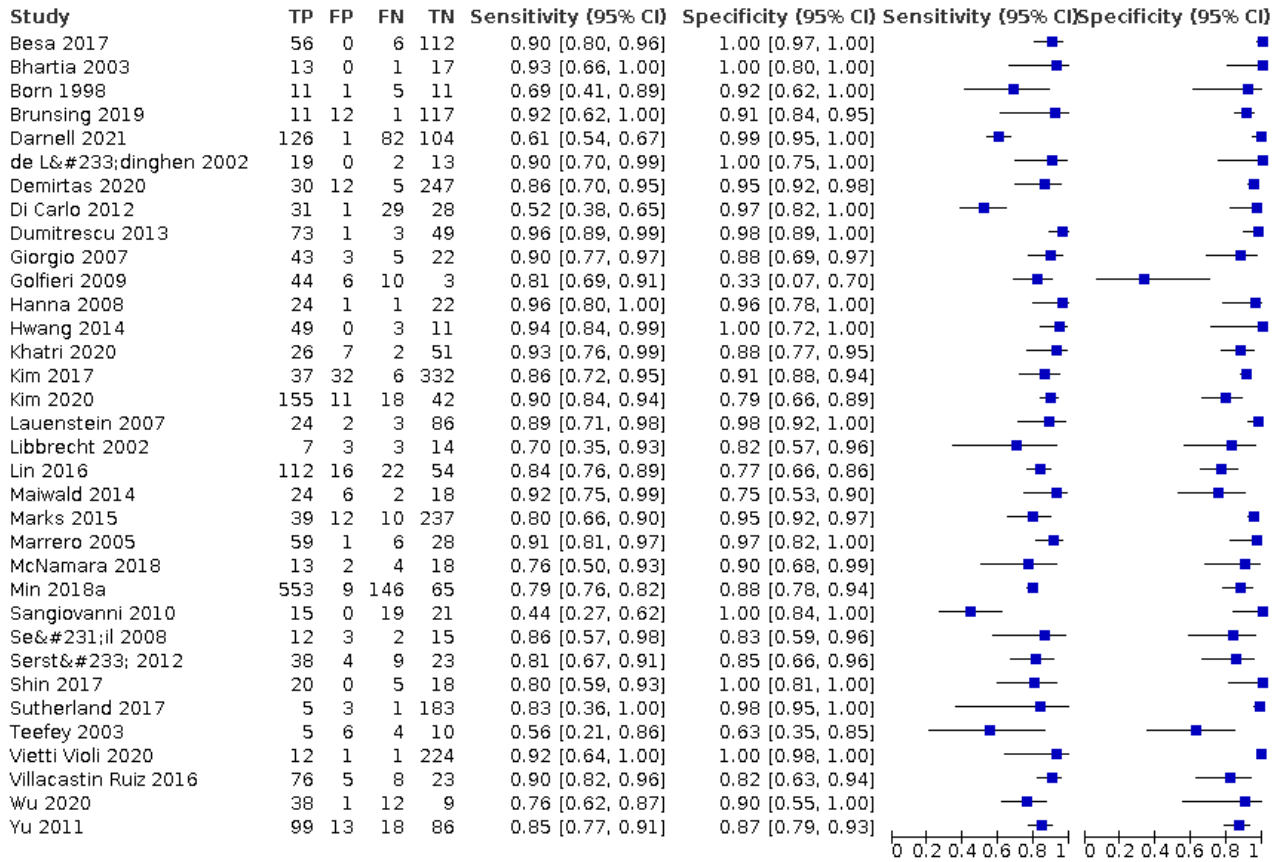
Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 MRI	34	4841
2 Secondary outcome	16	2150

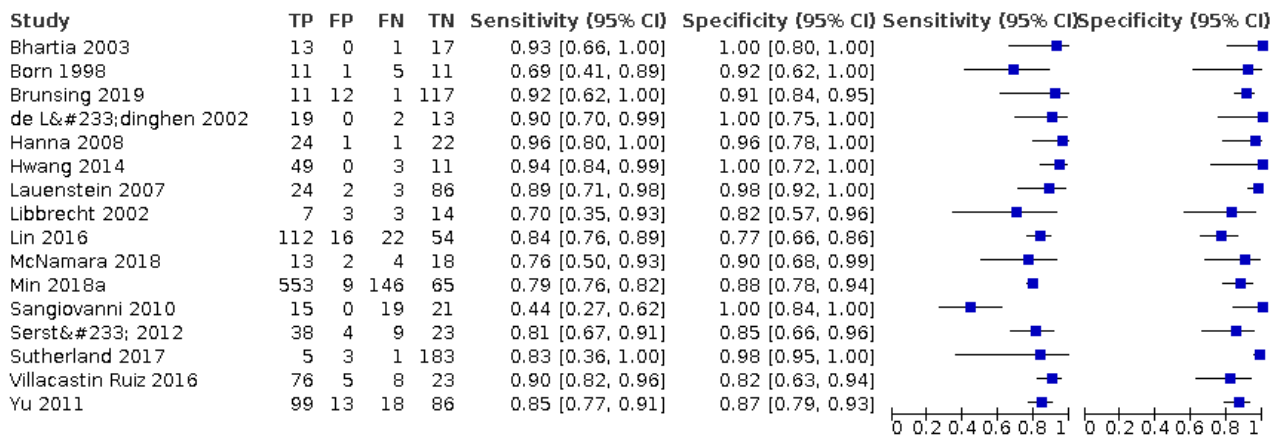
Test 1. MRI

MRI



Test 2. Secondary outcome

Secondary outcome



ADDITIONAL TABLES

Table 1. Post-test probabilities

Pre-test probability		Likelihood ratio	Post-test probability
36%	if MRI positive	13.5 ^a	88%
36%	if MRI negative	0.17 ^b	9%
56%	if MRI positive	13.5 ^a	95%
56%	if MRI negative	0.17 ^b	18%
66%	if MRI positive	13.5 ^a	96%
66%	if MRI negative	0.17 ^b	25%

MRI: magnetic resonance imaging

^aPositive likelihood ratio.

^bNegative likelihood ratio.

Table 2. Heterogeneity and sensitivity analyses for magnetic resonance imaging

Analyses	Studies	No of studies	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	P value	
—	All	34	84.4 (80.1 to 87.9)	93.8 (90.1 to 96.1)	—	
Sensitivity analyses	—	Secondary outcome (resectability 100%)	16	84.3 (77.6 to 89.3)	92.9 (88.3 to 95.9)	—
		Positivity criteria clearly defined	31	83.9 (79.3 to 87.5)	93.8 (89.9 to 96.3)	—
		Reference standard blinded	7	76.8 (66.4 to 84.7)	89.3 (81.5 to 94.0)	—
Subgroups analyses	<i>Publication year</i>	Published before 2011	12	82.4 (72.8 to 89.1)	91.3 (80.6 to 96.4)	0.621
		Published after 2011	22	85.2 (80.2 to 89.0)	94.6 (90.9 to 96.8)	
	<i>Prevalence of cirrhosis</i>	Cirrhosis > 90%	26	83.9 (78.7 to 88.1)	93.3 (89.4 to 95.8)	0.745
		Cirrhosis < 90%	6	87.8 (80.1 to 93.0)	96.8 (85.7 to 99.0)	
	<i>Study location</i>	Europe	15	81.9 (73.0 to 88.4)	93.0 (85.6 to 96.7)	0.833
		America	11	87.1 (82.5 to 90.6)	95.4 (89.5 to 98.0)	
		Asia	8	84.4 (78.5 to 89.0)	92.3 (83.7 to 96.5)	
	<i>Setting</i>	Setting clinical	20	86.8 (82.9 to 90.0)	93.3 (88.1 to 96.3)	0.243
		Setting confirmatory test after screening	8	77.1 (63.7 to 86.6)	91.7 (75.1 to 97.6)	

Table 2. Heterogeneity and sensitivity analyses for magnetic resonance imaging (Continued)

	Setting initial screening test	6	85.0 (78.0 to 90.1)	96.2 (92.2 to 98.2)	
<i>Prevalence of resectable HCC</i>	HCC resectable < 20%	6	86.7 (82.6 to 89.9)	87.6 (70.4 to 95.5)	0.593
	HCC resectable > 20%	20	85.4 (79.5 to 89.9)	93.6 (88.8 to 96.4)	
<i>Reference standard</i>	Histology	5	78.5 (62.8 to 88.8)	93.2 (85.3 to 97.0)	0.620
	OLT	10	88.6 (83.7 to 92.2)	93.4 (87.0 to 96.8)	
	OLT and histology	6	80.1 (76.1 to 83.6)	86.9 (53.4 to 97.5)	
	Mix	13	86.0 (77.8 to 91.5)	95.2 (90.9 to 97.6)	
<i>Prevalence of viral aetiology</i>	Viral < 80%	23	87.6 (83.8 to 90.7)	94.5 (90.7 to 96.8)	0.195
	Viral > 80%	6	74.9 (64.4 to 83.1)	96.5 (68.6 to 99.7)	
<i>LI-RADS positivity criteria</i>	LI-RADS positivity criteria not used	29	84.4 (79.9 to 88.1)	93.0 (88.7 to 95.8)	0.829
	LI-RADS positivity criteria used	5	83.7 (67.3 to 92.8)	96.3 (88.2 to 98.9)	
<i>Type of contrast media</i>	Gd extracellular + SPIO	16	86.2 (80.1 to 90.7)	93.0 (87.7 to 96.1)	0.793
	Gd intracellular	14	82.5 (74.4 to 88.4)	95.5 (88.0 to 98.4)	
<i>Operator expertise</i>	Operator expertise reported	3	75.2 (55.2 to 88.2)	94.4 (85.3 to 98.0)	0.667
	Operator expertise not reported	31	85.2 (80.9 to 88.6)	93.9 (89.8 to 96.4)	

CI: confidence interval; Gd: gadolinium; HCC: hepatocellular carcinoma; LI-RADS: Liver Imaging Reporting and Data System criteria; OLT: orthotopic liver transplantation; SPIO: superparamagnetic iron oxide.

Table 3. Other systematic reviews on diagnostic accuracy of magnetic resonance imaging for hepatocellular carcinoma

Systematic review	Analysistype	No of included studies	No of participants analysed	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Use of bivariate statistical model
Colli 2006	Per-patient	9	498	80.6 (70 to 91)	84.8 (77 to 93)	No
Chen 2014	Per-patient	18	Not reported	91 (90 to 93)	94 (93 to 96)	Unclear
Chou 2015	Per-patient	12	Not reported	86 (79 to 91)	89 (82 to 93)	Yes

Table 3. Other systematic reviews on diagnostic accuracy of magnetic resonance imaging for hepatocellular carcinoma (Continued)

Lee 2015	Per-patient	8	Not reported	88 (83 to 92)	94 (85 to 98)	Unclear
Chan 2021	Per-patient	22	1685	86.8 (83.9 to 89.4)	90.3 (87.3 to 92.7)	No
Gupta 2021	Per-patient	15	2807	86 (84 to 88)	94 (91 to 96)	Unclear
Xie 2011	Per-lesion	8	520	85 (82 to 88)	87 (83 to 91)	—
Chen 2013	Per-lesion	15	Not reported	91 (87 to 94)	95 (93 to 97)	
Junqiang 2014	Per-lesion	11	605	92 (89 to 94)	95 (93 to 97)	
Liu 2013	Per-lesion	10	852	91 (89 to 93)	95 (94 to 96)	
Wu 2013	Per-lesion	10	570	91 (77 to 97)	93 (85 to 97)	
Ye 2015	Per-lesion	9	469	95 (88 to 96)	96 (94 to 97)	
Chou 2015	Per-lesion	20	Not reported	83 (80 to 86)	87 (79 to 93)	
Lee 2015	Per-lesion	33	2489	79 (74 to 83)	Not estimated	
Li 2015a	Per-lesion	5	460	88 (85 to 91)	96 (94 to 97)	
Li 2015b	Per-lesion	15	670	85 (82 to 88)	78 (73 to 83)	
Hanna 2016	Per-lesion	74	Not reported	77.5 (73.1 to 79.3)	Not estimated	
Guo 2016	Per-lesion	12	627	86 (76 to 93)	94 (92 to 96)	
Kierans 2016	Per-lesion	22	1908	78 (68 to 85)	92 (88 to 95)	
Liu 2017	Per-lesion	18	1735	92 (90 to 93)	89 (87 to 91)	
Roberts 2018	Per-lesion	19	Not reported	82 (75 to 87)	91 (82 to 95)	
Li 2019	Per-lesion	8	498	85 (77 to 90)	94 (88 to 97)	
Feng 2021	Per-lesion	8	1002	74 (69 to 78)	93 (77 to 98)	
Gupta 2021	Per-lesion	15	2807	77 (74 to 81)	Not estimated	

CI: confidence interval.

APPENDICES

Appendix 1. Search strategies

Database	Time span	Search strategy
The Cochrane Hepato-Biliary Group Controlled Trials Register (via the Cochrane Register of Studies Web)	9 November 2021	((magnetic resonance or MRI or 'MR imag*' or NMR or gadolinium* or Gadoteric* or ECA or HBA or Gd-EOB* or eovist or primovist) or ((Contrast or radiocontrast or radiopaque) and (agent* or media or medium or material*))) and (((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumor* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or 'focal liver lesion*') and (((advanc* or chronic or end* or terminal* or unresect* or 'late stage*') and (liver* or hepat*)) or cirrho*)
The Cochrane Hepato-Biliary Group Diagnostic Test of Accuracy Studies Register (via the Cochrane Register of Studies Web)	9 November 2021	((magnetic resonance or MRI or 'MR imag*' or NMR or gadolinium* or Gadoteric* or ECA or HBA or Gd-EOB* or eovist or primovist) or ((Contrast or radiocontrast or radiopaque) and (agent* or media or medium or material*))) and (((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumor* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or 'focal liver lesion*') and (((advanc* or chronic or end* or terminal* or unresect* or 'late stage*') and (liver* or hepat*)) or cirrho*)
The Cochrane Library	2021, Issue 11	<p>#1 MeSH descriptor: [Magnetic Resonance Imaging] this term only</p> <p>#2 MeSH descriptor: [Diffusion Magnetic Resonance Imaging] this term only</p> <p>#3 MeSH descriptor: [Multiparametric Magnetic Resonance Imaging] this term only</p> <p>#4 (magnetic resonance or MRI or 'MR imag*' or NMR or gadolinium* or Gadoteric* or ECA or HBA or Gd-EOB* or eovist or primovist):ti,ab,kw</p> <p>#5 ((Contrast or radiocontrast or radiopaque) NEAR/2 (agent* or media or medium or material*)):ti,ab,kw</p> <p>#6 #1 or #2 or #3 or #4 or #5</p> <p>#7 MeSH descriptor: [Carcinoma, Hepatocellular] this term only</p> <p>#8 MeSH descriptor: [Liver Neoplasms] this term only</p> <p>#9 (((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumor* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or 'focal liver lesion*'):ti,ab,kw</p> <p>#10 #7 or #8 or #9</p> <p>#11 (((advanc* or chronic or end* or terminal* or unresect* or 'late stage*') NEAR/4 (liver* or hepat*)) or cirrho*):ti,ab,kw</p> <p>#12 #6 and #10 and #11</p>
MEDLINE Ovid	1946 to 9 November 2021	<p>1. magnetic resonance imaging/ or diffusion magnetic resonance imaging/ or multiparametric magnetic resonance imaging/</p> <p>2. (magnetic resonance or MRI or 'MR imag*' or NMR or gadolinium* or Gadoteric* or ECA or HBA or Gd-EOB* or eovist or primovist).tw,kf.</p> <p>3. ((Contrast or radiocontrast or radiopaque) adj2 (agent* or media or medium or material*)).tw,kf.</p> <p>4. 1 or 2 or 3</p> <p>5. carcinoma, hepatocellular/ or liver neoplasms/</p>

(Continued)

		<p>6. (((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumo* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or 'focal liver lesion*').tw,kf.</p> <p>7. 5 or 6</p> <p>8. (((advanc* or chronic or end* or terminal* or unresect* or 'late stage*') adj4 (liver* or hepat*)) or cirrho*).tw,kf.</p> <p>9. 4 and 7 and 8</p>
Embase Ovid	1974 to 9 November 2021	<p>1. nuclear magnetic resonance imaging/ or diffusion weighted imaging/ or multiparametric magnetic resonance imaging/</p> <p>2. (magnetic resonance or MRI or 'MR imag*' or NMR or gadolinium* or Gadoxetic* or ECA or HBA or Gd-EOB* or eovist or primovist).tw,kw.</p> <p>3. ((Contrast or radiocontrast or radiopaque) adj2 (agent* or media or medium or material*)).tw,kw.</p> <p>4. 1 or 2 or 3</p> <p>5. exp liver cancer/ or liver tumor/</p> <p>6. (((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumo* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or 'focal liver lesion*').tw,kw.</p> <p>7. 5 or 6</p> <p>8. (((advanc* or chronic or end* or terminal* or unresect* or 'late stage*') adj4 (liver* or hepat*)) or cirrho*).tw,kw.</p> <p>9. 4 and 7 and 8</p>
LILACS (Bireme)	1982 to 9 November 2021	<p>(magnetic resonance or MRI or MR imag\$ or NMR or gadolinium\$ or Gadoxetic\$ or ECA or HBA or Gd-EOB\$ or eovist or primovist) or ((Contrast or radiocontrast or radiopaque) and (agent\$ or media or medium or material\$)) [Words] and (((liver or hepato\$) and (carcinom\$ or cancer\$ or neoplasm\$ or malign\$ or tumo\$ or adeno\$ or angiom\$ or sarcoma\$ or angiosarcoma\$)) or HCC\$ or hepatoma\$ or focal liver lesion\$) [Words] and (((advanc\$ or chronic or end\$ or terminal\$ or unresect\$ or late stage\$) and (liver\$ or hepat\$)) or cirrho\$) [Words]</p>
Science Citation Index – Expanded (Web of Science)	1900 to 9 November 2021	<p>#5 (#1 or #2) and #3 and #4</p> <p>#4 TS=(((advanc* or chronic or end* or terminal* or unresect* or (late*stage*)) near (liver* or hepat*)) or cirrho*)</p> <p>#3 TS=(((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumo* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or focal liver lesion*)</p> <p>#2 TS=((Contrast or radiocontrast or radiopaque) NEAR (agent* or media or medium or material*))</p> <p>#1 TS=(magnetic resonance or MRI or MR imag* or NMR or gadolinium* or Gadoxetic* or ECA or HBA or Gd-EOB* or eovist or primovist)</p>
Conference Proceedings Citation Index – Science (Web of Science)	1990 to 9 November 2021	<p>#5 (#1 or #2) and #3 and #4</p> <p>#4 TS=(((advanc* or chronic or end* or terminal* or unresect* or (late*stage*)) near (liver* or hepat*)) or cirrho*)</p>

(Continued)

#3 TS=(((liver or hepato*) and (carcinom* or cancer* or neoplasm* or malign* or tumo* or adeno* or angiom* or sarcoma* or angiosarcoma*)) or HCC* or hepatoma* or focal liver lesion*)

#2 TS=((Contrast or radiocontrast or radiopaque) NEAR (agent* or media or medium or material*))

#1 TS=(magnetic resonance or MRI or MR imag* or NMR or gadolinium* or Gadoxetic* or ECA or HBA or Gd-EOB* or eovist or primovist)

Appendix 2. QUADAS-2

Domain	1. Participant selection	2. Index test	3. Reference standard	4. Flow and timing
Signalling questions and criteria	<p>Q1: "Was a consecutive or random sample of participants enrolled?"</p> <p>Yes – if the study reported on a consecutive or a random selection of participants.</p> <p>No – if the study reported on another form of selection of participants.</p> <p>Unclear – if the study did not report on how the participants were enrolled.</p> <p>Q2: "Did the study avoid inappropriate exclusions?"</p> <p>Yes – if definitions of exclusion criteria were appropriate (i.e. previous surgery or treatment for HCC; people with cholangiocarcinoma) and all exclusions were reported.</p> <p>No – if exclusion criteria were inappropriate and exclusions were not reported.</p> <p>Unclear – if the study did not report causes of exclusions.</p>	<p>Q1: "Were the index test results interpreted without knowledge of the results of the reference standard?"</p> <p>Yes – if the study reported that the results of the index test were interpreted without the knowledge of the results of the reference standard.</p> <p>No – if the study reported that results of the index test were interpreted with the results of the reference standard.</p> <p>Unclear – if the study did not report information about blinding of the results of the index test and reference standard.</p> <p>Q2: "Were positivity criteria clearly defined?"</p> <p>Yes – if the study clearly reported positivity criteria (i.e. non-rim-like hyperenhancement in late arterial phase (defined as arterial phase hyperenhancement) and subsequent non-peripheral washout on portal venous or delayed phases, or both).</p>	<p>Q1: "Is the reference standard likely to correctly classify the target condition?"</p> <p>Yes – if the reference standard correctly defined the presence/absence of HCC (pathology of explanted liver in a transplant cohort).</p> <p>No – if the study used other reference tests than pathology of explanted liver.</p> <p>Unclear – if the study did not report enough information on the reference standard used.</p> <p>Q2: "Were the reference standard results interpreted without the knowledge of the results of the index test?"</p> <p>Yes – if the study reported that the results of the reference standard were interpreted without the knowledge of the results of the index test.</p> <p>No – if the study reported that the results of the reference standard were interpreted with the knowledge of the results of the index test.</p> <p>Unclear – if the study did not report information about blinding of the results of the reference standard and the index test.</p>	<p>Q1: "Was there an appropriate interval between the index test and the reference standard?"</p> <p>Yes – if the study reported the range of intervals between the index test and the reference standard, and the maximum interval was ≤ 3 months.</p> <p>No – if the study reported the range of intervals between the index test and the reference standard, and the maximum interval was > 3 months.</p> <p>Unclear – if the study did not report the range of intervals between the index test and the reference standard.</p> <p>Q2: "Did all participants receive the same reference standard?"</p> <p>Yes – if the study had only 1 reference standard for all the participants.</p> <p>No – if the study had > 1 reference standard.</p> <p>Unclear – if the study information regarding the</p>

(Continued)

No – if the study did not report the positivity criteria.

use of reference standard was unclear.

Q3: "Were all participants included in the analysis?"

Yes – if all participants meeting the inclusion criteria were included in the analysis, or data on all included participants were available so a 2 × 2 table including all included participants could be constructed.

No – not all participants meeting the inclusion criteria were included in the analysis or the 2 × 2 table could not be constructed using data on all included participants.

Unclear – insufficient data were reported to permit a judgement.

<p>Risk of bias</p>	<p><i>Could the selection of participants have introduced bias?</i></p> <p>Low risk: 'Yes' for all signalling questions.</p> <p>High risk: 'No' or 'Unclear' for ≥ 1 signalling question.</p>	<p><i>Could the conduct or interpretation of the index test have introduced bias?</i></p> <p>Low risk: 'Yes' for all signalling questions.</p> <p>High risk: 'No' or 'Unclear' for ≥ 1 signalling question.</p>	<p><i>Could the reference standard, its conduct, or its interpretation have introduced bias?</i></p> <p>Low risk: 'Yes' for all signalling questions.</p> <p>High risk: 'No' or 'Unclear' for ≥ 1 signalling question.</p>	<p><i>Could the participant flow have introduced bias?</i></p> <p>Low risk: 'Yes' for all signalling questions.</p> <p>High risk: 'No' or 'Unclear' for ≥ 1 signalling question.</p>
<p>Concerns about applicability</p>	<p><i>Are there concerns that included participants and setting do not match the review question?</i></p> <p>Low concern: the participants included in the review represent the participants in whom the tests were used in clinical practice (i.e. surveillance programme in people with advanced chronic liver disease; clinical cohort of people with advanced chronic liver disease).</p>	<p><i>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</i></p> <p>Low concern: the index test, its conduct, or its interpretation did not differ from the way it is used in clinical practice.</p> <p>High concern: the index test, its conduct, or its interpretation differed from the way it is used in clinical practice.</p>	<p><i>Are there concerns that the target condition as defined by the reference standard does not match the question?</i></p> <p>High concern: the definition of the target condition as defined by the reference standard did not match the question (i.e. pathology of the explanted liver is feasible only in the case of liver transplant; the natural history and prognosis of HCC detected in explanted liver might be different).</p> <p>Low concern: the definition of the target condition as defined by the reference stan-</p>	<p>—</p>

(Continued)

High concern: the participants included in the review differed from the participants in whom the tests were used in clinical practice (cohort of people with advanced and decompensated liver disease, candidates for orthotopic liver transplantation).

Standard matched the question for all included study participants.

HCC: hepatocellular carcinoma.

HISTORY

Protocol first published: Issue 5, 2021

CONTRIBUTIONS OF AUTHORS

TN: wrote the protocol and performed searches for references, evaluated references for obtaining the full reports, evaluated studies for inclusion, extracted data from studies, assessed the risk of bias, and wrote the final review.

AC: co-ordinated the protocol design and designed the final review.

VG: commented on the protocol and performed searches for references, evaluated references for obtaining the full reports, evaluated studies for inclusion, extracted data from studies, assessed the risk of bias, and wrote the final review.

MF: performed searches for references and critically commented on the review.

GC: wrote the protocol, provided statistical expert opinion, and critically commented on the final review.

CM: critically commented on the protocol and performed searches for references, evaluated references for obtaining the full reports, evaluated studies for inclusion, extracted data from studies, and assessed the risk of bias.

DŠ: critically commented on the protocol, acted as arbiter when review authors could not reach a consensus, and critically commented on the final review.

DM: commented on the protocol and critically commented on the final review.

All authors approved the review for publication.

DECLARATIONS OF INTEREST

TN: none.

AC: I am a Deputy-Coordinating Editor at the CHBG Editorial Team, but as an author, I was not involved in the editorial process.

VG: none.

MF: I am an Editor at the CHBG Editorial Team, but as an author, I was not involved in the editorial process.

GC: I am a Statistical Editor at the CHBG Editorial Team, but as an author, I was not involved in the editorial process.

CM: none.

DŠ: I am an Editor at the CHBG Editorial Team, but as an author, I was not involved in the editorial process.

DM: none.

SOURCES OF SUPPORT

Internal sources

- No sources of support, Other

External sources

- No sources of support, Other

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We did not use Covidence to manage the selection of studies ([Covidence 2020](#)).

We did not perform the planned comparison between groups "positivity criteria clearly defined" compared to "positivity criteria not defined" as a possible source of heterogeneity but, to assess the robustness of our results, we performed a sensitivity analysis including studies with clearly defined positivity criteria only.

We did not perform the planned comparison studies using LI-RADS 5 only as MR positivity criteria compared to studies using LI-RADS 4 and 5 as positivity criteria due to the paucity of studies using these criteria.

We did not perform the planned sensitivity analysis in which studies published only in abstract or letter form are excluded because only one study was published in abstract form ([Di Carlo 2012](#)). All other included studies were published as full-texts.

We did not perform the planned sensitivity analysis in which studies at high risk of bias are excluded as all the included studies were judged to be at high risk of bias.

INDEX TERMS

Medical Subject Headings (MeSH)

*Carcinoma, Hepatocellular [complications] [diagnostic imaging]; Cross-Sectional Studies; *Liver Neoplasms [complications] [diagnostic imaging]; Magnetic Resonance Imaging; Sensitivity and Specificity; Ultrasonography

MeSH check words

Adult; Humans