European Radiology (2019) 29:1022–1031
https://doi.org/10.1007/s00330-018-5557-1

HEPATOBILIARY-PANCREAS

Optimal criteria for hepatocellular carcinoma diagnosis using CT in patients undergoing liver transplantation

Nieun Seo[1] · Myoung Soo Kim[2] · Mi-Suk Park[1] · Jin-Young Choi[1] · Chansik An[1] · Kyunghwa Han[3] · Seung Up Kim[4] · Dong Jin Joo[2] · Myeong-Jin Kim[1]

Received: 12 March 2018 / Revised: 27 April 2018 / Accepted: 23 May 2018 / Published online: 4 July 2018

Abstract

Objective To compare the diagnostic performance of various guidelines for hepatocellular carcinoma (HCC) diagnosis using computed tomography (CT) in patients undergoing liver transplantation (LT).

Methods In total, 216 patients who underwent preoperative CT and subsequent LT were included. Two radiologists retrospectively evaluated focal hepatic lesions independently according to various guidelines and allocated patients according to the Milan criteria. The diagnostic performance of the guidelines was compared using alternative free-response receiver-operating characteristics (AFROC) analysis with bootstrapping. Comparisons of sensitivity, specificity, and accuracy of patient allocation based on the Milan criteria between guidelines were performed using logistic regression with generalized estimating equations (GEE).

Results Fifty-two of 216 patients had 87 HCCs. The reader-averaged figure of merit obtained using AFROC analysis was 0.738 for the AASLD/EASL or KLCSG-NCC guidelines and 0.728 for the LI-RADS v2014 or OPTN/UNOS (bootstrapping, $p = 0.005$). The per-lesion sensitivity for HCCs (all and 1–2-cm lesions) was significantly higher with the AASLD/EASL (37.9–41.4% and 30.8–41.0%) than with LI-RADS (28.7% and 15.4–18.0%) (logistic regression with GEE, $p = 0.008$ and 0.030 for reader 1 and $p = 0.005$ for reader 2). The per-patient specificity (98.8–99.4%) was the same for all guidelines. The accuracy of the Milan criteria was 81.5–83.3% without significant differences among the four guidelines (logistic regression with GEE, $p > 0.05$).

Conclusion AASLD/EASL showed higher diagnostic performance and sensitivity, particularly for 1–2-cm HCCs, and the same specificity with LI-RADS. All guidelines are comparable for patient allocation based on the Milan criteria for LT.

Key Points

• *The overall diagnostic performance of CT for HCC diagnosis was highest with AASLD/EASL.*
• *AASLD/EASL showed higher sensitivity for diagnosis of 1–2-cm HCCs than LI-RADS.*
• *The accuracy of the Milan criteria using CT was comparable among the four guidelines.*

Keywords Carcinoma, hepatocellular · Liver transplantation · ROC curve · Sensitivity and specificity · Tomography, X-ray computed

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5557-1) contains supplementary material, which is available to authorized users.

✉ Mi-Suk Park
radpms@yuhs.ac

[1] Department of Radiology, Severance Hospital, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea

[2] Department of Surgery, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea

[3] Department of Radiology, Yonsei Biomedical Research Institute, Research Institute of Radiological Science, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea

[4] Department of Internal Medicine, Institute of Gastroenterology, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea

Abbreviations

AASLD	American Association for the Study of Liver Diseases
AFROC	Alternative free-response receiver operating characteristics
CT	Computed tomography
EASL	European Association for the Study of the Liver
EORTC	European Organization for Research and Treatment of Cancer
FOM	Figure of merit
GEE	Generalized estimating equation
HCC	Hepatocellular carcinoma
KLCSG-NCC	Korean Liver Cancer Study Group and the National Cancer Center
LI-RADS	Liver Imaging Reporting and Data System
LT	Liver transplantation
MC	Milan criteria
MRI	Magnetic resonance imaging
OPTN	Organ Procurement and Transplant Network
UNOS	United Network for Organ Sharing

Introduction

Liver transplantation (LT) is an ideal treatment option for suitable patients with hepatocellular carcinoma (HCC), because it not only eliminates HCC but also cures the underlying liver disease [1, 2]. Following the introduction of the Milan criteria (one HCC ≤ 5 cm, up to three HCCs ≤ 3 cm), the long-term survival rate for patients with HCC who undergo LT has dramatically increased, with overall survival rates of up to 80%–85%, which are not different from those for patients without HCC [3, 4]. Because the Milan criteria are dependent on the pretransplantation imaging diagnosis, an accurate diagnosis of HCC using high-quality imaging is important.

HCC is the only malignancy that can be confirmed on the sole basis of imaging findings. According to written communication at the 2012 United Network for Organ Sharing (UNOS), more than 95% LT candidates with HCC are diagnosed on the basis of imaging alone, without any confirmatory biopsy [5]. Considering this situation, standardized language and reporting of findings are essentially required. Substantial efforts have resulted in the formulation of several guidelines for the noninvasive diagnosis of HCC on cross-sectional imaging. These include the American Association for the Study of Liver Diseases (AASLD), European Association for the Study of the Liver-European Organization for Research and Treatment of Cancer (EASL-EORTC), Asian-Pacific Association for the Study of the Liver (APASL), Organ Procurement and Transplantation Network (OPTN) system, Liver Imaging Reporting and Data System (LI-RADS), Japan Society of Hepatology (JSH), and Korean Liver Cancer Study Group-National Cancer Center (KLCSG-NCC) guidelines [5–14].

However, all these guidelines differ with regard to their background and criteria, which can result in different staging and treatment decisions. Furthermore, there is no evidence-based agreement on the appropriate guidelines for different conditions. From these perspectives, we designed the present study to investigate and compare the diagnostic performance of the AASLD/EASL, OPTN/UNOS, LI-RADS, and KLCSG-NCC guidelines for the diagnosis of HCC using computed tomography (CT) in at-risk patients requiring LT.

Materials and methods

Study population

This study was approved by our institutional review board, which waived the requirement for informed consent because of the retrospective design. Between January 2007 and December 2014, 493 adult patients who underwent LT were identified from the prospective database at the Department of Surgery in our institution. The inclusion criteria were as follows: performance of dynamic contrast-enhanced liver CT within 3 months before LT and no history of any treatment for HCC before LT. From a total of 493 patients, 277 were excluded for the following reasons: lack of preoperative CT data ($n = 22$), lack of four-phase dynamic liver CT data ($n = 19$), time interval of > 3 months between CT and LT ($n = 25$), presence of intrahepatic cholangiocarcinoma on the explanted liver ($n = 1$), and history of treatment for HCC before LT ($n = 210$). Finally, 216 patients were included in our study (Fig. 1).

CT image acquisition

CT was performed using 16-, 64-, or 128-channel multidetector CT scanners (Siemens Healthineers and GE Healthcare). The routine four-phase dynamic liver CT protocol at our institution includes nonenhanced, late arterial, portal venous, and delayed phases. After precontrast scanning, patients received an intravenous injection of 2.0 ml/kg of iodinated contrast medium, followed by a 20-ml saline bolus at a fixed injection duration of 30 s. Using the bolus-tracking method, late arterial phase images were acquired 18 s after the attenuation value reached 100 HU in the abdominal aorta. The portal venous and delayed phases began with a delay time of 30 s and 150 s after the late arterial and portal venous phases, respectively. The scanning parameters were as follows: 120 kV; 240 mAs; rotation time, 0.5 s; beam pitch, 2; and slice thickness, 3–5 mm.

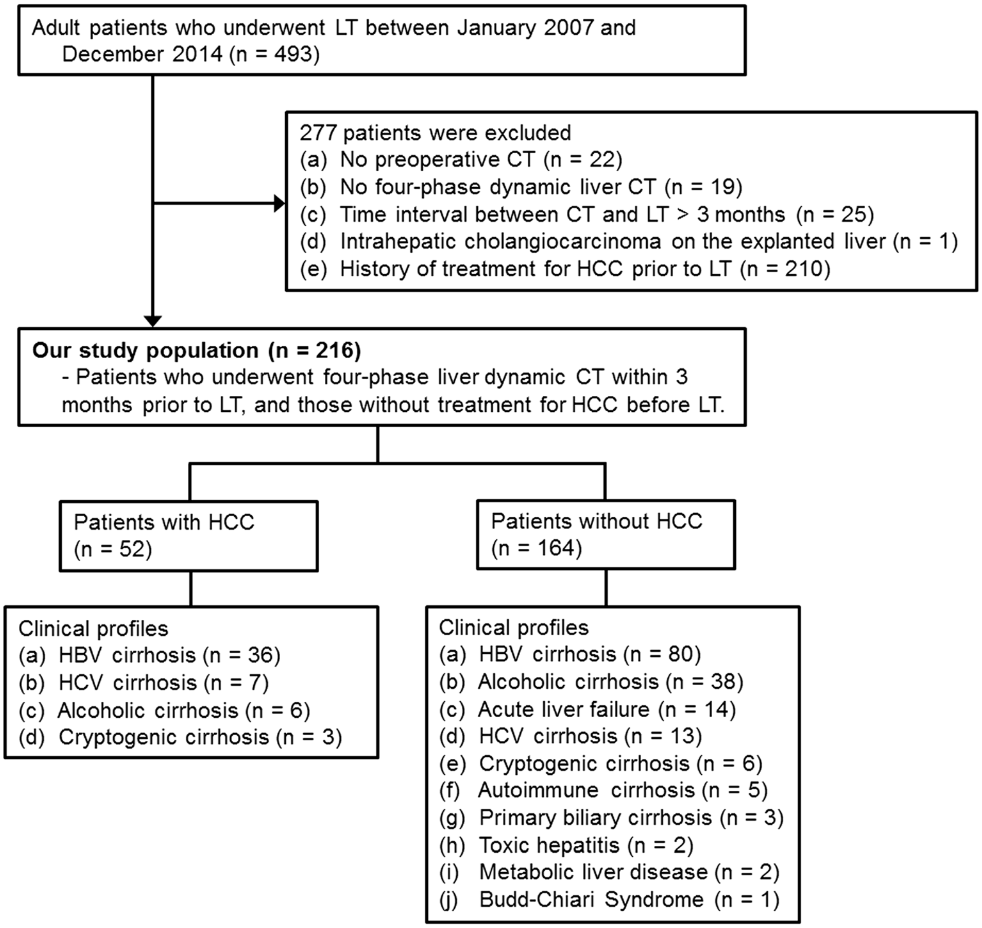

Fig. 1 Flow diagram showing the selection of study participants. LT, liver transplantation; HBV, hepatitis B virus; HCV, hepatitis C virus

Guidelines for HCC diagnosis

In the present study, the diagnostic performances of various guidelines for noninvasive HCC diagnosis were compared. According to the AASLD/EASL-EORTC guidelines, HCC is diagnosed when the lesion (≥ 1 cm) shows arterial hyperenhancement with washout in the venous or delayed phase [8, 9]. The KLCSG-NCC guidelines have set similar criteria for HCC diagnosis. Notably, subcentimeter-sized HCCs can be diagnosed using the KLCSG-NCC guidelines when the lesions fulfill all of the following criteria: typical features of HCC observed with two or more imaging modalities [dynamic CT, magnetic resonance imaging (MRI), and/or gadoxetic acid-enhanced MRI] and a serial increase in serum alpha-fetoprotein (AFP) levels with suppressed hepatitis activity [10, 11]. The LI-RADS and OPTN/UNOS guidelines additionally include capsule appearance and threshold growth as major criteria and subdivide arterial phase-enhancing nodules according to the size, i.e., < 1 cm, 1–2 cm, and ≥ 2 cm. For the diagnosis of HCCs measuring 1–2 cm, at least two of the following three findings should be observed in addition to arterial phase enhancement: washout, capsule/pseudocapsule, and threshold growth. The LI-RADS and OPTN/UNOS guidelines are almost the same, except the fine definition of threshold growth and the exclusion of LR-5us from HCC in the OPTN/UNOS guidelines [5–7, 15]. In LI-RADS v2014, threshold growth is defined by a diameter increase of ≥ 50% in ≤ 6 months, a diameter increase of ≥ 100% in > 6 months, or a newly observed lesion measuring ≥ 1 cm, regardless of the time interval. In contrast, the OPTN/UNOS guidelines solely recognize threshold growth as a diameter increase of ≥ 50% in ≤ 6 months [5–7].

Image analysis

Two board-certified abdominal radiologists (J.Y.C. and C.A.) with 15 and 4 years of experience in liver imaging, respectively, independently reviewed the CT images using a picture archiving and communication system (PACS). Both reviewers were aware that the patients had undergone LT, although they were blinded to other clinical findings and the final histopathologic results for the explanted livers. The radiologists first detected all hepatic focal lesions, except definitely or probably benign lesions such as cysts or typical hemangiomas, and recorded the liver segment, with the image number on PACS for exact matching of the locations. Then, they measured the longest diameter of the focal lesion in the phase in which the lesion was best demarcated [6, 7]. For every focal lesion, the following parameters were assessed for HCC diagnosis according to each guideline: attenuation (hypo-/iso-/hyperattenuation) in the late arterial, portal venous, and delayed phases; presence of a capsule; ancillary features of LI-RADS (mosaic appearance, intralesional fat, and blood products); and threshold growth as defined by the LI-RADS and OPTN/UNOS guidelines [5–7]. Threshold growth was evaluated as a trichotomous variable (present, absent, and no applicable prior examination). Applicable prior examinations were defined as liver dynamic CT or MRI examinations that were obtained at least 3 months prior to the primary CT scans. To diagnose a subcentimeter-sized HCC using the KLCSG-NCC guidelines, the abovementioned parameters were also evaluated on preoperative dynamic liver MRI images, if available. Subsequently, the reviewers rated the possibility of HCC according to each guideline. When the LI-RADS was used, LR-5/LR-5g/LR-5us was considered to represent HCC. Finally, the patients were allocated into one of four groups: no HCC, within the Milan criteria without additional priority, within the Milan criteria with additional priority (single 2–5-cm HCC OR two or three 1–3-cm HCCs), and beyond the Milan criteria according to the OPTN/UNOS guidelines [5, 16].

Pathologic reference standard

Histopathologic evaluation of explanted livers was performed by a senior hepatic pathologist with 25 years of experience in liver pathology. The explanted livers were routinely sectioned into 5–9-mm-thick slices in the axial plane. During macroscopic evaluation, HCCs were suspected when lesions were distinct from the surrounding liver or regenerative nodules in size, color, and texture or exhibited a bulging contour beyond the surface of the liver. If the tumors were small, thinner slices were obtained for exact analysis. All suspected macroscopic nodules identified during gross examination were histopathologically evaluated and matched with preoperative images using a lesion-by-lesion analysis. If a focal hepatic lesion noted on preoperative CT was not detected in the explanted liver, areas that best corresponded with CT findings were cut from the explanted liver and carefully examined. The pathologist examined each lesion on slides stained with hematoxylin and eosin at the level of the largest tumor diameter. After thorough examination, the pathologist reported the tumor location and size, presence of microvascular invasion, and grade of tumor differentiation according to the Edmondson-Steiner grading system [17]. A study coordinator (N.S., 4 years of experience in liver imaging) who did not participate in the image analysis correlated CT findings with pathologic findings based on the size and segment location of lesions on explant. If the difference between radiologically measured size and pathologically measured size of the lesion is < 5 mm, without another similar-sized lesion in the same segment, it was considered to be matched.

Statistical analysis

To evaluate the diagnostic performance of each guideline for the diagnosis of HCC, alternative free-response receiver-operating characteristic (AFROC) analysis was performed using the "RJafroc" package of R version 3.3.3 (R Foundation for Statistical Computing) [18, 19]. The AFROC curve is a plot of the per-lesion sensitivity against the false-positive fraction per patient for each reader and each guideline. It allows the analysis of the multiple lesions per patient unlike the conventional ROC curve. The diagnostic performance assessed using AFROC analysis was represented by the figure of merit (FOM), which is defined as the probability that a lesion is rated higher than the highest rated nonlesion seen on normal CT images. FOM is analogous to the area under the standard ROC curve, which is estimated by calculating the area under the AFROC curve using the trapezoidal rule. Bootstrapping on 1,000 resampling with replacement was used to estimate the 95% confidence interval (CI) for each FOM and to compare FOMs among the four guidelines. The per-lesion and per-patient sensitivity and the per-patient specificity for HCC detection were calculated. The per-lesion sensitivity was also evaluated according to the pathologically measured size (< 1 cm, 1–2 cm, and ≥ 2 cm). Sensitivity and specificity values for the four guidelines were compared using logistic regression with the generalized estimating equation (GEE). The accuracy of the guidelines with regard to patient allocation based on the Milan criteria was also analyzed and was compared using logistic regression with GEE. To adjust the effect by multiple comparisons, Bonferroni correction was applied. The corrected p values were calculated by multiplying the original p values by the number of comparisons. The interobserver agreement for patient allocation based on the Milan criteria using the four guidelines was assessed using κ statistics as follows: κ value < 0.20, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, good; and 0.81–1.00, excellent. For pathologically proven HCCs, the presence of each CT finding was compared according to the size (< 1 cm, 1–2 cm, and ≥ 2 cm) and histologic grade (grade I vs. grades II–III) of the lesion using chi-square or Fisher's exact tests.

All statistical analyses except AFROC analysis were performed using SAS (version 9.4, SAS Institute Inc.). $P < 0.05$ was considered statistically significant. All p values except those described in the Supplementary table are adjusted p values.

Results

Clinical and pathologic characteristics of the study population

The patient demographics and baseline factors are summarized in Table 1. Of 216 patients, 148 were male, and mean

Table 1 Patient demographics

Variable	
Clinical factors ($n = 216$)	
Age (years)	51.4 ± 8.7
Male gender	148 (67.9%)
Etiology of liver disease	
HBV	116 (53.7%)
HCV	20 (9.3%)
Alcoholic	44 (20.4%)
Others	36 (16.7%)
Serum AFP (ng/ml)	87.8 ± 389.0
Child-Pugh class	
A	47 (21.8%)
B	74 (34.3%)
C	95 (44.0%)
Type of liver transplantation	
Living donor	127 (58.8%)
Deceased donor	89 (41.2%)
Pathologic factors (no. of HCC = 87)	
Microvascular invasion	11 (12.6%)
Diameter of largest tumor (cm)	2.1 ± 1.0
Tumor size (cm)	1.8 ± 1.0
< 1	24 (27.6%)
1 ≤ < 2	39 (44.8%)
≥ 2	24 (27.6%)
Number of HCCs per patient (total number of patients with HCCs = 52)	1.7 ± 1.1
1	33 (63.5%)
2	10 (19.2%)
3	3 (5.8%)
4	5 (9.6%)
5	1 (1.9%)
Edmondson-Steiner grading	
I	47 (54.0%)
II	37 (42.5%)
III	3 (3.4%)
Classification of explanted liver based on Milan criteria ($n = 216$)	
No HCC	164 (75.9%)
Within Milan criteria without priority	22(10.2%)
Within Milan criteria with priority	19 (8.8%)
Beyond Milan criteria	11 (5.1%)

Data are expressed as means ± standard deviations

Data in parentheses are percentages. Percentages may not sum to 100% because of rounding off

HBV, hepatitis B virus; HCV, hepatitis C virus; AFP, alpha-fetoprotein

age of patients was 51.4 years. The most common cause of chronic liver disease was hepatitis B virus (HBV) infection (116/216, 53.7%). In total, 52 of 216 patients had 87 HCCs with 33 exhibiting a single lesion and 19 exhibiting multiple lesions. From the 87 HCCs, 24 measured < 1 cm, 39 measured 1–2 cm, and 24 measured ≥ 2 cm. The Edmondson-Steiner grade was mostly I or II (84/87, 96.5%). When the Milan criteria were applied to the explanted livers, 22 patients were categorized in the within the Milan criteria without an additional priority group, 19 in the within the Milan criteria with an additional priority group, and 11 in the beyond the Milan criteria group. For 52 patients with HCCs, previous CT or MR examinations to assess threshold growth were available in 48 patients; 34 of 52 patients had preoperative MRI within 3 months prior to LT for the evaluation using the KLCSG-NCC guidelines. Additionally, serial AFP values were available for 39 of 52 patients with HCCs.

Diagnostic performance of the different guidelines for HCC diagnosis

The reader-averaged FOM was 0.738 (95% CI, 0.670–0.805) for the AASLD/EASL-EORTC guidelines, 0.738 (95% CI, 0.671–0.804) for the KLCSG-NCC guidelines, 0.728 (95% CI, 0.663–0.793) for the LI-RADS v2014 guidelines, and 0.728 (95% CI, 0.658–0.799) for the OPTN/UNOS guidelines. There was a significant difference between the AASLD/EASL-EORTC/KLCSG-NCC and the LI-RADS/OPTN/UNOS guidelines (bootstrapping, $p = 0.005$). For both readers, there was no difference in the diagnostic performance between the AASLD/EASL-EORTC and KLCSG-NCC guidelines and between the LI-RADS and OPTN/UNOS guidelines. The per-lesion and per-patient sensitivity and per-patient specificity values are demonstrated in Table 2. For both readers, the per-lesion sensitivity of the AASLD/EASL-EORTC and KLCSG-NCC guidelines (37.9%–41.4%) was significantly higher than that of the LI-RADS and OPTN/UNOS guidelines (28.7%; $p = 0.008$ for reader 1 and $p = 0.005$ for reader 2, using logistic regression with GEE). The per-patient sensitivity and specificity were comparable among guidelines (logistic regression with GEE, $p > 0.05$).

The per-lesion sensitivity for HCC detection according to the pathologic size of HCCs is summarized in Table 3. The pathologically measured size of the lesions was < 1 cm in 24, 1–2-cm in 39, and ≥ 2 cm in 24 patients. Among the 24 subcentimeter-sized HCCs, 2 and 3 HCCs were measured as > 10 mm on CT by reader 1 and reader 2, respectively, and they were diagnosed as HCCs according to the AASLD/EASL-EORTC and KLCSG-NCC guidelines. Using the KLCSG-NCC guidelines, two- and one-subcentimeter-sized nodules showed arterial enhancement and washout on both CT and MRI by reader 1 and reader 2, respectively. However, as all of these patients did not show a serial increase of serum AFP, these subcentimeter-sized nodules were not considered to be HCC according to the KLCSG-NCC guidelines. For both readers, there was no significant difference

Table 2 Diagnostic performance of the AASLD/EASL, KLCSG-NCC, LI-RADS, and OPTN/UNOS guidelines for the detection of HCC

	AASLD/EASL		KLCSG-NCC		LI-RADS		OPTN/UNOS		*p* value*
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	
Per-lesion sensitivity									
Reader 1	37.9 (33/87)	29.3–47.4	37.9 (33/87)	29.3–47.4	28.7 (25/87)	20.9–38.1	28.7 (25/87)	20.9–38.1	0.008
Reader 2	41.4 (36/87)	32.1–51.4	41.4 (36/87)	32.1–51.4	28.7 (25/87)	20.5–38.6	28.7 (25/87)	20.5–38.6	0.005
Per-patient sensitivity									
Reader 1	50.0 (26/52)	36.7–63.3	50.0 (26/52)	36.7–63.3	40.4 (21/52)	28.0–54.1	40.4 (21/52)	28.0–54.1	0.116
Reader 2	48.1 (21/52)	35.0–61.5	48.1 (21/52)	35.0–61.5	38.4 (20/52)	26.3–52.2	38.4 (20/52)	26.3–52.2	0.312
Per-patient specificity									
Reader 1	99.4 (163/164)	95.8–99.9	99.4 (163/164)	95.8–99.9	99.4 (163/164)	95.8–99.9	99.4 (163/164)	95.8–99.9	NA
Reader 2	98.8 (162/164)	95.3–99.7	98.8 (162/164)	95.3–99.7	98.8 (162/164)	95.3–99.7	98.8 (162/164)	95.3–99.7	NA

*The *p* values are obtained by the comparison of parameters for either of AASLD/EASL or KLCSG-NCC AND either LI-RADS or OPTN/UNOS, because there was no difference in the diagnostic performance between AASLD/EASL and KLCSG-NCC and between LI-RADS and OPTN/UNOS. These *p* values are obtained by using logistic regression with generalized estimating equations.

AASLD, American Association for the Study of Liver Diseases; EASL, European Association for the Study of the Liver; KLCSG-NCC, Korean Liver Cancer Study Group and the National Cancer Center; LI-RADS, Liver Imaging Reporting and Data System; OPTN, Organ Procurement and Transplant Network; UNOS, United Network for Organ Sharing; CI, confidence interval; NA, not applicable

between guidelines for HCCs measuring < 1 cm (0%–12.5%) and those measuring ≥ 2 cm (70.8%–79.2%). In contrast, for 1–2-cm HCCs, the sensitivity of the AASLD/EASL-EORTC and KLCSG-NCC guidelines (30.8%–41.0%) was significantly higher than that of the LI-RADS and OPTN/UNOS guidelines (15.4%–18.0%; $p = 0.030$ for reader 1 and $p = 0.005$ for reader 2 using logistic regression with GEE; Fig. 2).

Accuracy of the different guidelines for patient allocation based on the Milan criteria

The accuracies of the different guidelines for patient allocation based on the pretransplantation Milan criteria are demonstrated in Table 4. The accuracy of classifying patients into the beyond Milan criteria (36.4%) and within Milan criteria with

Table 3 Per-lesion sensitivity of the AASLD/EASL, KLCSG-NCC, LI-RADS, and OPTN/UNOS guidelines according to the pathologic size of HCCs

Pathologic size (cm)†	AASLD/EASL		KLCSG-NCC		LI-RADS		OPTN/UNOS		*p* value*
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	
< 1									
Reader 1	8.3 (2/24)‡	2.5–24.7	8.3 (2/24)	2.5–24.7	0 (0/24)	0–0	0 (0/24)	0–0	0.577
Reader 2	12.5 (3/24)‡	4.6–29.8	12.5 (3/24)	4.6–29.8	0 (0/24)	0–0	0 (0/24)	0–0	0.238
1 ≤ < 2									
Reader 1	30.8 (12/39)	19.7–44.5	30.8 (12/39)	19.7–44.5	15.4 (6/39)	7.8–28.1	15.4 (6/39)	7.8–28.1	0.030
Reader 2	41.0 (16/39)	29.7–53.4	41.0 (16/39)	29.7–53.4	18.0 (7/39)	9.3–31.8	18.0 (7/39)	9.3–31.8	0.005
≥ 2									
Reader 1	79.2 (19/24)	59.4–90.8	79.2 (19/24)	59.4–90.8	79.2 (19/24)	59.4–90.8	79.2 (19/24)	59.4–90.8	NA
Reader 2	70.8 (17/24)	50.7–85.1	70.8 (17/24)	50.7–85.1	75.0 (18/24)	55.0–88.0	75.0 (18/24)	55.0–88.0	> 0.999

*The *p* values are obtained by the comparison of parameters for either of AASLD/EASL or KLCSG-NCC AND either LI-RADS or OPTN/UNOS, because there was no difference in the diagnostic performance between AASLD/EASL and KLCSG-NCC and between LI-RADS and OPTN/UNOS. These *p*-values are obtained by using logistic regression with generalized estimating equations.

† This was the pathologically determined size of HCC, which could be different from the size of HCC that was measured on CT images. Radiologically measured sizes were used for evaluation of the guidelines.

‡ The pathologic size of all of these nodules was all 9 mm; however readers measured them on CT images as >10 mm. Therefore, they could be evaluated using AASLD/EASL guidelines.

AASLD, American Association for the Study of Liver Diseases; EASL, European Association for the Study of the Liver; KLCSG-NCC, Korean Liver Cancer Study Group and the National Cancer Center; LI-RADS, Liver Imaging Reporting and Data System; OPTN, Organ Procurement and Transplant Network; UNOS, United Network for Organ Sharing; CI, confidence interval; NA, not applicable

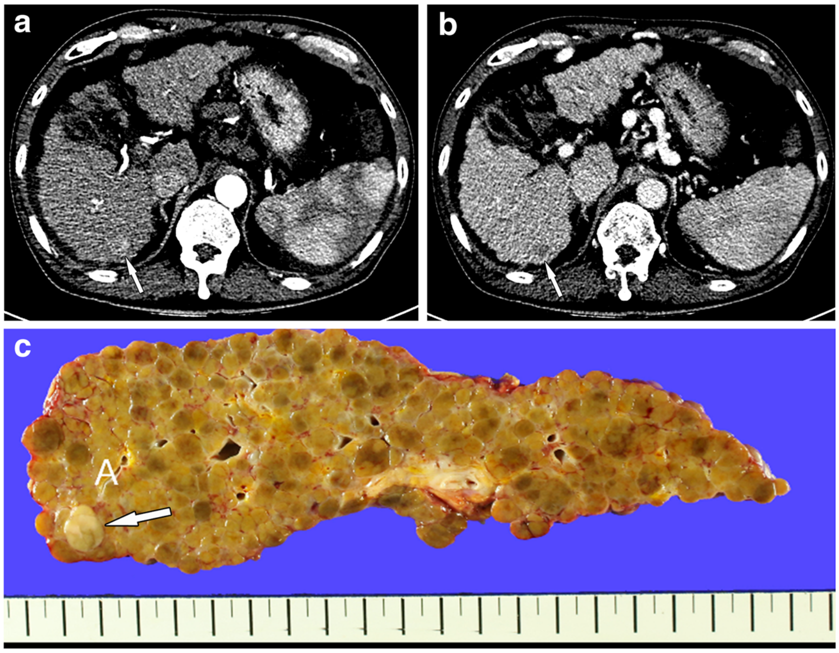

Fig. 2 A 1.3-cm pathologically proven HCC lesion in a 61-year-old male. (**a**) An axial computed tomography (CT) image obtained in the late arterial phase shows a 1.3-cm enhancing nodule in segment VI of the liver (arrow). (**b**) CT image obtained in the portal venous phase shows washout of the arterial phase-enhancing nodule (arrow). On the basis of preoperative CT, this nodule was diagnosed as HCC using the AASLD/EASL and KLCSG-NCC guidelines. However, it was graded as LR-4 using LI-RADS v2014, and the OPTN/UNOS guidelines indicated that it was not HCC. (**c**) Histopathologic analysis of the explanted liver specimen confirms a 1.3-cm grade II (Edmondson-Steiner) HCC in segment VI (arrow)

Table 4 Accuracy of the AASLD/EASL, KLCSG-NCC, LI-RADS, and OPTN/UNOS guidelines for patient allocation based on the Milan criteria

	AASLD/EASL		KLCSG-NCC		LI-RADS		OPTN/UNOS		*p* value*
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	
Beyond MC									
Reader 1	36.4 (4/11)	14.3–66.1	36.4 (4/11)	14.3–66.1	36.4 (4/11)	14.3–66.1	36.4 (4/11)	14.3–66.1	NA
Reader 2	36.4 (4/11)	14.3–66.1	36.4 (4/11)	14.3–66.1	36.4 (4/11)	14.3–66.1	36.4 (4/11)	14.3–66.1	NA
MC with priority									
Reader 1	52.6 (10/19)	31.1–73.2	52.6 (10/19)	31.1–73.2	52.6 (10/19)	31.1–73.2	52.6 (10/19)	31.1–73.2	NA
Reader 2	52.6 (10/19)	31.1–73.2	52.6 (10/19)	31.1–73.2	52.6 (10/19)	31.1–73.2	52.6 (10/19)	31.1–73.2	NA
MC without priority									
Reader 1	13.6 (3/22)	0–28.0	13.6 (3/22)	0–28.0	0 (0/22)	0–0	0 (0/22)	0–0	0.372
Reader 2	9.1 (2/22)	0–21.1	9.1 (2/22)	0–21.1	0 (0/22)	0–0	0 (0/22)	0–0	0.828
No HCC									
Reader 1	99.4 (163/164)	95.8–99.9	99.4 (163/164)	95.8–99.9	99.4 (163/164)	95.8–99.9	99.4 (163/164)	95.8–99.9	NA
Reader 2	98.2 (161/164)	94.5–99.4	98.2 (161/164)	94.5–99.4	98.8 (162/164)	95.3–99.7	98.8 (162/164)	95.3–99.7	>0.999
Overall accuracy									
Reader 1	83.3 (180/216)	77.8–87.7	83.3 (180/216)	77.8–87.7	81.9 (177/216)	76.2–86.5	81.9 (177/216)	76.2–86.5	0.487
Reader 2	81.9 (177/216)	76.2–86.5	81.9 (177/216)	76.2–86.5	81.5 (176/216)	75.7–86.1	81.5 (176/216)	75.7–86.1	> 0.999

*The *p* values are obtained by the comparison of parameters for either of AASLD/EASL or KLCSG-NCC AND either LI-RADS or OPTN/UNOS, because there was no difference in the diagnostic performance between AASLD/EASL and KLCSG-NCC and between LI-RADS and OPTN/UNOS. These *p* values are obtained by using logistic regression with generalized estimating equations. Data in parentheses indicate the number of radiologic allocations divided by the number of pathologic allocations

MC, Milan criteria; AASLD, American Association for the Study of Liver Diseases; EASL, European Association for the Study of the Liver; KLCSG-NCC, Korean Liver Cancer Study Group and the National Cancer Center; LI-RADS, Liver Imaging Reporting and Data System; OPTN, Organ Procurement and Transplant Network; UNOS, United Network for Organ Sharing; CI, confidence interval; NA, not applicable additional priority (52.6%) groups was not different among the four guidelines according to both readers. However, the accuracy of classifying patients in the within Milan criteria without priority group was higher with the AASLD/EASL-EORTC and KLCSG-NCC (9.1% and 13.6% for readers 1 and 2, respectively) guidelines than with the LI-RADS and OPTN/UNOS (0% for both readers) guidelines, without significant difference. The overall accuracy for patient allocation based on the Milan criteria was 81.5%–83.3%, with no significant differences among the four guidelines according to both readers (logistic regression with GEE, $p > 0.05$). The interobserver agreement for patient allocation was excellent for all guidelines, with the following κ values: AASLD/EASL-EORTC, 0.858; KLCSG-NCC, 0.874; LI-RADS, 0.907; and OPTN/UNOS, 0.907.

CT findings according to the pathologic size and histologic grade of HCCs

CT findings according to the pathologic size and histologic grade of HCCs were analyzed and are summarized in Supplementary Tables 1 and 2, respectively. For both readers, there was no significant difference regarding all CT findings according to the pathologic size (chi-square test or Fisher's exact test, $p > 0.05$). Capsule appearance was seldom observed in 1–2-cm HCCs on CT (0/39 for reader 1 and 1/39 for reader 2), while it was present in a small number of HCCs > 2 cm (16.7%, 4/24 for both readers). Ancillary features of LI-RADS evaluated in this study were rarely present on the CT images. Regarding the histologic grade, arterial hyperenhancement was more frequently observed for grade II–III HCCs than for grade I HCCs according to both readers (Fisher's exact test, $p < 0.001$ for reader 1 and $p = 0.006$ for reader 2). There were no significant differences in other CT findings according to the histologic grade of HCCs.

Discussion

In the present study, we assessed and compared the diagnostic performance of the AASLD/EASL-EORTC, KLCSG-NCC, LI-RADS v2014, and OPTN/UNOS guidelines for the diagnosis of HCC and evaluation of the eligibility for LT using MDCT. The overall diagnostic performance and per-lesion sensitivity for HCC diagnosis was significantly higher for the AASLD/EASL-EORTC and KLCSG-NCC guidelines than for the LI-RADS and OPTN/UNOS guidelines. Our results correspond well with the results from a recent prospective study, which showed that LI-RADS v2014 was not superior to the AASLD guidelines for the diagnosis of small (1–3 cm) HCCs in high-risk patients [20]. In particular, for the detection of 1–2-cm HCCs, which accounted for the largest proportion of HCCs in the present study (44.8%), the AASLD/EASL-EORTC and KLCSG-NCC guidelines showed a significantly higher sensitivity (30.8%–41.0%) compared with the LI-RADS and OPTN/UNOS guidelines (15.4%–18.0%), with no differences in the specificity. The overall diagnostic accuracy of patient allocation based on the Milan criteria was comparable among all guidelines and ranged from 81.5% to 83.3%.

In terms of per-lesion sensitivity, the AASLD/EASL-EORTC and KLCSG-NCC guidelines were superior to the LI-RADS and OPTN/UNOS guidelines. In general, the AASLD/EASL-EORTC guidelines are simpler and easier to apply compared with the other guidelines. The AASLD/EASL-EORTC guidelines for HCC diagnosis have been validated in many studies, although studies including LT recipients are rare [21–24]. Our results showed that the AASLD/EASL-EORTC guidelines can be implemented along with MDCT for HCC diagnosis before LT, with good diagnostic performance. The recently established KLCSG-NCC guidelines have not been widely validated yet. Of note, the KLCSG-NCC guidelines specify criteria for subcentimeter-sized HCCs [11, 12]. However, subcentimeter-sized nodules showing a typical enhancement pattern of HCC on both of CT and MRI could not be diagnosed as HCC, because the serial increase in serum AFP levels was not detected in these patients. Additionally, as the detection of subcentimeter-sized nodules is difficult on CT, sophisticated criteria were not useful for the additional detection of subcentimeter-sized HCCs in the present study. Compared with CT, the overall per-lesion sensitivity of MRI has been reported to be higher, especially for the subcentimeter-sized HCCs [25]. From this perspective, the KLCSG-NCC guidelines could be helpful to diagnose subcentimeter-sized HCCs on MRI, although this factor was not evaluated in the present study.

In the LI-RADS v2014 and OPTN/UNOS guidelines, criteria for the diagnosis of 1–2-cm HCCs are more stringent than those in the AASLD/EASL-EORTC and KLCSG-NCC guidelines. In the LI-RADS v2014, 1–2-cm arterial phase-enhancing nodules can be considered as HCC when CT shows at least two of the following three findings: washout, capsule, and threshold growth [7]. In the present study, approximately 50% HCCs measuring 1–2 cm were missed when the LI-RADS and OPTN/UNOS guidelines were used. Most of the missed 1–2-cm HCCs, using LI-RADS, were categorized as LR-4, because these nodules showed only arterial enhancement and washout. The low sensitivity for the detection of 1–2-cm HCCs can be a disadvantage of these guidelines, considering that 1–2-cm HCCs account for a considerable proportion of HCCs in explanted livers [26, 27].

Our results revealed that a capsule appearance was seldom observed in HCCs measuring ≤ 2 cm on CT. Although a capsule appearance was present in a small number of HCCs measuring > 2 cm (16.7%), this feature was not additionally beneficial for the diagnosis of HCC, because all these lesions showed arterial hyperenhancement and venous-delayed phase washout. One previous study using MRI showed results concordant with those of our study [28]. A capsule appearance did not add sensitivity or specificity to the contrast enhancement pattern for the diagnosis of small HCCs (≤ 2 cm) on CT. Threshold growth is another major finding adapted only in the LI-RADS and OPTN/UNOS guidelines. Excluding the single LR-5g observation, threshold growth did not additionally contribute to the diagnosis of HCC in the present study. Therefore, capsule appearance and threshold growth rarely enhance the diagnostic performance of guidelines for HCC detection in patients undergoing LT. Moreover, ancillary features of LI-RADS were rarely detected on CT in LT candidates, although these features did not affect the diagnosis of HCC (LR-5). The per-patient specificity of the LI-RADS and OPTN/UNOS guidelines was similar to that of the AASLD/EASL-EORTC and KLCSG-NCC guidelines. Because the specificity of the AASLD/EASL-EORTC guidelines is almost 100% [8, 9, 29], it seems that there are few additional benefits in terms of the specificity of other guidelines.

Our study has several limitations. First, accurate lesion-by-lesion matching in the explanted liver is known to be particularly demanding. However, the routine plane of histologic sections is the same as the CT imaging plane (axial) at our institution, and our study coordinator correlated CT images with explanted liver specimens, based on the size and the segment of the lesions, following strict criteria. Second, target lesions defined in the AASLD/EASL-EORTC guidelines were not considered in the present study. These guidelines define target lesions as nodules (≥ 1 cm) that are initially detected by surveillance ultrasound. However, we could not strictly adhere to this criterion because our study patients were retrospectively recruited from the LT database and surveillance ultrasound was not available for most patients. Finally, we solely evaluated the diagnostic performance of CT according to various guidelines. It has been proven that MRI has higher sensitivity than CT for HCC diagnosis [25, 30–32]. However, the use of MRI is often limited for LT candidates, because a poor general condition and large ascites volume deteriorate the MR image quality [33, 34].

In conclusion, among the AASLD/EASL-EORTC, KLCSG-NCC, LI-RADS v2014, and OPTN/UNOS guidelines, the AASLD/EASL-EORTC guidelines may be the most useful for HCC diagnosis on CT in patients requiring LT because of the relatively simple diagnostic criteria and higher sensitivity without any compromise in the specificity.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Mi-Suk Park.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry Kyunghwa Han from Yonsei Biomedical Research Institute, who is one of the coauthors, performed statistical analysis.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- diagnostic study
- performed at one institution

References

1. Clavien PA, Lesurtel M, Bossuyt PM, Gores GJ, Langer B, Perrier A (2012) Recommendations for liver transplantation for hepatocellular carcinoma: an international consensus conference report. Lancet Oncol 13:e11–e22
2. Yu SJ (2016) A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010-2016. Clin Mol Hepatol 22:7–17
3. Mazzaferro V, Regalia E, Doci R et al (1996) Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med 334:693–699
4. Befeler AS, Hayashi PH, Di Bisceglie AM (2005) Liver transplantation for hepatocellular carcinoma. Gastroenterology 128: 1752–1764
5. Wald C, Russo MW, Heimbach JK, Hussain HK, Pomfret EA, Bruix J (2013) New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. Radiology 266:376–382
6. Mitchell DG, Bruix J, Sherman M, Sirlin CB (2015) LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. Hepatology 61:1056–1065
7. American College of Radiology. Liver imaging reporting and data system (LI-RADS). American College of Radiology. Web site. http://www.acr.org/Quality-Safety/Resources/LIRADS/LIRADS-v2014. Accessed June 1, 2017
8. Bruix J, Sherman M (2011) Management of hepatocellular carcinoma: an update. Hepatology 53:1020–1022
9. European Association for the Study of the Liver, European Organisation for Research and Treatment of Cancer (2012) EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 56:908–943
10. Lee JM, Park JW, Choi BI (2014) 2014 KLCSG-NCC Korea Practice Guidelines for the management of hepatocellular carcinoma: HCC diagnostic algorithm. Dig Dis 32:764–777
11. Korean Liver Cancer Study Group (KLCSG) and National Cancer Center, Korea (NCC) (2015) 2014 Korean Liver Cancer Study Group-National Cancer Center Korea practice guideline for the management of hepatocellular carcinoma. Korean J Radiol 16:465–522
12. Korean Society of Abdominal Radiology (2017) Diagnosis of Hepatocellular Carcinoma with Gadoxetic Acid-Enhanced MRI: 2016 Consensus Recommendations of the Korean Society of Abdominal Radiology. Korean J Radiol 18:427–443
13. Kokudo N, Hasegawa K, Akahane M et al (2015) Evidence-based Clinical Practice Guidelines for Hepatocellular Carcinoma: The Japan Society of Hepatology 2013 update (3rd JSH-HCC Guidelines). Hepatol Res 45

14. Omata M, Lesmana LA, Tateishi R et al (2010) Asian Pacific Association for the Study of the Liver consensus recommendations on hepatocellular carcinoma. Hepatol Int 4:439–474
15. Tang A, Fowler KJ, Chernyak V, Chapman WC, Sirlin CB (2017) LI-RADS and transplantation for hepatocellular carcinoma. Abdom Radiol (NY). https://doi.org/10.1007/s00261-017-1210-8
16. Lee DH, Lee JM, Baek JH, Shin CI, Han JK, Choi BI (2015) Diagnostic performance of gadoxetic acid-enhanced liver MR imaging in the detection of HCCs and allocation of transplant recipients on the basis of the Milan criteria and UNOS guidelines: correlation with histopathologic findings. Radiology 274:149–160
17. Edmondson HA, Steiner PE (1954) Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. Cancer 7:462–503
18. Chakraborty DP, Berbaum KS (2004) Observer studies involving detection and localization: modeling, analysis, and validation. Med Phys 31:2313–2330
19. Zhai X, Chakraborty DP (2015). RJafroc: Analysis of data acquired using the receiver operating characteristic paradigm and its extensions. R package version 0.1.1. https://CRAN.R-project.org/package=RJafroc
20. Ronot M, Fouque O, Esvan M, Lebigot J, Aube C, Vilgrain V (2017) Comparison of the accuracy of AASLD and LI-RADS criteria for the non-invasive diagnosis of HCC smaller than 3cm. J Hepatol. https://doi.org/10.1016/j.jhep.2017.12.014
21. Manini MA, Sangiovanni A, Fornari F et al (2014) Clinical and economical impact of 2010 AASLD guidelines for the diagnosis of hepatocellular carcinoma. J Hepatol 60:995–1001
22. Kim TK, Lee KH, Jang HJ et al (2011) Analysis of gadobenate dimeglumine-enhanced MR findings for characterizing small (1-2-cm) hepatic nodules in patients at high risk for hepatocellular carcinoma. Radiology 259:730–738
23. Aube C, Oberti F, Lonjon J et al (2017) EASL and AASLD recommendations for the diagnosis of HCC to the test of daily practice. Liver Int 37:1515–1525
24. Pahwa A, Beckett K, Channual S, Tan N, Lu DS, Raman SS (2014) Efficacy of the American Association for the Study of Liver Disease and Barcelona criteria for the diagnosis of hepatocellular carcinoma. Abdom Imaging 39:753–760
25. Lee YJ, Lee JM, Lee JS et al (2015) Hepatocellular carcinoma: diagnostic performance of multidetector CT and MR imaging-a systematic review and meta-analysis. Radiology 275:97–109
26. Ronzoni A, Artioli D, Scardina R et al (2007) Role of MDCT in the diagnosis of hepatocellular carcinoma in patients with cirrhosis undergoing orthotopic liver transplantation. AJR Am J Roentgenol 189:792–798
27. Luca A, Caruso S, Milazzo M et al (2010) Multidetector-row computed tomography (MDCT) for the diagnosis of hepatocellular carcinoma in cirrhotic candidates for liver transplantation: prevalence of radiological vascular patterns and histological correlation with liver explants. Eur Radiol 20:898–907
28. Rimola J, Forner A, Tremosini S et al (2012) Non-invasive diagnosis of hepatocellular carcinoma </= 2 cm in cirrhosis. Diagnostic accuracy assessing fat, capsule and signal intensity at dynamic MRI. J Hepatol 56:1317–1323
29. Forner A, Vilana R, Ayuso C et al (2008) Diagnosis of hepatic nodules 20 mm or smaller in cirrhosis: Prospective validation of the noninvasive diagnostic criteria for hepatocellular carcinoma. Hepatology 47:97–104
30. Rostambeigi N, Taylor AJ, Golzarian J et al (2016) Effect of MRI Versus MDCT on Milan Criteria Scores and Liver Transplantation Eligibility. AJR Am J Roentgenol 206:726–733
31. Addley HC, Griffin N, Shaw AS et al (2011) Accuracy of hepatocellular carcinoma detection on multidetector CT in a transplant liver population with explant liver correlation. Clin Radiol 66: 349–356
32. Park MS, Kim S, Patel J et al (2012) Hepatocellular carcinoma: detection with diffusion-weighted versus contrast-enhanced magnetic resonance imaging in pretransplant patients. Hepatology 56: 140–148
33. Heimbach J, Kulik LM, Finn R et al (2017) Aasld guidelines for the treatment of hepatocellular carcinoma. Hepatology. https://doi.org/10.1002/hep.29086
34. Roberts LR, Sirlin CB, Zaiem F et al (2017) Imaging for the Diagnosis of Hepatocellular Carcinoma: a Systematic Review and Meta-analysis. Hepatology. https://doi.org/10.1002/hep.29487