https://doi.org/10.1016/j.hpb.2019.04.007 HPB

ORIGINAL ARTICLE

Expanding the Liver Imaging Reporting and Data System (LI-RADS) v2018 diagnostic population: performance and reliability of LI-RADS for distinguishing hepatocellular carcinoma (HCC) from non-HCC primary liver carcinoma in patients who do not meet strict LI-RADS high-risk criteria

Daniel R. Ludwig[1], Tyler J. Fraum[1], Roberto Cannella[2,3], Richard Tsai[1], Muhammad Naeem[1], Maverick LeBlanc[1], Amber Salter[4], Allan Tsung[5], Jaquelyn Fleckenstein[6], Anup S. Shetty[1], Amir A. Borhani[2], Alessandro Furlan[2] & Kathryn J. Fowler[7]

[1]Mallinckrodt Institute of Radiology, Washington University School of Medicine, 510 S. Kingshighway Blvd, Campus Box 8131, Saint Louis, MO 63104, [2]Department of Radiology, University of Pittsburgh Medical Center, 3708 Fifth Ave, Pittsburgh, PA, 15213, USA, [3]University of Palermo, Palermo, Italy, [4]Division of Biostatistics, Washington University School of Medicine, 660 S. Euclid Ave., CB 8067, St. Louis, MO 63110, [5]Department of Surgery, The Ohio State University Medical Center, N924 Doan Hall, 410 W 10h Ave, Columbus, OH 43210, [6]Department of Gastroenterology, Washington University School of Medicine, 660 S Euclid Ave., St. Louis, MO 63110, and [7]Department of Radiology, University of California San Diego, 200 W Arbor Dr., San Diego, CA, 92103, USA

Abstract

Background: Hepatocellular carcinoma (HCC) can be diagnosed using imaging criteria in patients at high-risk for HCC, according to Liver Imaging Reporting and Data System (LI-RADS) guidelines. The aim of this study was to determine the diagnostic performance and inter-rater reliability (IRR) of LI-RADS v2018 for differentiating HCC from non-HCC primary liver carcinoma (PLC), in patients who are at increased risk for HCC but not included in the LI-RADS 'high-risk' population.

Methods: This retrospective HIPAA-compliant study included a 10-year experience of pathologically-proven PLC at two liver transplant centers, and included patients with non-cirrhotic hepatitis C infection, non-cirrhotic non-alcoholic fatty liver disease, and fibrosis. Two readers evaluated each lesion and assigned an overall LI-RADS diagnostic category, additionally scoring all major, LR-M, and ancillary features.

Results: The final study cohort consisted of 27 HCCs and 104 non-HCC PLC in 131 patients. The specificity of a 'definite HCC' designation was 97% for reader 1 and 100% for reader 2. The IRR was fair for overall LI-RADS category and substantial for most major features.

Conclusion: In a population at increased risk for HCC but not currently included in the LI-RADS 'high-risk' population, LI-RADS v2018 demonstrated very high specificity for distinguishing pathologically-proven HCC from non-HCC PLC.

Received 22 December 2018; accepted 12 April 2019

Correspondence

Daniel R. Ludwig, 510 S. Kingshighway Blvd, Campus Box 8131, Saint Louis, MO 63104, USA. E-mail: ludwigd@wustl.edu

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common cancer and a leading cause of cancer-related mortality worldwide.[1] Important risk factors for HCC include cirrhosis, chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV), and non-alcoholic steatohepatitis (NASH).[2] Major societies recommend routine surveillance for HCC in patients with cirrhosis of any etiology and non-cirrhotic HBV infection. However, there is little consensus regarding the screening of patients with non-cirrhotic HCV, non-cirrhotic nonalcoholic fatty liver disease (NAFLD), and mild to moderate hepatic fibrosis.[3,4] Indeed, screening programs must consider survival benefit, cost effectiveness, and healthcare priorities, which tend to have a high degree of regional variability.[5]

However, diagnostic algorithms need not apply only to the formally targeted screening population, as long as the pre-test

probability of the disease is sufficiently increased.[3] The Liver Imaging Reporting and Data System (LI-RADS) was developed by the American College of Radiology (ACR) to standardize the diagnosis of HCC through the application of a diagnostic algorithm and a standardized lexicon.[6] LI-RADS specifically distinguishes between the screening population and the diagnostic population, the latter of which also includes patients with current or prior HCC, or recipients of orthotopic liver transplantation (OLT).[7] In the LI-RADS diagnostic population, a diagnosis of HCC on imaging may be sufficient to qualify a patient for HCC model for end stage liver disease (MELD) exception points, conferring priority on the OLT waiting list.[8] This serves to minimize or eliminate the role of percutaneous biopsy, which has a major adverse event rate of approximately 1% in patients with advanced chronic liver disease.[9] Differentiating HCC from non-HCC primary liver carcinoma (PLC) such as intrahepatic cholangiocarcinoma (iCCA) and combined hepatocellular-cholangiocarcinoma (cHCC-CCA) is of primary importance given the poor outcomes after OLT with iCCA and cHCC-CCA.[10,11]

In clinical practice, it is a common scenario for a patient with some variety of chronic liver disease but without any LI-RADS-defined HCC high-risk factors to present with a malignant-appearing liver mass on imaging. There is scant evidence to support the use of the LI-RADS diagnostic algorithm in patients with these weaker risk factors, such as non-cirrhotic HCV, non-cirrhotic NAFLD, and hepatic fibrosis. For example, among patients with non-cirrhotic NAFLD, recent reports suggest that HCC may be more likely to have an infiltrative appearance[12] and, therefore, less likely to display classic features such as nonperipheral "washout" or an enhancing "capsule".[13]

Thus, the aim of this study was to determine the diagnostic performance and inter-rater reliability (IRR) of LI-RADS version 2018 (v2018) for distinguishing HCC from non-HCC in patients with pathologically-proven PLC who have risk factors for HCC but do not satisfy the strict LI-RADS criteria for inclusion in the 'high-risk' population.

Methods

Study design

This study was performed in a retrospective fashion in compliance with the Health Insurance Portability and Accountability Act (HIPAA), at two large liver transplant centers. At both centers, the institutional review board (IRB) waived the requirement for informed consent.

Study cohort

Pathologic diagnoses served as the standard of reference. At both institutions, the pathology databases were queried to identify all liver specimens logged between August 2007 and July 2017 with final diagnoses containing at least one of the following terms: hepatocellular carcinoma, cholangiocarcinoma, biphenotypic, and hepato-cholangiocarcinoma. Pathology reports often use 'biphenotypic' and 'hepato-cholangiocarcinoma' interchangeably, so both terms were used to identify all combined hepatocellular-cholangiocarcinoma (cHCC-CCA). A diagnosis of cHCC-CCA was based on morphologic features on routine histopathology with hematoxylin and eosin. Additional immunohistochemical testing was performed at the discretion of the interpreting pathologist, and supportive features included keratin 7 and 19 positivity, as well as biliary canalicular expression of CD10 and pCEA.[14,15] For this study, cHCC-CCA was considered a non-HCC PLC.

Pathology reports were reviewed by authors uninvolved in image interpretation to identify candidate liver masses for imaging review. Lesions were excluded if the tissue received by pathology was deemed inadequate to make a final pathologic diagnosis or if the final diagnosis was inconclusive (e.g., poorly differentiated carcinoma or adenocarcinoma not otherwise specified). In patients with multiple lesions satisfying the inclusion criteria, only the largest lesion was selected for LI-RADS assessment, as the largest lesion most commonly guides initial patient management. Following identification of candidate lesions, a subset of HCCs from this same period was selected at random to achieve a number of HCCs equal to one third of the total number of cases. This was performed to facilitate a more robust analysis of non-HCC PLC, due the high frequency of HCCs relative to non-HCC PLC among patients with chronic liver disease. Note that the ratio of HCCs to non-HCCs was lower in final cohort due to differential exclusion of patients on the basis of underling high-risk status (i.e., many HCCs occurred in patients considered 'high-risk' by LI-RADS).

For the lesions meeting the above criteria, relevant clinical history and imaging studies were reviewed, again by authors uninvolved in image interpretation. Lesions without a clear radiologic correlate or without imaging prior to locoregional therapy (LRT) were excluded. Additionally, lesions were excluded if the patient did not undergo a liver-protocol magnetic resonance imaging (MRI) or computed tomography (CT) that satisfied the Organ Procurement and Transplantation Network (OPTN) technical requirements.[8] CT and MRI was performed at both participating institutions according to previously published protocols.[16,17] Given the 10-year interval from which eligible studies were identified, there was minor year-to-year modifications, however these protocols were generally representative of our scanning techniques. If multiple imaging studies were available, the study immediately prior to tissue acquisition or before the first LRT, if performed, was selected for LI-RADS assessment.

Clinical, pathologic, laboratory, and imaging data were used to identify patients with chronic liver disease who have increased risk of HCC but are not currently within the LI-RADS diagnostic population, as the goal of the study was to assess the applicability of LI-RADS in this patient population. Patients

satisfying the formal LI-RADS definition of 'high-risk' for HCC, specifically those with cirrhosis and/or chronic HBV, were excluded (Fig. 1). Hepatic fibrosis was preferentially assessed using histopathology of background liver tissue; if none was available (i.e., biopsy only included mass), patients were considered cirrhotic if the liver demonstrated unequivocal imaging findings of cirrhosis according to the interpreting radiologist or laboratory values were suggestive of cirrhosis (see Fig. 1 caption for more details). Additionally, patients were excluded if they had cirrhosis secondary to a vascular disorder. Patients were included if they had non-cirrhotic HCV, non-cirrhotic NAFLD, and/or fibrosis without frank cirrhosis. The degree of HCC risk conferred by hepatic steatosis without superimposed non-alcoholic steatohepatitis (NASH) or fibrosis is unknown; however, we elected to include such cases due to the established link between non-cirrhotic NALFD and HCC, and the known spatial and temporal heterogeneity of inflammation associated with steatosis.[12,18]

LI-RADS assessment

Image interpretation was performed by two fellowship-trained abdominal radiologists (KJF and ASS) with 7 and 3 years of post-fellowship experience, serving as reader 1 (R1) and reader 2 (R2), respectively. Readers were blinded to most clinical information, such as the original imaging interpretation and pathologic diagnosis, but did have access to patient age and gender. Readers also had access to information from prior studies, when available, to permit assessment of threshold growth. The presence of the lesion as a discrete nodule on antecedent ultrasound was provided to the reviewer. Readers were directed to the lesion of interest by means of a series/image number, liver segment, and additional spatial identifying information when multiple lesions were present. Readers evaluated only the lesion of interest and did not score additional lesions. Each lesion was scored with respect to all major, LR-M, and ancillary features and assigned an overall LI-RADS category. Readers applied tie-breaking rules and category adjustments according to LI-RADS methodology.

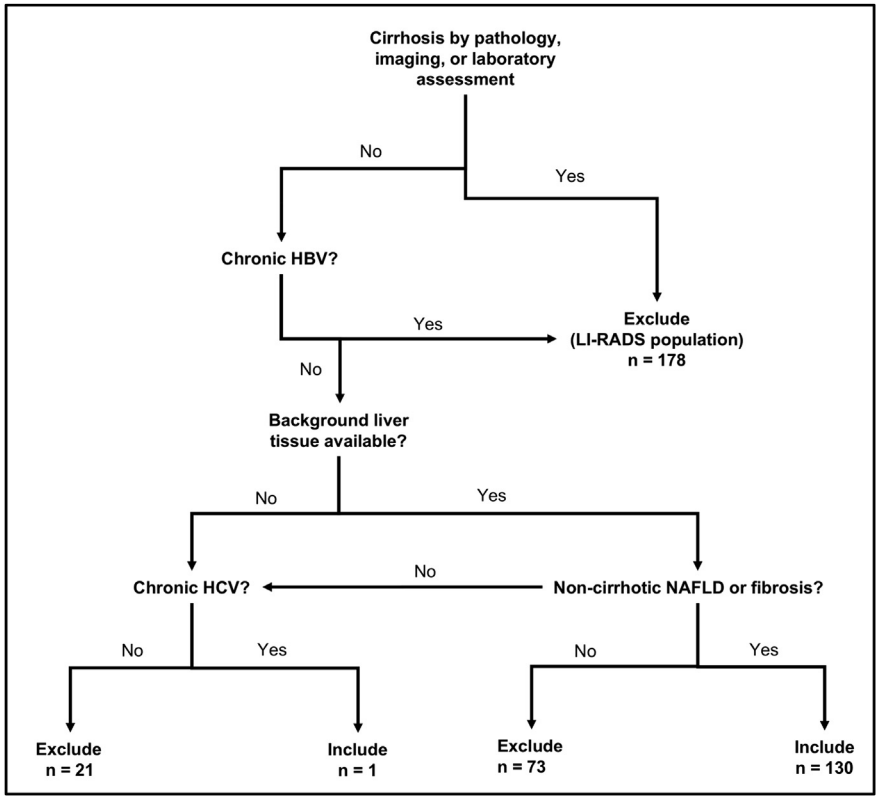

Figure 1 Algorithm for patient inclusion. Patients with chronic hepatitis B viral (HBV) infection and/or cirrhosis by pathology, imaging, or laboratory assessment were excluded, as these patients represent the LI-RADS 'at-risk' population. Cirrhosis was determined using background liver tissue, if available. If no background liver tissue was available (i.e., biopsy tissue included only mass), patients were considered cirrhotic if the original interpreting radiologist identified unequivocal imaging findings of cirrhosis. However, due to the limited sensitivity of imaging to diagnose cirrhosis,[37,38] laboratory values (if available) were used to calculate a FIB-4 score in patients without cirrhosis by imaging. A FIB-4 score greater than 3.25 was used to identify cirrhosis, as a FIB-4 greater than 3.25 has a 97% specificity for the identification of advanced fibrosis.[39] Patients were included if they had chronic hepatitis C viral infection (HCV), non-cirrhotic non-alcoholic fatty liver disease (NAFLD), and/or fibrosis

Readers assigned overall LI-RADS category according to the v2017 methodology; however, given the recent update to LI-RADS (v2018), a LI-RADS v2018 score was generated from the reader-provided data by an author uninvolved in image interpretation using the new definition of threshold growth and change in major feature criteria of LR-5 for 10–19 mm observations.[6,19] Notably, no observations changed category with application of the v2018 criteria. The LI-RADS v2018 score was used for all analyses.

Statistical analysis

Descriptive statistics were used to summarize the imaging features using means ± standard deviation (SD) for continuous variables and frequency (percentage) for categorical variables. Confidence intervals for sensitivity and specificity were calculated according to the methods in Mercaldo *et al.*[20] Positive and negative predictive value were not assessed due to our methodologic decision to enrich our population for non-HCC PLC. Cohen κ test was used to assess the IRR for categorical variables, and intraclass correlation coefficient was used to assess the IRR for continuous variables. Agreement was scored as poor (<0.00), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), or almost perfect (0.81–1.00).[21] Differences in frequencies of categorical variables (e.g., frequency of LI-RADS features) between HCC and non-HCC PLC, were assessed using the Pearson χ^2 or Fisher exact test. Correction for multiple comparisons was performed to achieve a false discovery rate of 5%, using the methods of Benjamini and Hochberg[22]; $p < 0.02$ was indicative of a significant difference. All statistical analyses were performed using R Studio (version 1.1.456, R Development Core Team, New Zealand).

Results

Study cohort

Query of the pathology database and subsequent random exclusion of HCCs to limit their number to one third of the total cases resulted in 571 candidate liver specimens. Of these, 168 (29%) were eliminated based on predefined exclusion criteria, most commonly a lack of liver-protocol CT or MRI (100 of 168, 60%) or no intraparenchymal mass on imaging (e.g., extrahepatic cholangiocarcinoma; 32 of 168, 19%). Additionally, 2 (1%) liver specimens were excluded due to their occurrence in cirrhosis secondary to a vascular etiology. Of the remaining candidate masses, 178 (44%) were excluded due to high-risk status (i.e. LI-RADS population), and 94 (23%) were excluded due to lack of previously mentioned risk factors for HCC, specifically a lack of chronic HCV by history or pathologic evidence of NAFLD and/or fibrosis, as shown in Fig. 1. The final cohort included 131 masses in 131 patients (Table 1). Age was 65.2 ± 10.6 years (mean ± SD), and 50% of patients were female (n = 66). The most common histopathologic risk factor was cryptogenic fibrosis (n = 42, 32%), followed by NASH with fibrosis (n = 33, 25%). Because we enriched our population for non-HCC PLC, most lesions were iCCA (n = 84, 64%). HCC comprised 27 (21%) of the included masses. In this population of patients without cirrhosis, the source of tissue for pathologic diagnosis was most commonly resection (26 [96%] of HCC; 97 [93%] of non-HCC). Few lesions were treated for LRT prior to pathologic diagnosis (6 [22%] among HCC; 11 [11%] among non-HCC). HCC was more frequently characterized on MRI (22 of 27, 81%), whereas non-HCC PLC was more frequently characterized on CT (53 of 104, 51%).

Diagnostic performance of LI-RADS v2018 for differentiating HCC from Non-HCC primary liver carcinoma

The numbers of lesions in each LI-RADS category stratified by reader and pathologic diagnosis are shown in Fig. 2, and the diagnostic performance of LI-RADS v2018 by reader is shown in Table 2. Specificity of LR-5 as a predictor of HCC was 97% and 100% for R1 and R2, respectively. Supplementary Figure 1 shows a representative HCC scored as LR-5 by both readers. All three of the false positive LR-5 observations for R1 were iCCA (100%); R2 had no false positive LR-5 observations. Supplementary Figure 2 shows an example iCCA scored as LR-5 and LR-4 by R1 and R2, respectively. The combination of LR-5 or LR-TIV (definitely due to HCC) as a predictor of HCC did not change specificity (97% and 100% for R1 and R2, respectively), as no non-HCC PLC were scored as LR-TIV (definitely due to HCC). Sensitivity of LR-5 as a predictor of HCC was limited, 67% and 37% for R1 and R2, respectively.

Sensitivity of LR-M as a predictor for non-HCC PLC was high, 91% and 84% for R1 and R2, respectively. Combining LR-M or LR-TIV (may be due to non-HCC malignancy) as a predictor of non-HCC PLC increased sensitivity to 95% and 98%, respectively. Supplementary Figure 3 shows a representative iCCA scored by both readers as LR-M. Specificity of LR-M as a predictor of non-HCC PLC was more limited, 78% and 44% for R1 and R2, respectively. Supplementary Figure 4 shows an example HCC scored by both readers as LR-M.

Inter-rater reliability of LI-RADS categories

Table 3 shows the distribution of overall LI-RADS categories assigned by R1 and R2, with the results of the IRR analysis. Agreement for overall LI-RADS category was fair (κ of 0.37). However, agreement for LR-5 versus other categories was moderate (κ of 0.53). Similarly, agreement for LR-M versus other categories, and LR-5 versus LR-M or LR-TIV (i.e., likely malignant but not eligible for OPTN exception points) was moderate (κ of 0.45 and 0.55, respectively). Agreement on LR-5 or LR-TIV (definitely HCC) versus LR-M or LR-TIV (may be due to non-HCC malignancy) was fair (κ of 0.38).

Table 1 Patient and mass characteristics

Patient Characteristics (N = 131)	
Gender	N (%)
Male	65 (50%)
Female	66 (50%)
Age	Mean ± SD (range) in years
All	65.2 ± 10.6 (28–88)
Male	66.7 ± 10.2 (28–88)
Female	63.8 ± 10.9 (31–84)
Etiology of chronic liver disease	N (%)
Steatosis without fibrosis	25 (19%)
Steatosis with fibrosis	10 (7%)
NASH without fibrosis	8 (6%)
NASH with fibrosis	33 (25%)
Hepatitis C without fibrosis	1 (1%)
Hepatitis C with fibrosis	5 (4%)
EtOH with fibrosis	2 (2%)
PSC with fibrosis	5 (4%)
Cryptogenic fibrosis	42 (32%)
Mass Characteristics (N = 131)	
Pathologic diagnosis	N (%)
Hepatocellular carcinoma (HCC)	27 (21%)
Hepatocellular-cholangiocarcionoma (cHCC-CCA)	20 (15%)
Intrahepatic cholangiocarcinoma (iCCA)	84 (64%)
Source of tissue for pathologic diagnosis	N (%)
Hepatocellular carcinoma (N = 27)	–
Biopsy	0 (0%)
Resection	26 (96%)
Explant	1 (4%)
cHCC-CCA and iCCA (N = 104)	–
Biopsy	6 (6%)
Resection	97 (93%)
Explant	1 (1%)
LRT between imaging and pathology	N (%)
Hepatocellular carcinoma (n = 27)	6 (22%)
cHCC-CCA and iCCA (n = 104)	11 (11%)
Imaging modality for LI-RADS	N (%)
Hepatocellular carcinoma (n = 27)	–
MRI	22 (81%)
CT	5 (19%)
cHCC-CCA and iCCA (n = 104)	–
MRI	51 (49%)
CT	53 (51%)

Abbreviations: CT – computed tomography; EtOH – alcohol; LRT – locoregional therapy; MRI – magnetic resonance imaging; N – number; NAFLD – non-alcoholic fatty liver disease; NASH – non-alcoholic steatohepatitis; PSC – primary sclerosing cholangitis; SD – standard deviation.

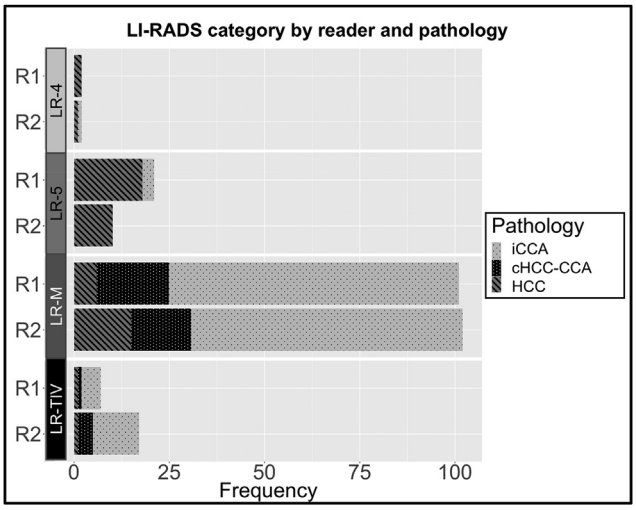

Figure 2 Number of observations in each LI-RADS category, stratified by reader and pathologic diagnosis. No observations were scored as LR-1, LR-2, or LR-3. Observations scored as LR-TIV were further stratified as definitely due to HCC, probably due to HCC, or may be due to non-HCC malignancy (not shown)

Frequency and inter-rater reliability of LI-RADS major, LR-M, and ancillary features

Supplementary Table 1 shows the frequencies of major features by reader among HCCs versus non-HCC PLC, along with results from the IRR analysis. Agreement on nonrim arterial phase hyperenhancement (APHE), nonperipheral "washout", and enhancing "capsule" was substantial (κ of 0.64, 0.62, and 0.63, respectively), and, as expected, all of these features were significantly more common among HCCs. Agreement for size was also substantial (intraclass correlation coefficient of 0.74). Notably, there was no difference in size between HCC and non-HCC PLC. Agreement for tumor in vein was only moderate (κ of 0.46).

Frequencies of LR-M features by reader among HCCs versus non-HCC PLC, along with results from the IRR analysis, are shown in Supplementary Table 2. Targetoid mass, rim APHE, and delayed central enhancement were the most frequently observed LR-M criteria, and agreement on these features was fair to moderate. Peripheral "washout" was more commonly identified by R1 compared to R2 (54% vs. 10%), and agreement was only slight. Agreement on the remaining LR-M features was poor to fair; however, these features were infrequently present. Supplementary Table 3 shows the frequencies of ancillary features favoring malignancy by reader among HCC versus non-HCC PLC, along with the results of the IRR analysis. IRR analysis was not performed for ancillary features favoring benignity due to the rarity of these features among malignant lesions.

Discussion

We have shown that LI-RADS v2018 is capable of differentiating HCC from non-HCC with a high degree of specificity in patients with pathologically-proven PLC and non-cirrhotic HCV, non-cirrhotic NAFLD, and hepatic fibrosis without cirrhosis. The specificity of LR-5 reported in this study (97% and 100% for R1 and R2, respectively) is moderately higher than that previously reported for 'high-risk' patients using v2014,[16] and slightly higher than that recently reported for 'high-risk' patients using v2018.[23,24] The degree of improvement from v2014 is likely, in part, due to revision of the diagnostic criteria for HCC with v2017, specifically the introduction of a distinction between nonrim APHE (as a major feature for HCC) and rim APHE (as criteria for LR-M) and the addition of the LR-TIV category.[19] The sensitivity of LR-5 for HCC, on the other hand, was limited (67% and 37% for R1 and R2, respectively), an expected finding given the design of the LI-RADS algorithm, which seeks to maximize specificity in order to avoid improper allocation of a

Table 2 Diagnostic Performance of LI-RADS v2018 by Reader for Differentiating HCC from Non-HCC Primary Liver Carcinoma

LI-RADS Category	Sensitivity (95% CI)	Specificity (95% CI)
LR-5 as a predictor of HCC	–	–
R1	66.7 (49.1–83.5)	97.1 (92.7–99.4)
R2	37.0 (21.6–57.6)	100 (97.2–100)
LR-5 or LR-TIV (definitely due to HCC) as a predictor for HCC	–	–
R1	70.3 (52.9–86.2)	97.1 (92.7–99.4)
R2	37.0 (21.6–57.6)	100 (97.2–100)
LR-M as a predictor for non-HCC	–	–
R1	91.1 (85.4–96.0)	77.8 (60.8–91.4)
R2	83.7 (76.5–90.2)	44.4 (28.0–64.7)
LR-M or LR-TIV (may be due to non-HCC malignancy) as a predictor for non-HCC	–	–
R1	95.2 (90.2–98.4)	77.8 (60.8–91.4)
R2	98.1 (94.1–99.8)	40.7 (24.8–61.2)

Abbreviations: CI - confidence interval; HCC – hepatocellular carcinoma; LI-RADS – Liver Imaging Reporting and Data System; LR-5 – definitely HCC; LR-M – probably or definitely malignant but not HCC specific; TIV – tumor in vein.

Table 3 Overall agreement by category

LI-RADS Category for R1	LI-RADS Category for R2							κ Value[a]	Agreement
	LR-4	LR-5	LR-M	LR-TIV (probably HCC)	LR-TIV (def. HCC)	LR-TIV (may be non-HCC)	All		
LR-4	0	0	2	0	0	0	2	–	–
LR-5	1	9	10	0	1	0	21	–	–
LR-M	1	1	89	1	9	0	101	–	–
LR-TIV (probably HCC)	0	0	0	0	2	0	2	–	–
LR-TIV (def. HCC)	0	0	0	0	1	0	1	–	–
LR-TIV (may be non-HCC)	0	0	1	0	3	0	4	–	–
All	2	10	102	1	16	0	131	0.37 (0.18, 0.56)	Fair
Agreement on LR-5 versus Other Categories									
	LI-RADS Category for R2								
LI-RADS Category for R1	LR-5				Other				Agreement
LR-5	9				12			–	–
Other	1				109			0.53 (0.29–0.77)	Moderate
Agreement on LR-M versus Other Categories									
	LI-RADS Category for R2								
LI-RADS Category for R1	LR-M				Other				Agreement
LR-M	89				12			–	–
Other	13				17			0.45 (0.26–0.65)	Moderate
Agreement on LR-5 versus LR-M or LR-TIV									
	LI-RADS Category for R2								
LI-RADS Category for R1	LR-5				LR-M or LR-TIV				Agreement
LR-5	9				11			–	–
LR-M or LR-TIV	1				106			0.55 (0.31–0.79)	Moderate
Agreement on LR-5 or LR-TIV (def. HCC) versus LR-M or LR-TIV (may be non-HCC malignancy)									
	LI-RADS Category for R2								
LI-RADS Category for R1	LR-5 or LR-TIV (def. HCC)				LR-M or LR-TIV (non-HCC)				Agreement
LR-5 or LR-TIV (def. HCC)	11				10			–	–
LR-M or LR-TIV (non-HCC)	13				90			0.38 (0.15–0.61)	Fair

Abbreviations: HCC – hepatocellular carcinoma; LR-4 – probably HCC; LR-5 – definitely HCC; LR-M – probably or definitely malignant but not HCC specific; R1 – reader 1; R2 – reader 2; TIV – tumor in vein.
[a] Data in parentheses represent 95% confidence intervals.

transplant livers to patients with false positive HCC imaging diagnoses.[6] However, the comparably lower sensitivity of LR-5 for the diagnosis of HCC in this population may relate to the variable appearance of HCC in the setting of hepatic steatosis[13] and to the tendency of HCC to present at a later stage or with a more infiltrative appearance in an unscreened population.[12]

Accurate performance of the LI-RADS diagnostic algorithm requires a sufficiently high pre-test probability that an observation represents HCC. Despite an unequivocally increased risk of HCC in non-cirrhotic HCV, non-cirrhotic NAFLD, and hepatic fibrosis, the pre-test probability in these populations has not yet been precisely established.[2,3,12,25] Specifically, the incidence of benign lesions that can mimic HCC is unknown. For instance, hepatocellular adenoma, which occurs at a higher rate in patients with NAFLD compared to the general population, has the potential to mimic HCC on imaging.[26–28] To a similar effect, LI-RADS excludes patients with cirrhosis secondary to vascular etiologies from the diagnostic population due to a high incidence of regenerative nodules which mimic HCC in these patients.[29] Due to our inclusion of only PLCs, the true specificity and positive predictive value of LR-5 for HCC in this population remains unclear. Our study supports the notion that LI-RADS can distinguish HCC from non-HCC with a high specificity in patients with pathologically-proven PLC and non-cirrhotic HCV, non-cirrhotic NAFLD, and hepatic fibrosis without cirrhosis. Additionally, MR elastography may be prove valuable in establishing a diagnosis of fibrosis in a non-invasive manner, potentially eliminating the need for a biopsy entirely in a patient with a

LR-5 observation and elevated liver stiffness.[30] However, a broad study of all lesions (benign and malignant, primary liver versus hepatic metastases) arising in this patient population is warranted before expanding the LI-RADS diagnostic population to include patients with these weaker HCC risk factors.

The sensitivity of LR-M for the prediction of non-HCC PLC was high (91% and 84% for R1 and R2, respectively). Combining LR-M and LR-TIV (may be due to non-HCC malignancy) improved the sensitivity to 95% and 98% for R1 and R2, respectively. Categorization of an observation as LR-M suggests that it is likely to represent malignancy but does not have features specific for HCC, and a biopsy may be necessary for definitive diagnosis.[6] Thus, high sensitivity is desired over a high specificity; accordingly, only a single LR-M feature is sufficient for categorization of an observation as LR-M, irrespective of major or ancillary features. The high degree of sensitivity of LR-M for non-HCC malignancy in our study may be explained, at least in part, by the large average size of observations in our study (approximately 6 cm). Smaller non-HCC PLCs, such as those arising in a screening population, may be less likely to develop a targetoid appearance, and thus may be more likely to mimic HCC.[31]

Agreement on overall LI-RADS v2018 category was fair (κ of 0.37), slightly less than the agreement reported in prior studies utilizing LI-RADS v2014.[16,32] One explanation for the lower agreement observed in this study is the modification of the tumor in vein algorithmic pathway; previously, all lesions with definite tumor in vein on imaging were categorized as LR-5V, whereas LI-RADS v2017 and v2018 entail assignment of a new LR-TIV category, with subsequent assessment of the associated parenchymal mass and assignment of one of three subcategories: probably due to HCC, definitely due to HCC, or may be due to non-HCC malignancy.[6,19] Further study is needed to investigate the accuracy and reliability of the tumor in vein subcategories introduced in v2017. Indeed, agreement on LR-5 versus other categories, as well as LR-5 versus LR-M or LR-TIV (i.e. likely malignant but not eligible for OPTN exception points) were moderate (κ of 0.53 and 0.55, respectively), suggesting that a number of the disagreements occurred between categories that have similar management strategies (i.e. LR-TIV [probably due to HCC] versus LR-TIV [definitely due to HCC]). Agreement for nearly all major criteria including nonrim APHE, nonperipheral "washout", enhancing "capsule", and size was substantial, whereas agreement for most LR-M criteria was fair to moderate. Surprisingly, agreement for peripheral "washout" was only slight, and more frequently scored as present by R1. However, the small discrepancy in diagnostic performance between R1 and R2, as well as the moderate agreement on LR-M versus other categories, suggests that this feature was rarely scored as a solitary LR-M feature, and given its coexistence with other LR-M criteria was unlikely to result in a change in overall LI-RADS category when scored as present.

Our study had several important limitations, most notably the inclusion of only pathologically-proven PLCs, as described above. Inclusion of benign masses remains a challenge in retrospective studies utilizing a pathologic reference standard, as masses that appear benign by imaging are rarely managed surgically and infrequently require biopsy to establish a diagnosis. Furthermore, our study did not include intrahepatic metastases, a more common occurrence than HCC in patients without cirrhosis. Prospective cohort studies focusing on the patient population in our study will likely be needed to overcome this limitation. Secondly, non-HCC PLCs were enriched in our population due to our methodologic choice to limit HCC to one third of the total number of malignant lesions. This choice was made to facilitate a more robust analysis of non-HCC PLC, and precluded calculation of positive and negative predictive values due to the over-representation of non-HCC malignancy relative to the frequencies encountered in practice. Thirdly, we included patients with steatosis without associated NASH or fibrosis. An established link exists between non-cirrhotic NALFD and HCC[33,34]; however, it is unclear whether NASH or fibrosis is a required mediator of HCC risk.[12] Several studies have demonstrated an association between metabolic syndrome and type II diabetes with HCC, raising the possibility that steatosis may, in fact, be sufficient to increase HCC risk.[3,34] Furthermore, spatial and temporal heterogeneity of inflammation may occur with hepatic steatosis.[18,35] Thus, sampling error associated with the location and timing of biopsy or resection may influence whether inflammation is identified pathologically in a patient with diffuse hepatic steatosis. Additionally, our population may have been biased by our requirement for a liver-protocol MRI or CT, a prerequisite for the application of the LI-RADS algorithm. Many lesions arising in unscreened patients are detected incidentally on routine contrast-enhanced CT, and subsequent liver-protocol MRI or CT is performed only if the patient is a candidate for surgical treatment or LRT (i.e. no extrahepatic metastatic disease). Our results, therefore, may only be applicable patients without advanced disease, rather than all patients with these HCC risk factors. Finally, our use of background liver tissue adjacent to the mass to establish risk factors for HCC has the potential for further sampling error, as peritumoral desmoplastic reaction or sinusoidal congestion can results in changes mimicking fibrosis.[36]

Conclusion

In a population at increased risk for HCC but not currently included in the LI-RADS 'high-risk' population, LI-RADS v2018 demonstrated fair inter-reader agreement for overall LI-RADS category and very high specificity for distinguishing HCC from non-HCC in patients with pathologically-proven PLC. Our study supports the notion of utilizing LI-RADS for distinguishing HCC from other primary malignant masses in this patient population. However, further study including non-malignant masses and

hepatic metastases is warranted before the LI-RADS diagnostic population can be expanded to include patients with non-cirrhotic HCV, non-cirrhotic NAFLD, and hepatic fibrosis.

Previous communication to a society or meeting

None.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Dr. Ludwig receives salary support from the Training OPportunities in Translational Imaging Education and Research (TOP-TIER) Holden Thorp grant at Washington University.

Conflicts of interest

None declared.

References

1. Ferenci P, Fried M, Labrecque D, Bruix J, Sherman M, Omata M *et al*. (2010) Hepatocellular carcinoma (HCC): a global perspective. *J Clin Gastroenterol* 44:239–245.
2. Bruix J, Sherman M. (2011) Management of hepatocellular carcinoma: an update. *Hepatology* 53:1020–1022.
3. Tang A, Hallouch O, Chernyak V, Kamaya A, Sirlin CB. (2018) Epidemiology of hepatocellular carcinoma: target population for surveillance and diagnosis. *Abdom Radiol (NY)* 43:13–25.
4. 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.
5. Marrero JA, Kulik LM, Sirlin C, Zhu AX, Finn RS, Abecassis MM *et al*. (2018) Diagnosis, staging and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. *Hepatology* 68:723–750.
6. American College of Radiology: CT/MRI LI-RADS v2018 Core. Available from: https:// www.acr.org/-/media/ACR/Files/RADS/LI-RADS/LI-RADS-2018-Core.pdf. [Accessed 25 July 2018].
7. Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR *et al*. (2018) Liver imaging reporting and data system (LI-RADS) version 2018: imaging of hepatocellular carcinoma in at-risk patients. *Radiology* 289:816–830.
8. 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.
9. Seeff LB, Everson GT, Morgan TR, Curto TM, Lee WM, Ghany MG *et al*. (2010) Complication rate of percutaneous liver biopsies among persons with advanced chronic liver disease in the HALT-C trial. *Clin Gastroenterol Hepatol* 8:877–883.
10. Sapisochin G, Fidelman N, Roberts JP, Yao FY. (2011) Mixed hepatocellular cholangiocarcinoma and intrahepatic cholangiocarcinoma in patients undergoing transplantation for hepatocellular carcinoma. *Liver Transplant* 17:934–942.
11. Vilchez V, Shah MB, Daily MF, Pena L, Tzeng CW, Davenport D *et al*. (2016) Long-term outcome of patients undergoing liver transplantation for mixed hepatocellular carcinoma and cholangiocarcinoma: an analysis of the UNOS database. *HPB* 18:29–34.
12. Piscaglia F, Svegliati-Baroni G, Barchetti A, Pecorelli A, Marinelli S, Tiribelli C *et al*. (2016) Clinical patterns of hepatocellular carcinoma in nonalcoholic fatty liver disease: a multicenter prospective study. *Hepatology* 63:827–838.
13. Thompson SM, Garg I, Ehman EC, Sheedy SP, Bookwalter CA, Carter RE *et al*. (2018) Non-alcoholic fatty liver disease-associated hepatocellular carcinoma: effect of hepatic steatosis on major hepatocellular carcinoma features at MRI. *Br J Radiol* 91:20180345.
14. Bosman FT, Carneiro F, Hruban RH, Theise ND. (2010) *WHO classification of tumours of the digestive system*. Lyon, France: IARC Press.
15. Brunt E, Aishima S, Clavien PA, Fowler K, Goodman Z, Gores G *et al*. (2018) cHCC-CCA: consensus terminology for primary liver carcinomas with both hepatocytic and cholangiocytic differentation. *Hepatology* 68:113–126.
16. Fraum TJ, Tsai R, Rohe E, Ludwig DR, Salter A, Nalbantoglu I *et al*. (2017) Differentiation of hepatocellular carcinoma from other hepatic malignancies in patients at risk: diagnostic performance of the liver imaging reporting and data system version 2014. *Radiology* 286: 158–172.
17. Furlan A, Almusa O, Yu RK, Sagreiya H, Borhani AA, Bae KT *et al*. (2018) A radiogenomic analysis of hepatocellular carcinoma: association between fractional allelic imbalance rate index and the liver imaging reporting and data system (LI-RADS) categories and features. *Br J Radiol* 91:1086.
18. Benedict M, Zhang X. (2017) Non-alcoholic fatty liver disease: an expanded review. *World J Hepatol* 9:715–732.
19. American College of Radiology: CT/MRI LI-RADS v2017 Core. Available from: https://www.acr.org/-/media/ACR/Files/RADS/LI-RADS/LIRADS_2017_Core.pdf. [Accessed 25 July 2018].
20. Mercaldo ND, Lau KF, Zhou XH. (2007) Confidence intervals for predictive values with an emphasis to case–control studies. *Stat Med* 26: 2170–2183.
21. Landis JR, Koch GG. (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174.
22. Benjamini Y, Yekutieli D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188.
23. Ludwig DR, Fraum TJ, Cannella R, Ballard DH, Tsai R, Naeem M *et al*. (2019) Hepatocellular carcinoma (HCC) versus non-HCC: accuracy and reliability of liver imaging reporting and data system v2018. *Abdom Radiol (NY)*. https://doi.org/10.1007/s00261-019-01948-x.
24. Ren AH, Zhao PF, Yang DW, Du JB, Wang ZC, Yang ZH. (2019) Diagnostic performance of MR for hepatocellular carcinoma based on LI-RADS v2018, compared with v2017. *J Magn Reson Imaging*. https://doi.org/10.1002/jmri.26640.
25. Mittal S, El-Serag HB, Sada YH, Kanwal F, Duan Z, Temple S *et al*. (2016) Hepatocellular carcinoma in the absence of cirrhosis in United States veterans is associated with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* 14:124–131. e1.
26. Bioulac-Sage P, Laumonier H, Couchy G, Le Bail B, Sa Cunha A, Rullier A *et al*. (2009) Hepatocellular adenoma management and phenotypic classification: the Bordeaux experience. *Hepatology* 50: 481–489.
27. Katabathina VS, Menias CO, Shanbhogue AK, Jagirdar J, Paspulati RM, Prasad SR. (2011) Genetics and imaging of hepatocellular adenomas: 2011 update. *Radiographics* 31:1529–1543.
28. Ba-Ssalamah A, Antunes C, Feier D, Bastati N, Hodge JC, Stift J *et al*. (2015) Morphologic and molecular features of hepatocellular adenoma with gadoxetic acid–enhanced MR imaging. *Radiology* 277:104–113.

29. Brancatelli G, Federle MP, Grazioli L, Golfieri R, Lencioni R. (2002) Benign regenerative nodules in Budd-Chiari syndrome and other vascular disorders of the liver: radiologic-pathologic and clinical correlation. *Radiographics* 22:847–862.
30. Venkatesh SK, Yin M, Ehman RL. (2013) Magnetic resonance elastography of liver: technique, analysis, and clinical applications. *J Magn Reson Imaging* 37:544–555.
31. Fowler KJ, Potretzke TA, Hope TA, Costa EA, Wilson SR. (2018) LI-RADS M (LR-M): definite or probable malignancy, not specific for hepatocellular carcinoma. *Abdom Radiol (NY)* 43:149–157.
32. Fowler KJ, Tang A, Santillan C, Bhargavan-Chatfield M, Heiken J, Jha RC *et al*. (2018) Interreader Reliability of LI-RADS version 2014 algorithm and imaging features for diagnosis of hepatocellular carcinoma: a large international multireader study. *Radiology* 286:173–185.
33. Stine JG, Wentworth BJ, Zimmet A, Rinella ME, Loomba R, Caldwell SH *et al*. (2018) Systematic review with meta-analysis: risk of hepatocellular carcinoma in non-alcoholic steatohepatitis without cirrhosis compared to other liver diseases. *Aliment Pharmacol Ther* 48:696–703.
34. Welzel TM, Graubard BI, Zeuzem S, El-Serag HB, Davila JA, McGlynn KA. (2011) Metabolic syndrome increases the risk of primary liver cancer in the United States: a study in the SEER-Medicare database. *Hepatology* 54:463–471.
35. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E *et al*. (2005) Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 128:1898–1906.
36. Kanematsu M, Matsuo M, Yamada Y, Semelka RC, Kondo H, Goshima S *et al*. (2003) Perilesional hyperintense rim of malignant hepatic tumors on ferumoxide-enhanced T1-weighted gradient-echo MR images: correlation between MR imaging and histopathologic findings. *J Magn Reson Imaging* 18:40–48.
37. Horowitz JM, Venkatesh SK, Ehman RL, Jhaveri K, Kamath P, Ohliger MA *et al*. (2017) Evaluation of hepatic fibrosis: a review from the society of abdominal radiology disease focus panel. *Abdom Radiol (NY)* 42:2037–2053.
38. Kudo M, Zheng RQ, Kim SR, Okabe Y, Osaki Y, Iijima H *et al*. (2008) Diagnostic accuracy of imaging for liver cirrhosis compared to histologically proven liver cirrhosis. *Intervirology* 51:17–26.
39. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J *et al*. (2006) Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43:1317–1325.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.hpb.2019.04.007.