Ann Surg Oncol (2019) 26:2560–2567
https://doi.org/10.1245/s10434-019-07386-9

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – HEPATOBILIARY TUMORS

Predictors and Cumulative Frequency of Hepatocellular Carcinoma in High and Intermediate LI-RADS Lesions: A Cohort Study from a Canadian Academic Institution

Ephraim Shin-Tian Tang, MD, MSc, FRCSC[1], Grayson Hall, MD, FRCPC[2], David Yu, MD[3], Alexandre Menard, MD, FRCPC[2], Wilma Hopman, MA[4], and Sulaiman Nanji, MD, PhD, FRCSC[3]

[1]Department of Surgery, University of Western Ontario, London, ON, Canada; [2]Department of Radiology, Queen's University, Kingston, ON, Canada; [3]Department of Surgery, Kingston General Hospital, Queen's University, Kingston, ON, Canada; [4]Kingston General Hospital Research Institute, Kingston, ON, Canada

ABSTRACT

Background. The frequency and predictors of hepatocellular carcinoma (HCC) within each liver imaging reporting and data system (LI-RADS) category remains unclear. We sought to estimate the cumulative frequency of HCC in LI-RADS observations of high/intermediate category and identify clinical/radiographic features associated with HCC.

Methods. Our diagnostic imaging database was searched for computed tomography/magnetic resonance imaging reports of patients with evidence of cirrhosis and liver observations. LI-RADS categories were determined by imaging review, while demographic and clinical outcomes were assigned by chart review. A composite outcome of clinical/radiographic confirmation of HCC was used. We used multivariable analysis to identify features associated with HCC, and competing risks regression to estimate the cumulative frequency of HCC in each category.

Results. Our search returned 95 patients with 137 observations (LR2 = 4, LR3 = 53, LR4 = 37, and LR5 = 43). On multivariable analysis, increasing age (hazard ratio [HR] 1.76 per 10 years, $p = 0.049$), washout (HR 5.34, $p < 0.002$), and increasing size (size < 10 mm reference, 10–20 mm, HR 3.93, $p = 0.014$; size > 20 mm, HR 21.69, $p < 0.001$) were associated with HCC. Median time to diagnosis was 6.13 months (interquartile range [IQR] 4.6–13.1), 4.7 months (IQR 2.5–14.5), and 3.6 months (IQR 1.9–6.6) for LR3, 4, and 5 category observations, respectively. The cumulative frequency of HCC was 59.8% in LR3, 84.62% in LR4, and 99.84% in LR5, at last follow-up.

Conclusion. The frequency of HCC within each LI-RADS category reflects the intended purpose, intermediate probability for LR3, probable HCC for LR4, and definite HCC for LR5.

While the non-invasive diagnosis of HCC has been well-validated,[1] many observations do not display the full spectrum of features. To address this issue, the American College of Radiologists (ACR) developed the Liver Imaging Reporting and Data System (LI-RADS) to standardize the reporting and interpretation of liver observations on cross-sectional imaging of patients at risk for HCC.[2]

Developed in 2011, the LI-RADS was most recently updated in 2018.[3] Categories range from 1 to 5 and reflect the probability of an underlying HCC.[2,3] LR1 and 2 lesions represent 'definite' and 'probably benign' observations, respectively, LR 3 lesions represent 'intermediate probability', LR 4 lesions represent 'probable HCC', and LI-RADS 5 lesions represent 'definite HCC'. However, the true incidence of HCC in each category remains poorly defined. Estimates of the frequency of HCC in LR 3 and 4 categories vary widely, ranging from 38%[4] to 96%.[5] To this end, the LI-RADS management group has identified an ongoing need for studies to confirm the proportion of HCCs

Electronic supplementary material The online version of this article (https://doi.org/10.1245/s10434-019-07386-9) contains supplementary material, which is available to authorized users.

First Received: 17 November 2018;
Published Online: 25 April 2019

S. Nanji, MD, PhD, FRCSC
e-mail: Sulaiman.Nanji@kingstonhsc.ca

within each category.[2] We sought to estimate the cumulative frequency of HCC in intermediate-and high-risk LI-RADS category observations and determine which radiographic and clinical features are most suggestive of underlying HCC.

METHODS

Study Design and Population

We performed a retrospective cohort study to determine the cumulative risk of HCC within intermediate and high LI-RADS categories. We also sought to identify predictors of underlying HCC within these categories. This study was approved by and conducted in accordance with our institutional Ethics Review Board.

Our institutional database was queried using Montage (now mPower, Nuance, Burlington MA, USA[6]) for reports of computed tomography/magnetic resonance imaging (CT/MRI) studies with both cirrhosis and liver observations, between 2009 and 2015. The following Boolean terms were used: (cirrhosis AND lesion) OR (cirrhosis AND mass) OR (cirrhosis AND nodule) OR (cirrhosis AND observation). Patients were only included if they had two or more imaging studies to permit longitudinal follow-up and increase the yield of screened studies. Studies also had to be appropriately protocoled as per the LI-RADS criteria.

Chart review and image screening was performed by three authors (ET, GH, DY). Clinical variables abstracted included primary etiology of cirrhosis, vital status, oncologic outcome, and pathologic information from resected/biopsied lesions. Radiographic features (size, arterial enhancement, washout, pseudocapsule, fat content, and location) of each observation at the index imaging study were also determined. Many studies pre-dated LI-RADS categories, therefore all LI-RADS categories were assigned de novo, as per LI-RADS 2011. Observations were followed individually up to a maximum of four observations per patient. Those with more than four observations were excluded for feasibility. All initial radiographic features, LI-RADS classifications, and final radiographic outcomes were reviewed and adjudicated by a fellowship trained staff radiologist with 8 years of academic practice experience (AM).

Imaging Protocols

CT liver protocols were performed on a GE LightSpeed VCT 64 (GE Healthcare, Chicago, IL, USA) and an Aquilion One 320 CT scanner (Canon Medical Systems, previously Toshiba, Otawara, Japan) with acquisition of an unenhanced liver followed by post-contrast imaging at 10, 65, and 180 s. MRI liver protocols were performed on an Avento 1.5T MRI scanner (Siemens Healthcare, Malvern, PA, USA) with acquisition of the following sequences: coronal T2, axial T1 in and out of phase, axial T2, axial DWI (b0, b400, and b800), axial T2 with fat saturation, axial T1 with fat saturation pre-and post-gadolinium with Gadovist (Bayer, Leverkusen, Germany) at the immediate arterial phase and 60, 180, and 300 s.

Measures and Outcomes

A composite outcome was used as a surrogate for the presence or progression to HCC within an observation. Outcomes were classified as *unknown*, *definitely HCC*, or *definitely not HCC*. *Unknown* was defined as the absence of radiographic progression on follow-up imaging, or until the last point of clinical follow-up or death. Lesions that were treated by ablation or transarterial chemoembolization (TACE) without biopsy were also considered *unknown*. Lesions were considered *definitely HCC* when there was radiographic progression (increase in LI-RADS category, recurrence after local therapy, or, in the case of LR5 observations, clear progression, such as increase in size, or emergence of multifocal HCC), confirmation of the diagnosis on biopsy, surgical resection, or explant at the time of transplant. Lesions that were not visualized on follow-up imaging, or that were negative on pathology, were considered *definitely not HCC*.

Statistical Analysis

All analysis was performed on either SPSS version 24.0 for Windows (IBM Corporation, Armonk, NY, USA), or STATA version 14.0 for Windows (StataCorp LLC, College Station, TX, USA). Analysis was conducted at the observation level. Univariable and multivariable analysis was performed to measure the association of variables with *definitely HCC* versus all other outcomes. Univariable analysis was performed using the Chi square test for categorical data, and *t* test for continuous data; alpha was set at 0.05. Medians were compared using the Wilcoxon rank-sum test, and multivariable analysis was performed using logistic regression. Variables significant on univariable analysis were run forward and backward until the final model was reached. Age divided by ten was used to estimate effect size by decade. The median time to diagnosis was calculated for each LR category based on observations with an outcome of *definite HCC*.

Given our desire to produce an unbiased estimate of the cumulative frequency of HCC within each category, we avoided using a fixed period of radiographic stability to define observations as non-HCC. Instead, we used

competing risks regression, a form of survival analysis that calculates the cumulative index function rather than the survival function. This method is able to account for competing risks that prevent the primary failure event from occurring.[7] In our study, *definite HCC* was the primary failure event, while observations that were *definitely not HCC* were the competing risk; *unknown* observations are censored. This approach permits an estimate of the cumulative frequency of HCC (*definite HCC*) in each LI-RADS category at maximum follow-up. We did not treat death as a competing risk since patients who died may still have had an underlying HCC at the time of death.

As a sensitivity analysis, we recalculated the cumulative index function by coding lesions treated with ablation of TACE as definite HCC, based on the high index of suspicion in these lesions. A second sensitivity analysis was performed where observations that were stable for varying minimum durations (6, 9, or 12 months) were treated as *definitely not HCC* to determine the impact of different cut-offs of radiographic stability on our estimates.

RESULTS

Study Population and Radiographic Characteristics

Our search returned 637 patients. Most (n = 542) were excluded for no properly protocoled studies (n = 170), no cirrhosis (n = 27), no observations (n = 263), or an established diagnosis of HCC or other malignancy at the time of initial imaging (n = 82). Therefore, 95 patients comprised our study cohort, with a range of one to four observations, for a total of 137 observations (electronic supplementary Fig. 1). Demographic and radiographic features are summarized in Table 1. Of the 137 evaluated lesions, 14 were not noted on follow-up imaging, and were classified as *definitely not HCC* (electronic supplementary Fig. 1). Clinical and radiographic outcomes are summarized in Table 2.

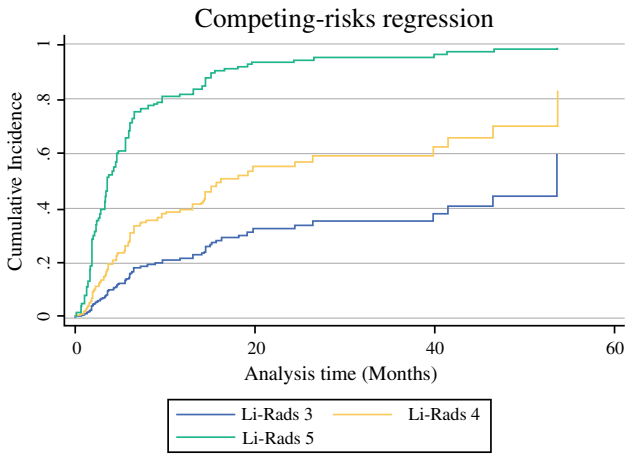

FIG. 1 Cumulative index function of *definite HCC* by LI-RADS category from competing risks regression. Final estimates for the cumulative frequency of HCC within each category: LR3: 59.78%; LR4: 84.62%; LR5: 99.84%. *HCC* hepatocellular carcinoma, *LI-RADS* liver imaging reporting and data system

TABLE 1 Clinical characteristics and initial radiographic characteristics of patients with suspected HCC, and observations identified in patients with suspected HCC

Patient characteristics [n = 95]	n (%)
Demographics	
Mean age [years (SD)]	61.6 (± 9.2)
Male sex	74 (77.9)
Liver disease	
Hepatitis C	39 (41.1)
Alcohol	23 (24.2)
NASH	21 (22.1)
Hepatitis B	3 (3.2)
Cholestasis	2 (2.1)
Autoimmune	1 (1.1)
Other	5 (5.3)
Number of lesions	
1	66 (69.5)
2	18 (18.9)
3	9 (9.5)
4	2 (2.1)
Median duration of follow-up, months (quartiles)	6.5 (2.7–17.9)
Median number of imaging studies (quartiles)	3 (2–5)
Radiographic features per observation (n = 137)[a]	**n (%)**
LI-RADS category	
2	4 (2.9)
3	53 (38.7)
4	37 (27.0)
5	43 (31.4)
Imaging characteristics	
Mean size [mm (SD)]	19.2 (15)
Arterial hyperenhancement	110 (80.3)
Washout	64 (46.7)
Fat content	5 (3.6)
Pseudocapsule	14 (10.2)
Threshold growth	23 (16.8)

Data are expressed as n (%) unless otherwise specified

NASH non-alcoholic steatohepatitis, *HCC* hepatocellular carcinoma, *SD* standard deviation, *LI-RADS* liver imaging reporting and data system

[a]Multiple characteristics per lesion included; denominator for the proportion is the total number of lesions (n = 137)

TABLE 2 Summary of radiographic and clinical outcomes by LI-RADS category

LI-RADS category	Definitely not HCC	Definitely HCC	Unknown
Radiographic outcomes			
2 (4)	1	0	3
3 (53)	9	12	32
4 (37)	4	11	22
5 (43)	0	19	24
Clinical outcomes			
2 (4)	0	0	4
3 (53)	5	12	36
4 (37)	1	19	17
5 (43)	0	34	9
Composite outcome			
2 (4)	1	0	3
3 (53)	10	17	26
4 (37)	5	21	11
5 (43)	0	39	4

Data are presented at the individual observation level (n = 137)

LI-RADS liver imaging reporting and data system, *HCC* hepatocellular carcinoma

Risk Factors Associated with Hepatocellular Carcinoma (HCC)

Univariable analysis was performed to evaluate the association of clinical and radiographic features with *definite HCC* (Table 3). Notably, male sex and underlying cause for cirrhosis were associated with HCC. Radiographic features associated with HCC included hyperenhancement, washout, pseudocapsule, increasing LI-RADS category, and increasing size. On multivariable regression, *definite HCC* was associated with increasing age, alcoholic cirrhosis, washout, and increasing size category (Table 4). Pseudocapsule could not be included in the multivariable model given its presence in the HCC group alone.

Cumulative Frequency of HCC and Time to Diagnosis Within Each LI-RADS Category

Using competing risks regression, the cumulative frequency of HCC within each LI-RADS category was 59.8% for LR3, 84.6% for LR4, and 99.8% for LR5 (Fig. 1) at a maximum follow-up (53 months). As part of a sensitivity analysis, cumulative frequencies were calculated under two additional conditions: (1) locally treated lesions without pathologic diagnosis or evidence of recurrence were also *definite HCC* (LR3: 60.0%; LR4: 87.0%; LR5: 99.9%); and (2) observations with stability for 6 (LR3: 39.2%; LR4: 71.1%; LR5: 97.6%), 9 (LR3: 42.7%; LR4: 71.1%; LR5: 98.4%), or 12 (LR3: 42.1%; LR4: 70.4%; LR5: 98.3%) months were treated as *definitely not HCC*.

We also calculated the median time to diagnosis in *definite HCC*. This was 6.13 months for LR3 observations (interquartile range [IQR] 4.6–13.1 months), 4.7 months for LR4 observations (IQR 2.5–14.5 months), and 3.6 months for LR5 observations (IQR 1.9–6.6 months). Comparison of the median time to progression between these categories was significant for the comparison between LR3 and LR5 observations (p = 0.04), and non-significant for others (LR3 vs. LR4, p = 0.428; LR4 vs. LR5, p = 0.084).

DISCUSSION

The primary objective of our study was to determine the cumulative frequency of HCC within each LI-RADS category. We found that, essentially, all LR5 lesions represent HCC (99.8%), and that while the majority of LR4 (84.6%) and 3 (59.8%) lesions are estimated to develop/harbor HCC, this proportion decreases in each category.

In designating the LR5 category, the ACR intended to provide high specificity, such that non-HCC lesions would only very rarely be included in this category. Other authors have reported estimates within this category, ranging from 91 to 98% within the LR5 category.[5,8,9] Findings from a recent series of 595 observations, of which 341 were LR5, HCC frequency was estimated at 94–96%.[8] Differences in estimates may be related to the way non-HCC observations were defined, since radiographic stability is frequently used as a criterion for non-HCC and an unknown proportion of these may be false negatives.

TABLE 3 Univariable analysis of initial imaging features associated with *definite HCC*

Continuous variables	Unknown/negative [n = 46]	HCC [n = 79]	p value
Size, mm [mean (SD)]	11.04 (5.7)	24.5 (16.9)	< 0.001*
Age, years [mean (SD)]	61.09 (9.3)	61.48 (9.4)	0.821
Categorical variables [n]	**Unknown/negative (%)**	**HCC (%)**	**p value**
Sex			
Male [92]	26 (28.3)	66 (71.7)	0.001*
Female [33]	30 (60.6)	13 (39.4)	
Etiology			
Hepatitis C [53]	16 (30.2)	37 (69.8)	0.004*
NAFLD [31]	17 (54.8)	14 (45.2)	
ETOH [25]	4 (16.0)	21 (84.0)	
Other [12]	6 (50.0)	6 (50.0)	
Multiple lesions			
No [57]	19 (33.3)	38 (66.7)	0.640
Yes [67]	27 (40.3)	40 (59.7)	
Hyperenhancement			
Absent [25]	14 (56.0)	11 (44.0)	0.026*
Present [100]	32 (32.0)	68 (68.0)	
Washout			
Absent [67]	36 (53.7)	31 (46.3)	0.001*
Present [58]	10 (17.2)	48 (82.8)	
Fat content			
Absent [120]	46 (38.3)	74 (61.7)	0.175
Present [3]	0 (0.0)	3 (100)	
Psuedocapsule			
Absent [114]	46 (40.4)	68 (59.6)	0.008*
Present [11]	0 (0.0)	11 (100)	
Growth			
Absent [78]	36 (46.2)	42 (53.8)	0.017*
Present [23]	4 (17.4)	19 (82.6)	
LI-RADS category			
2 [4]	4 (100)	0 (0)	< 0.001*
3 [48]	31 (64.6)	17 (35.4)	
4 [34]	11 (32.4)	23 (67.6)	
5 [39]	0 (0.0)	39 (100)	

Unless otherwise noted, statistical comparisons were performed using the t test for continuous variables and the Chi square test for categorical variables

HCC hepatocellular carcinoma, *NAFLD* non-alcoholic fatty liver disease, *ETOH* ethyl alcohol, *LI-RADS* liver imaging reporting and data system

*$p < 0.05$

Similarly, the goal of the LR4 category, i.e. *probable HCC*, is supported by a cumulative estimate of 84.6%. Prior estimates have ranged as widely as 29–96% within this category.[4,5,10,11] Of these estimates, the study that was at the highest end of this spectrum was restricted to observations that were identified on screening ultrasound, a feature that has been demonstrated to predict HCC.[12] These other estimates have used radiographic outcomes alone or variable follow-up, potentially underestimating the frequency of HCC. Others have attempted to further define the outcomes within the LR4 category, and noted T2 hyperintensity, threshold growth, and hepatitis C made subsequent upgrade more likely.[10,11] These findings support the concept that the majority of LR4 observations represent HCC; with a median time to diagnosis of 4.7 months, aggressive pursuit of a diagnosis in these lesions, or even treatment, may be warranted.

TABLE 4 Initial and selective multivariable logistic regression model of features associated with HCC

Variable	Initial model			Selective model		
	OR (95% CI)		*p* value	OR (95% CI)		*p* value
Age [divided by 10]	1.74	(0.97–3.12)	0.065	1.76	(1.00–3.07)	0.049*
Sex						
Female	0.77	(0.23–2.62)	0.674			
Etiology						
Viral hepatitis	1.00	Ref	Ref	1.00	Ref	0.035*
Alcoholism	2.25	(0.52–9.68)	0.275	2.33	(0.56–9.73)	0.245
NASH	0.38	(0.10–1.38)	0.375	0.31	(0.10–1.01)	0.053
Other	0.18	(0.01–2.27)	0.184	0.13	(0.01–1.49)	0.102
Imaging features						
Arterial enhancement	1.55	(0.46–5.18)	0.481			
Washout	5.00	(1.67–14.97)	< 0.002*	5.34	(1.90–15.03)	< 0.002*
Growth	1.65	(0.56–4.91)	0.367			
Size (mm)						
< 10	1.00	Ref	Ref	1.00	Ref	< 0.001*
10–19	3.65	(1.18–11.25)	0.024*	3.93	(1.32–11.73)	0.014*
20 +	18.13	(4.39–74.82)	< 0.001*	21.69	(5.36–87.74)	< 0.001*

*$p < 0.05$

Features significant in the univariable analysis were added forwards and backwards in a stepwise fashion until the final model was generated. For variables where reference is not specified, absence of the feature has been used as reference

HCC hepatocellular carcinoma, *OR* odds ratio, *CI* confidence interval, *NASH* non-alcoholic steatohepatitis, *Ref* reference

Even within the *intermediate probability* LR3 category, we noted that most observations (59.8%) are expected to be, or develop into, HCC. Estimates within this category vary the most[12] and prior studies have estimated anywhere from 9 to 69% HCC.[4,5] A recent large series with 249 LR3 observations estimated this frequency at 33–41%,[8] using a 6-month radiographic stability cut-off for determining non-HCC lesions. We performed two sensitivity analysis to test the effects of our assumptions on the generated frequency estimates. In the first, observations treated by radiofrequency ablation or TACE without clear radiographic progression or pathologic diagnosis were considered *definite HCC*. With this assumption, frequency estimates were only slightly increased. In the second sensitivity analysis, we did note that an a priori cut-off of radiographic stability markedly decreased the cumulative frequency estimates, especially in LR category 3 and 4 observations. Different stability thresholds of 6, 9, or 12 months had only marginal effects on these estimates. However, we noted the median time to diagnosis of HCC within LR3 lesions was 6.13 months, suggesting even stable lesions might ultimately represent HCC.

The median time to diagnosis also varied depending on LR category. Median time to diagnosis was higher in lower LR category lesions, and, for LR4 and LR3 observations, the upper quartile of time to diagnosis/progression was 13.1 and 14.5 months, respectively. These findings imply that even radiographic stability over the period of several months to 1 year does not guarantee benignity. Whether this reflects the presence of a low-grade neoplasm with slow growth, or the progression of a pre-neoplastic lesion to frank malignancy remains unclear. While these findings could also reflect differences in imaging interval or diagnostic strategy, previous work has demonstrated a relationship between the median doubling times of HCC and grade or intralesional blood supply.[13,14]

Regarding clinical and radiographic features associated with HCC, size > 2.0 cm was associated with the largest hazard ratio of all the major features in our analysis. In many other studies, this has been validated previously as a predictor of HCC.[1,14,15] Of interest, while washout was also associated with HCC in the multivariable analysis, arterial hyperenhancement was not (Table 3). This is in keeping with reports of high sensitivity, but poor specificity of arterial enhancement in diagnosing HCC, compared with washout appearance, which is highly specific but has poor sensitivity.[14,16] While we were unable to calculate a hazard ratio for pseudocapsule, others have noted high specificity.[14,16]

Strengths

In this study, we evaluated the incidence of HCC by LI-RADS category in a real-world population of imaging studies. By avoiding selection through screening US,[5] or confining analysis to only observations with pathology available,[17] our search strategy is less prone to selection bias and is applicable to patients with observations identified de novo on cross-sectional imaging. This increases the external validity and generalizability of our findings. Additionally, while prior estimates have used radiographic or pathologic outcomes alone,[4,18] we used a composite endpoint and competing risks regression to adjust for non-HCC observations and censor unknown outcomes. We believe this approach minimizes acquisition bias and assumptions of outcome based on missing data.

Weaknesses

Certain limitations are noteworthy. Our study is retrospective with a modest sample size. In addition, for many lesions we did not have a definite outcome and used censoring to account for this. While these approaches are appropriate, they provide only an estimate of the actuarial frequency. Negative biopsies are subject to sampling error, and radiographic regression or disappearance on a subsequent examination, while suggestive of artefact, do not guarantee that observations are *definitely not HCC*. It is also worth mentioning that our search strategy used the term 'cirrhosis' to identify relevant studies. While we likely missed eligible studies, we feel a broader strategy would not have been feasible given the large number of studies screened.

LI-RADS is also intended for use in select non-cirrhotic patients with chronic hepatitis B infection, or in patients with prior treated HCC now undergoing surveillance. Our results are not applicable to this population. While prior work has included all patients undergoing imaging investigation based on clinical concern for HCC, exact risk factors were not explicitly stated and only biopsied patients were included, potentially engendering selection bias.[16] Further studies to evaluate the performance of the LI-RADS categories and major radiographic features in these patient populations are required.

CONCLUSION

We performed a retrospective analysis to determine clinical and radiographic features associated with HCC in patients with cirrhosis, and to estimate the frequency of HCC within intermediate LI-RADS observations. While our findings support the intended purpose of each category, they also suggest most intermediate observations are likely HCC. Additionally, the time to diagnosis may vary substantially in intermediate observations, and even prolonged radiographic stability does not guarantee benignity.

DISCLOSURES Ephraim Shin-Tian Tang, Grayson Hall, David Yu, Alexandre Menard, Wilma Hopman, and Sulaiman Nanji have no conflicts of interest to declare.

FUNDING No specific funding sources were used for this study.

REFERENCES

1. Forner A, Vilana R, Ayuso C, 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.
2. Mitchell DG, Bruix J, Sherman M, Sirlin CB. LI-RADS (liver imaging reporting and data system): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. *Hepatology*. 2015;61(3):1056–65.
3. Bashir M, Chernyak V, Cui J, et al. CT/MRI LI-RADS V2017. 2017; CT and MRI LI-RADS algorithm. https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/CT-MRI-LI-RADS-v2017. Accessed 4 May 2018.
4. Tanabe M, Kanki A, Wolfson T, et al. Imaging outcomes of liver imaging reporting and data system version 2014 category 2, 3, and 4 observations detected at CT and MR imaging. *Radiology*. 2016;281:152173.
5. Darnell A, Forner A, Rimola J, et al. Liver Imaging reporting and data system with MR imaging: evaluation in nodules 20 mm or smaller detected in cirrhosis at screening US. *Radiology*. 2015;275(3):698–707.
6. mPower Clinical Analytics for Diagnostic Imaging. 2018; Product webpage for Nuance mPower. https://www.nuance.com/healthcare/medical-imaging/mpower-clinical-analytics.html. Accessed 7 May 2018.
7. Fine JP, Gray RJ. A Proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496–509.
8. Ronot M, Fouque O, Esvan M, Lebigot J, Aube C, Vilgrain V. Comparison of the accuracy of AASLD and LI-RADS criteria for the non-invasive diagnosis of HCC smaller than 3 cm. *J Hepatol*. 2018;68(4):715–23.
9. Choi SH, Byun JH, Kim SY, et al. Liver imaging reporting and data system v2014 with gadoxetate disodium-enhanced magnetic resonance imaging: validation of LI-RADS category 4 and 5 criteria. *Invest Radiol*. 2016;51(8):483–90.
10. Burke LM, Sofue K, Alagiyawanna M, et al. Natural history of liver imaging reporting and data system category 4 nodules in MRI. *Abdom Radiol (NY)*. 2016;41(9):1758–66.
11. Sofue K, Burke LMB, Nilmini V, et al. Liver imaging reporting and data system category 4 observations in MRI: risk factors predicting upgrade to category 5. *J Magn Reson Imaging*. 2017;46(3):783–92.
12. Mitchell DG, Bashir MR, Sirlin CB. Management implications and outcomes of LI-RADS-2, -3, -4, and -M category observations. *Abdom Radiol (NY)*. 2018;43(1):143–8.
13. Kudo M, Tochio H. Intranodular blood supply correlates well with biological malignancy grade determined by tumor growth rate in pathologically proven hepatocellular carcinoma. *Oncology*. 2008;75 Suppl 1:55–64.
14. Tang A, Bashir MR, Corwin MT, et al. Evidence supporting LI-RADS major features for CT- and MR imaging-based diagnosis

of hepatocellular carcinoma: a systematic review. *Radiology*. 2018;286(1):29–48.
15. Ohkawa K, Imanaka K, Sakakibara M, et al. Factors related to shift from hepatic borderline lesion to overt HCC diagnosed by CT. *Hepatogastroenterology*. 2014;61(134):1680–7.
16. Cerny M, Bergeron C, Billiard JS, et al. LI-RADS for MR imaging diagnosis of hepatocellular carcinoma: performance of major and ancillary features. *Radiology*. 2018;288:171678.
17. Zhao W, Li W, Yi X, et al. Diagnostic value of liver imaging reporting and data system MRI on primary hepatocellular carcinoma [in Chinese]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban*. 2016;41(4):380–7.
18. Choi SH, Byun JH, Lim YS, et al. Liver imaging reporting and data system: patient outcomes for category 4 and 5 nodules. *Radiology*. 2018;287(2):515–24.