



Ancillary features in the Liver Imaging Reporting and Data System: how to improve diagnosis of hepatocellular carcinoma ≤ 3 cm on magnetic resonance imaging

Ji Hun Kang¹ · Sang Hyun Choi¹ · Jae Ho Byun¹ · Dong Hwan Kim^{1,2} · So Jung Lee¹ · So Yeon Kim¹ · Hyung Jin Won¹ · Yong Moon Shin¹ · Pyo-Nyun Kim¹

Received: 6 November 2019 / Revised: 6 November 2019 / Accepted: 18 December 2019 / Published online: 4 February 2020
© European Society of Radiology 2020

Abstract

Objectives To determine the strength of association with hepatocellular carcinoma (HCC) of each ancillary feature (AF) in LI-RADS version 2018, and to develop an appropriate strategy for applying AFs to improve the diagnosis of HCC ≤ 3 cm on gadoxetate-enhanced MRI.

Methods A total of 385 nodules (283 HCCs, 18 non-HCC malignancies, 84 benign nodules) of ≤ 3 cm in 266 patients at risk for HCC who underwent gadoxetate-enhanced MRI in 2016 were retrospectively evaluated. Two radiologists independently evaluated the presence/absence of AFs, and assigned a LI-RADS category to each nodule. Diagnostic odds ratio (DOR) of each AF was assessed. To improve the diagnostic performance for HCC, various criteria were developed based on the number of AFs favoring malignancy in general or HCC in particular. Generalized estimating equation models were used to compare the diagnostic performance of each criterion with that of the major features (MFs) only.

Result All AFs favoring HCC in particular and malignancy in general were more common in the HCC group than in the non-HCC group. Of these AFs, hepatobiliary-phase hypointensity had the strongest association with HCC (DOR, 21.82; 95% confidence interval, 5.59–85.20). When we applied AFs in addition to MFs, the new criterion (with a number of AFs ≥ 4) had significantly higher sensitivity (80.6% vs. 70.0%; $p < 0.001$) than MFs only, without significant lower specificity (85.3% vs. 90.2%; $p = 0.060$).

Conclusions The AFs varied in the strengths of association with HCC. More strict application of AFs (AFs ≥ 4) in LR-3 may improve the diagnostic performance for probable HCC ≤ 3 cm.

Key Points

- The ancillary features (AFs) in the Liver Imaging Reporting and Data System version 2018 showed variable frequencies of occurrence and strengths of association with hepatocellular carcinoma (HCC).
- Of the various AFs, hepatobiliary-phase hypointensity had the highest frequency and strongest association with HCC on gadoxetate disodium-enhanced MRI.
- When applying AFs in addition to major features, a criterion of four or more AFs significantly increased the sensitivity for diagnosing HCC, without a significantly decreased specificity, especially in LR-3 observations.

Ji Hun Kang and Sang Hyun Choi contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-019-06645-3>) contains supplementary material, which is available to authorized users.

✉ Jae Ho Byun
jhbyun@amc.seoul.kr

¹ Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-Ro 43-Gil, Songpa-Gu, Seoul 05505, Republic of Korea

² Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

Keywords Diagnosis · Hepatocellular carcinoma · Magnetic resonance imaging

Abbreviations

AASLD	American Association for the Study of Liver Diseases
AF	Ancillary feature
DOR	Diagnostic odds ratio
HCC	Hepatocellular carcinoma
LI-RADS	Liver Imaging Reporting and Data System
MF	Major feature
MRI	Magnetic resonance imaging
RFA	Radiofrequency ablation
TACE	Transcatheter arterial chemoembolization

Introduction

To standardize the performance of liver imaging in patients at risk for hepatocellular carcinoma (HCC) as well as interpretation and reporting of the results, the Liver Imaging Reporting and Data System (LI-RADS) was introduced in 2011 [1], recently updated in 2018 [2], and fully integrated into the American Association for the Study of Liver Diseases (AASLD) 2018 HCC clinical practice guidance [3]. LI-RADS assigns categories to liver observations based on the presence of major and ancillary imaging features [2]. These categories reflect the relative likelihood of benignity and HCC (i.e., LR-1 to LR-5), as well as the differentiation from other malignancies (LR-M). The major features (MFs) include size, non-rim arterial-phase hyperenhancement, non-peripheral washout, enhancing capsule, and threshold growth, which are used in the category assignment of LR-3, LR-4, and LR-5 observations, and the diagnosis of HCC currently relies on MFs alone [4]. In contrast to the MFs, ancillary features (AFs) may optionally be used to improve detection, increase confidence, or modify the category after the application of the MFs; i.e., AFs can upgrade LR category by one up to LR-4, but cannot be used to upgrade LR category to LR-5 [5].

Several rules for AF application are provided in LI-RADS v2018 [2]. However, the decision on whether to apply AFs when assigning a category is left to the radiologist's discretion [2]. Although this flexibility may encourage wide adoption of LI-RADS by minimizing its complexity [6], it may also result in a variable range of change in the LI-RADS category assignments, i.e., 20.5–97.1% of LR-3 being upgraded to LR-4 [5, 7–9]. Recently, a few studies have reported that the use of AFs in combination with the MFs could increase the diagnostic performance for HCC, but they did not elucidate how to apply the AFs when assigning a category [5].

Therefore, the purpose of this study was to determine both the frequency of occurrence of each AF in the LI-RADS v2018 and its strength of association with HCCs, and to

develop an appropriate strategy for applying AFs to improve the diagnosis of HCC ≤ 3 cm on gadoxetate disodium-enhanced magnetic resonance imaging (MRI) through the use of a historical cohort study.

Materials and methods

Our institutional review board approved this retrospective study, and the requirement for informed consent was waived.

Study subjects

From January to December 2016, 3854 patients at risk for HCC underwent surveillance ultrasound (Fig. 1). According to the AASLD practice guidance [3], patients who had hepatic nodules ≥ 1 cm in diameter detected on ultrasound were referred for further evaluations including dynamic computed tomography or MRI. A total of 746 patients underwent gadoxetate disodium-enhanced MRI for the evaluation of suspicious nodules detected during ultrasound surveillance. Patients were included according to the following criteria: (a) focal hepatic solid nodules on MRI, (b) nodule size ≤ 3.0 cm, (c) number of nodules ≤ 5 , and (d) the nodule not being definitely or probably benign (LR-1 or LR-2) on MRI, such as a cyst, hemangioma, perfusion alternation, hepatic fat deposition or sparing, hypertrophic pseudomass, confluent fibrosis, or focal scar [7]. Of these 569 available nodules in 396 patients, 184 nodules were excluded because of a lack of final diagnosis due to an insufficient follow-up period of less than 24 months (38 nodules in 18 patients), or immediate locoregional treatment such as transcatheter arterial chemoembolization (TACE) or radiofrequency ablation (RFA) without biopsy or marginal tumor recurrence (146 nodules in 112 patients). Finally, 385 nodules in 266 patients were analyzed in this study.

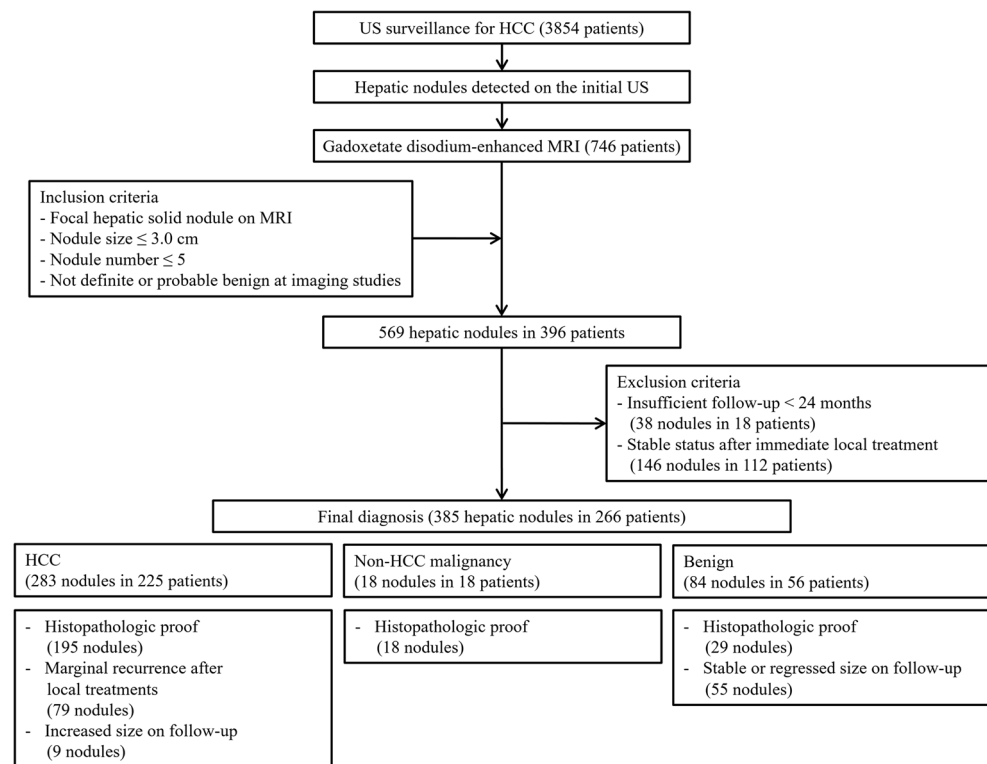
MRI techniques

MRI was performed on a 1.5-T or 3-T scanners. Detailed techniques are given in [Supplemental Method](#) and [Supplemental Table 1](#).

Image analysis

Two board-certified abdominal radiologists (each with > 7 years of experience in hepatic imaging) who were blinded to any information on clinical history or final diagnosis independently reviewed the MRI, and in the case of any discrepancies between the two readers, re-evaluated the MRI together to reach a consensus.

Fig. 1 Flowchart of study population. US, ultrasound; HCC, hepatocellular carcinoma; MRI, magnetic resonance imaging



The readers analyzed nodule size and location, and the presence or absence of MFs (non-rim arterial-phase hyperenhancement, non-peripheral washout, or enhancing capsule), targetoid mass features, and AFs, according to the LI-RADS v2018 [2]. AFs included findings favoring malignancy in general, favoring HCC in particular, and favoring benignity (Supplementary Table 2). Subthreshold growth, size stability over 2 or more years, and size reduction were not included, because hepatic nodules ≥ 1 cm in diameter initially detected on ultrasound were included, and no previous examination was available for comparison.

Reference standard

The final diagnoses for the 385 nodules in 266 patients were as follows: (a) 283 nodules in 225 patients were determined as HCC on the basis of pathological evidence (184 nodules by resection or explantation, and 11 by biopsy), marginal recurrence after RFA or TACE (79 nodules), or interval growth of the lesion on follow-up images (nine nodules; mean, 4.2 months [range, 1–9]); (b) 18 nodules in 18 patients were determined as non-HCC malignancies by pathological diagnosis (17 nodules by resection or explantation, and one by biopsy), including 11 combined HCC and cholangiocarcinomas, six cholangiocarcinomas, and one metastatic adenocarcinoma; and (c) 84 nodules in 56 patients were determined to be benign by pathological specimens (29 nodules, including 13 nodules by surgical resection, three nodules

by biopsy, and 13 nodules by explantation for liver transplantation) or a stable or regressed nodule size for at least 24 months (55 nodules; mean, 29.6 months [range, 24–37]). The 29 benign nodules determined by histopathologic proof were eight high-grade dysplastic nodules, 16 low-grade dysplastic nodules, one biliary adenoma, one hemangioma, one angiomyolipoma, one hyalinized nodule, and one chronic granulomatous inflammation. Of the 266 patients, 29 had both HCC (35 nodules) and benign lesions (39 nodules), three had both HCC (three nodules) and non-HCC malignancy (three nodules), and one had both non-HCC malignancy (one nodule) and benign lesions (two nodules).

Statistical analysis

All analyses were performed on a per-lesion basis. The frequency of occurrence of each AF was recorded for both the HCC and non-HCC groups. To determine the strength of association between HCC diagnosis and imaging features, the diagnostic odds ratios (DORs) of each MF and AF were calculated using logistic regression model with generalized estimating equation to adjust clustering effect [10]. Inter-observer agreement on the presence of AFs was assessed using the overall proportion of agreement and kappa statistics.

To develop the best strategy for applying AFs to improve diagnostic performance for HCC ≤ 3 cm, we counted the total number of all observed AFs favoring malignancy in general or HCC in particular. When both AFs favoring malignancy in

general or HCC in particular and AFs favoring benignity were simultaneously present, we did not adjust category determined using MFs only according to the LI-RADS version 2018 [2]. When only AFs favoring malignancy in general or HCC in particular were present, various criteria were developed using cutoff points based on the number of all observed AFs favoring malignancy in general or HCC in particular regardless of AFs' DORs and we adjusted category determined using MFs only according to each criterion.

The diagnostic performance for HCC ≤ 3 cm, i.e., sensitivity, specificity, positive predictive value, negative predictive value, and accuracy, was calculated when only MFs were considered. After then, the diagnostic performance of each criterion was calculated when AFs in addition to MFs were considered. To evaluate the diagnostic performance of each criterion, the sensitivity and specificity were compared with those values attained before addition of the AFs using generalized estimating equation models.

Subgroup analysis was performed in each LI-RADS category group (i.e., LR-3, LR-4, and LR-5) to determine the LI-RADS categories in which AFs can improve diagnostic accuracy. The change in diagnostic performance after applying AFs was also evaluated using McNemar's test and Bonferroni's correction to compare the sensitivity and specificity before and after application of the AFs.

Statistical analyses were performed using SAS v 9.4 (SAS Institute). $P < 0.05$ was considered to indicate a significant difference. In the case of multiple comparisons, Bonferroni's correction was applied to give an adjusted significance value of $p < 0.01$.

Results

Patients

The baseline characteristics of the 266 patients and 385 nodules are summarized in Table 1. There were 216 men (mean age, 61.2 years; range, 37–86 years) and 50 women (mean age, 61.7 years; range, 33–84 years). Hepatitis B was the most common cause of chronic liver disease ($n = 210$, 78.9%), followed by alcoholic liver disease ($n = 24$, 9.0%). The sizes of the 385 nodules ranged from 5 to 30 mm (mean, 18.3 ± 7.7 mm). The median value of total bilirubin in the 266 patients was 0.7 mg/dL (range, 0.2–6.0).

Major and ancillary imaging features

The LI-RADS category assignments using MFs only were as follows: LR-3, 154 nodules (40.0%); LR-4, 44 nodules (11.4%); LR-5, 164 nodules (42.6%); LR-TIV, 4 nodules (1.0%); and LR-M, 19 nodules (4.9%). Of the 283 HCCs, 159 (56.2%) were categorized as LR-5, while 82.4% (84/

Table 1 Baseline characteristics of the 266 patients and 385 nodules

Characteristic	Data
Age (years)*	61.3 \pm 9.1
Sex	
Male	216 (81.2)
Female	50 (18.8)
No. of nodules	
1	181 (68.0)
2	60 (22.6)
3	19 (7.1)
4	3 (1.1)
5	3 (1.1)
Risk factor	
Hepatitis B	210 (78.9)
Hepatitis C	13 (4.9)
Hepatitis B and C	2 (0.8)
Alcoholic liver disease	24 (9.0)
Others [†]	17 (6.4)
Child-Pugh classification [‡]	
A	241 (91.0)
B	20 (7.5)
C	4 (1.5)
BCLC stage	
0	55 (20.7)
A	181 (68.0)
B	15 (5.6)
C	11 (4.1)
D	4 (1.5)
Total bilirubin (mg/dL) [§]	0.7 (0.2–6.0)
Tumor marker (ng/mL)	
AFP [§]	7.1 (0.79–111,953)

Unless otherwise indicated, data are the number of patients with percentages in parentheses

AFP alpha-fetoprotein, BCLC Barcelona Clinic Liver Cancer

*Data are mean \pm standard deviation

[†] Others included liver cirrhosis caused by non-alcoholic fatty liver disease ($n = 3$) and cryptogenic liver cirrhosis ($n = 14$)

[‡] One patient had no available information

[§] Data are median value, and data in parentheses are range

102) of the non-HCCs were categorized as LR-3. The three MFs of non-rim arterial-phase hyperenhancement, non-peripheral washout, and enhancing capsule were more common in the HCC group than in the non-HCC group (Table 2). Of these three MFs, non-rim arterial-phase hyperenhancement had the strongest association with HCC (DOR 21.51; 95% confidence interval [CI] 10.18–45.47), followed by enhancing capsule (DOR 10.04; 95% CI 1.22–82.31) (Table 2).

Differences in the frequencies and DORs of AFs between the two groups are summarized in Table 3. Of the AFs favoring HCC in particular, three features (non-enhancing capsule,

Table 2 Frequencies and diagnostic odds ratios of the major features

Major feature	HCC (<i>n</i> = 283)	Non-HCC (<i>n</i> = 102)	DOR*
Size (mm)			
< 10	24 (8.5)	27 (26.5)	1
10–19	105 (37.1)	58 (56.9)	1.02 (0.53–1.97)
≥ 20	154 (54.4)	17 (16.7)	2.77 (1.18–6.50)
Non-rim arterial-phase hyperenhancement	228 (80.6)	12 (11.8)	21.51 (10.18–45.47)
Non-peripheral washout	192 (67.8)	35 (34.3)	2.76 (1.50–5.07)
Enhancing capsule	90 (31.8)	1 (1.0)	10.04 (1.22–82.31)

Data are number of nodules, and data in parentheses are percentages

DOR diagnostic odds ratio, HCC hepatocellular carcinoma

*Data in parentheses are 95% confidence intervals

nodule-in-nodule architecture, and mosaic architecture) were noted only in the HCC group. Mosaic architecture had the strongest association with HCC (DOR 15.95; 95% CI 0.89–285.37) among the five AFs favoring HCC in particular. All AFs favoring malignancy in general, not HCC in particular, were noted more frequently in the HCC group than in the non-HCC group. Of the AFs favoring malignancy in general, not HCC in particular, hepatobiliary-phase hypointensity had the strongest association with HCC (DOR 21.82; 95% CI 5.59–85.20), followed by restricted diffusion (DOR 16.45; 95% CI 8.85–30.57). The frequency of all AFs favoring benignity in

the HCC group was 1% or less than 1%, which was lower than that in the non-HCC group.

The proportions of inter-observer agreement and kappa values for each AF ranged from 85.5% (329/385) to 99.2% (382/385) and from 0.0 to 0.75, respectively (Supplementary Table 3). Of the 17 AFs, iron sparing in solid mass showed the highest proportion of agreement (99.2%), and corona enhancement and transitional-phase hypointensity showed the lowest proportion of agreement (85.5%). In addition, mild-moderate T2 hyperintensity showed the highest kappa value ($\kappa = 0.75$), and iron sparing in solid mass showed the lowest

Table 3 Frequencies and diagnostic odds ratios of the ancillary features

Ancillary feature	HCC (<i>n</i> = 283)	Non-HCC (<i>n</i> = 102)	DOR*
Favoring malignancy in general, not HCC in particular			
Corona enhancement	19 (6.7)	4 (3.9)	2.12 (0.60–7.49)
Restricted diffusion	232 (82.0)	21 (20.6)	16.45 (8.85–30.57)
Mild-moderate T2 hyperintensity	239 (84.5)	27 (26.5)	15.34 (8.59–27.41)
Iron sparing in solid mass	2 (0.7)	0 (0.0)	1.82 (0.04–75.31)
Fat sparing in solid mass	2 (0.7)	0 (0.0)	1.82 (0.04–75.31)
Transitional-phase hypointensity	262 (92.6)	70 (68.6)	5.96 (3.27–10.86)
Hepatobiliary-phase hypointensity	278 (98.2)	86 (84.3)	21.82 (5.59–85.20)
Favoring HCC in particular			
Non-enhancing capsule	17 (6.0)	0 (0.0)	13.49 (0.74–246.07)
Nodule-in-nodule architecture	13 (4.6)	0 (0.0)	10.24 (0.54–193.61)
Mosaic architecture	20 (7.1)	0 (0.0)	15.95 (0.89–285.37)
Fat in mass, more than adjacent liver	57 (20.1)	8 (7.8)	2.30 (0.98–5.43)
Blood products in mass	9 (3.2)	1 (1.0)	14.49 (0.01–95,779.12)
Favoring benignity			
Parallels blood pool enhancement	0 (0.0)	1 (1.0)	0.12 (0.01–11.11)
Undistorted vessels	3 (1.1)	4 (3.9)	0.61 (0.16–2.25)
Iron in mass, more than liver	3 (1.1)	2 (2.0)	0.57 (0.03–12.77)
Marked T2 hyperintensity	1 (0.4)	1 (1.0)	0.79 (0.35–1.76)
Hepatobiliary-phase isointensity	2 (0.7)	6 (5.9)	0.03 (0.01–0.59)

Data are number of nodules, and data in parentheses are percentages

DOR diagnostic odds ratio, HCC hepatocellular carcinoma

*Data in parentheses are 95% confidence intervals

kappa value ($\kappa = 0.0$). The reasons of discrepancy between kappa values and proportions of agreement are a well-known kappa paradox and very low prevalences of some ancillary features (Table 3) [11].

Diagnostic performance of criteria using ancillary features

The diagnostic performances of the LI-RADS categorizations for HCC ≤ 3 cm before and after adding the AFs to the MFs are summarized in Table 4. When we used MFs only, the sensitivity and specificity of the LR-5 category for diagnosis of HCC ≤ 3 cm were 56.2% (159/283) and 95.1% (97/102), respectively, while those of the LR-4 and LR-5 categories combined were 70.0% (198/283) and 90.2% (92/102), respectively.

We developed five criteria using the number of AFs favoring malignancy in general or HCC in particular regardless those DORs as cutoff points, i.e., a number of AFs ≥ 1 to ≥ 5 . As the number of AFs increased, the sensitivity decreased but the specificity increased. The sensitivities of the criteria with AFs ≥ 1 to AFs ≥ 3 were significantly higher than those of MFs only (86.6–93.6% [245/283–265/283] vs. 70.0% [198/283]; $p < 0.001$), while the specificities were significantly lower than those of MFs only (22.5–72.5% [23/102–74/102] vs. 90.2% [92/102]; $p < 0.001$). The sensitivity and specificity of the criterion of AFs ≥ 5 were not significantly different from those of MFs only ($p \geq 0.120$). However, the criterion of AFs ≥ 4 had significantly higher sensitivity than that of MFs only (80.6% [228/283] vs. 70.0% [198/283]; $p < 0.001$), while the specificity was not significantly different (85.3% [87/102] vs. 90.2% [92/102]; $p = 0.060$; Fig. 2). In this criterion (AFs ≥ 4), the combination of hepatobiliary-phase hypointensity, transitional-phase hypointensity, mild-moderate T2 hyperintensity, and restricted diffusion was the most common one (84.4%, 205/243).

Subgroup analyses for validating the effects of ancillary features

In the LR-3 category subgroup (Supplementary Table 4), significant increases in diagnostic accuracy were noted with the criteria of AFs ≥ 3 and AFs ≥ 4 ($p \leq 0.001$). The criterion of AFs ≥ 4 had significantly higher sensitivity than that of MFs only (42.9% [30/70] vs. 0.0% [0/70]; $p < 0.001$), although the specificity was not significantly different (94.0% [79/84] vs. 100.0% [84/84]; $p = 0.025$). The most common combination of AFs used in LR-3 being upgraded to LR-4 was hepatobiliary-phase hypointensity, transitional-phase hypointensity, mild-moderate T2 hyperintensity, and restricted diffusion (80.0%, 28/35).

However, in both the LR-4 (Supplementary Table 5) and LR-5 (Supplementary Table 6) category subgroups, the

Table 4 Diagnostic performance of LI-RADS for HCC in 385 nodules before and after the additional application of ancillary features to major features

Diagnostic criteria	No. of nodules				Diagnostic performance							
	TP	FN	FP	TN	Sensitivity (%)	p value*	Specificity (%)	p value†	Accuracy (%)	p value‡	PPV (%)	NPV (%)
Major features only (LR-5 as a positive finding)	159	124	5	97	56.2 (50.1–62.1)	–	95.1 (88.6–98.0)	–	66.5 (61.1–71.5)	–	97.0 (92.9–98.7)	43.9 (36.2–51.9)
Major features only (LR-4 or LR-5 as a positive finding)	198	85	10	92	70.0 (63.9–75.4)	Ref	90.2 (82.8–94.6)	Ref	75.3 (70.4–79.7)	Ref	95.2 (91.3–97.4)	52.0 (43.3–60.6)
Major and ancillary features combined (LR-4 or LR-5 as a positive finding)	265	18	79	23	93.6 (90.2–95.5)	< 0.001	22.5 (14.9–32.6)	< 0.001	74.8 (69.4–79.5)	> 0.999	77.0 (71.1–82.1)	56.1 (40.3–70.8)
≥ 1 ancillary features	260	23	67	35	91.9 (88.3–94.4)	< 0.001	34.3 (24.9–45.2)	< 0.001	76.6 (71.4–81.1)	0.995	79.5 (73.5–84.4)	60.3 (47.0–72.3)
≥ 2 ancillary features	245	38	28	74	86.6 (81.6–90.4)	< 0.001	72.5 (63.1–80.3)	< 0.001	82.9 (78.6–86.4)	0.001	89.7 (85.6–92.8)	66.1 (54.9–75.7)
≥ 3 ancillary features	228	55	15	87	80.6 (75.1–85.1)	< 0.001	85.3 (76.4–91.2)	0.060	81.8 (77.3–85.6)	< 0.001	93.8 (89.7–96.3)	61.3 (51.4–70.3)
≥ 4 ancillary features	204	79	10	92	72.1 (66.1–77.4)	0.120	90.2 (82.8–94.6)	> 0.999	76.9 (72.0–81.1)	0.123	95.3 (91.5–97.5)	53.8 (44.9–62.5)

Data in parentheses are 95% confidence intervals

FN false negative, FP false positive, HCC hepatocellular carcinoma, LI-RADS Liver Imaging Reporting and Data System, NPV negative predictive value, PPV positive predictive value, Ref reference, TN true negative, TP true positive

*Comparison of sensitivity in LR-3, LR-4, and LR-5 categories between major features only (reference) and major and ancillary features combined using generalized estimating equation (GEE) models with a significance level of $p < 0.05$

† Comparison of specificity in LR-3, LR-4, and LR-5 categories between major features only (reference) and major and ancillary features combined using GEE models with a significance level of $p < 0.05$

‡ Comparison of accuracy in LR-3, LR-4, and LR-5 categories between major features only (reference) and major and ancillary features combined using GEE models with a significance level of $p < 0.05$

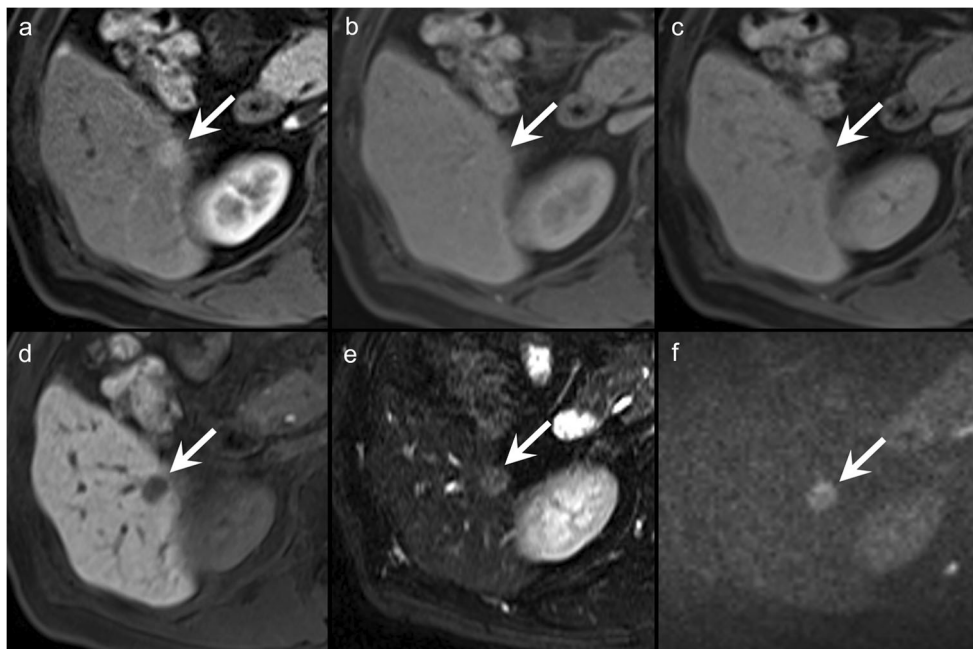


Fig. 2 A 60-year-old man with chronic hepatitis B and a surgically confirmed hepatocellular carcinoma. **a** A 15-mm nodule in hepatic segment VI shows non-rim arterial-phase hyperenhancement (arrow) without non-peripheral washout (**b**) on portal venous phase or enhancing capsule on portal venous phase or (**c**) transitional phase. The nodule was categorized as LR-3 when only major features were considered. The nodule (arrow)

shows (**c**) transitional-phase hypointensity, (**d**) hepatobiliary-phase hypointensity, (**e**) mild T2 hyperintensity, and (**f**) restricted diffusion ($b = 900 \text{ s/mm}^2$). After application of these four ancillary features in addition to the major features, this nodule was upgraded to LR-4. This nodule was confirmed as hepatocellular carcinoma after surgical resection

sensitivities decreased when AFs were used in addition to MFs, while the specificities increased, in comparison with values determined using MFs only. The diagnostic accuracy did not increase when AFs were applied in addition to MFs.

Discussion

This study demonstrated that AFs favoring HCC in particular and malignancy in general were more common in the HCC group than in the non-HCC group, with variable frequencies. Of these AFs, hepatobiliary-phase hypointensity had the strongest association with HCC $\leq 3 \text{ cm}$. When we applied the AFs in addition to the MFs, the criterion of four or more AFs favoring malignancy in general or HCC in particular significantly increased the sensitivity for diagnosis of probable HCC $\leq 3 \text{ cm}$ from 70.0 to 80.6% ($p < 0.001$), without a significant decrease in specificity (90.2% vs. 85.3%; $p = 0.060$), especially in LR-3 observations.

In LI-RADS v2018, there are nine AFs favoring malignancy in general, five AFs favoring HCC in particular, and seven AFs favoring benignity [2]. As these AFs are equally weighted in significance, the use of AFs might be unclear (e.g., whether an LR-3 nodule should be upgraded to LR-4 when only one rare AF favoring malignancy is present). A few recent studies noted the different frequencies of AFs and their associations with HCC and revealed that corona enhancement

and restricted diffusion showed high DORs on gadobenate dimeglumine-enhanced MRI [5, 12]. However, in this study, using gadoxetate disodium-enhanced MRI, hepatobiliary-phase hypointensity had the highest frequency and strongest association with HCC, followed by restricted diffusion and mild-moderate T2 hyperintensity. Considering the different frequencies of AFs and their different associations with HCC, we need to stratify the various AFs according to their strength of association, and in doing so, the use of AFs may be made clearer.

Previous studies have reported on the importance of hepatobiliary-phase hypointensity for diagnosing HCC, including increasing sensitivity and discriminating well-differentiated HCC from benign precursor nodules [13–15]. However, in this study, hepatobiliary-phase hypointensity showed high frequencies in both the non-HCC group (84.3%) and the HCC group (98.2%). Because of the high frequency of hepatobiliary-phase hypointensity in the non-HCC group, the criterion of AFs ≥ 1 decreased the specificity for diagnosing HCC from 90.2 to 22.5%. Therefore, in the case of gadoxetate disodium-enhanced MRI, a number of AFs ≥ 1 may be an insufficient criterion for a one-category upgrade up to LR-4. In other words, AFs need to be used in a more conservative and strict approach.

Several recent studies reported that AFs could improve the diagnostic performance for HCC, showing 10.0–12.9% increases in sensitivity [5, 16, 17]. However, although many

previous studies have reported on the LI-RADS category assignment and its diagnostic performance using AFs, the changes in the LI-RADS category assignments before and after the use of AFs were unclear [12, 18–20] or variable, i.e., 20.5–97.1% of LR-3 upgraded to LR-4 [5, 7, 9]. In the LI-RADS v2018, the LI-RADS category can be upgraded by one category up to LR-4 when one or more AFs favoring malignancy are present, although the use of AFs is at the radiologist's discretion [2]. This ambiguity might result in the variable range of change in LI-RADS category assignments and the conflicting results about whether AFs can increase the sensitivity of the LR-4 category for diagnosing HCC without a significant decrease in specificity [5, 17]. In this study, we found that diagnostic performance varied according to the number of AFs favoring malignancy in general or HCC in particular, and that the criterion of AFs ≥ 4 was the best criterion, as it increased sensitivity without significantly decreasing specificity. Therefore, the criterion of AFs ≥ 4 may help clarify the use of AFs when applying the LI-RADS for the diagnosis of HCC ≤ 3 cm.

According to the subgroup analyses, in both LR-5 and LR-4 nodules, the added application of AFs showed no benefit for HCC diagnosis in comparison with the use of MFs only. This result supports the current guideline, which does not allow LR-4 to be upgraded to LR-5, even though hepatic observations may have AFs favoring malignancy [2]. In contrast to both LR-4 and LR-5 subgroup analyses, we found a significant increase in diagnostic accuracy when the criteria of AFs ≥ 3 or ≥ 4 were applied to the LR-3 subgroup. Furthermore, the criterion of AFs ≥ 4 in the LR-3 subgroup significantly increased sensitivity (42.9% vs. 0.0%; $p < 0.001$), without significantly decreasing specificity (94.0% vs. 100.0%; $p = 0.025$). This suggests that the added application of AFs is important in LR-3 nodules, which have the potential to be upgraded to LR-4.

In this study, the LI-RADS category assignments using MFs only showed a bimodal distribution of LR categories, with LR-3 (40%) and LR-5 (43%) making up the most with a small proportion of LR-4 (11%). Our result is similar to those of previous studies regarding the LI-RADS category assignment determined by MFs only, i.e., 30–35% in LR-3, 14–20% in LR-4, and 21–44% in LR-5 [5, 7]. Considering these results, the use of AFs might undermine a bimodal distribution of LI-RADS categories, increasing confidence in category assignment.

This study has a few limitations. First, the retrospective nature of the study design may have led to selection bias. However, we tried to overcome this limitation by including a large number of subjects and using a historical cohort study design similar to a real clinical situation. Second, the clinical diagnosis including marginal tumor recurrence after TACE or RFA, or interval growth of the lesion on follow-up imaging, could be a limitation due to a lack of universal histological

confirmation as a reference standard. However, according to the AASLD practice guidance, observations that meet the typical imaging findings can be diagnosed as HCC without histopathological evidence [3]. Therefore, our final diagnoses correspond to clinical practice. Third, the predominance of patients with hepatitis B and preserved liver function (Child-Pugh classification A) in our study subjects may have a limitation to apply our results generally.

In conclusion, various AFs in the LI-RADS v2018 showed variable frequencies of occurrence and strengths of association with HCC ≤ 3 cm. More strict application of AFs in addition to MFs in LR-3 observations is needed to improve the diagnostic performance for probable HCC ≤ 3 cm on gadoxetate disodium-enhanced MRI, i.e., a criterion of four or more AFs favoring malignancy in general or HCC in particular.

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

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Jae Ho Byun.

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 No complex statistical methods were necessary for this paper

Informed consent Written informed consent was waived by the Institutional Review Board because of the retrospective nature of the study.

Ethical approval Institutional Review Board approval was obtained.

Study subjects or cohorts overlap There was subject overlap with an original article in *J Hepatol* (doi: <https://doi.org/10.1016/j.jhep.2019.05.005>) entitled “Arterial Subtraction Images of Gadoxetate-Enhanced MRI Improve Diagnosis of Early-Stage Hepatocellular Carcinoma”.

Methodology

- Retrospective
- Diagnostic or prognostic study
- Performed at one institution

References

1. 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
2. American College of Radiology CT/MRI LI-RADS v2018. Available via <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/CT-MRI-LI-RADS-v2018>. Accessed May 27, 2019

3. Marrero JA, Kulik LM, Sirlin CB 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
4. Tang A, Bashir MR, Corwin MT et al (2018) Evidence supporting LI-RADS major features for CT- and MR imaging-based diagnosis of hepatocellular carcinoma: a systematic review. *Radiology* 286:29–48
5. Cerny M, Bergeron C, Billiard JS et al (2018) LI-RADS for MR imaging diagnosis of hepatocellular carcinoma: performance of major and ancillary features. *Radiology* 288:118–128
6. Chernyak V, Tang A, Flusberg M et al (2018) LI-RADS® ancillary features on CT and MRI. *Abdom Radiol (NY)* 43:82–100
7. Choi SH, Byun JH, Kim SY et al (2016) 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* 51:483–490
8. Choi SH, Byun JH, Lim YS et al (2018) Liver Imaging Reporting and Data System: patient outcomes for category 4 and 5 nodules. *Radiology* 287:515–524
9. Joo I, Lee JM, Lee DH, Ahn SJ, Lee ES, Han JK (2017) Liver Imaging Reporting and Data System v2014 categorization of hepatocellular carcinoma on gadoxetic acid-enhanced MRI: comparison with multiphasic multidetector computed tomography. *J Magn Reson Imaging* 45:731–740
10. Bland JM (2004) Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol* 4:21
11. Feinstein AR, Cicchetti DV (1990) High agreement but low kappa: I. the problems of two paradoxes. *J Clin Epidemiol* 43:543–549
12. Fraum TJ, Tsai R, Rohe E et al (2018) 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
13. Cortis K, Liotta R, Miraglia R, Caruso S, Tuzzolino F, Luca A (2016) Incorporating the hepatobiliary phase of gadobenate dimeglumine-enhanced MRI in the diagnosis of hepatocellular carcinoma: increasing the sensitivity without compromising specificity. *Acta Radiol* 57:923–931
14. Lee MH, Kim SH, Park MJ, Park CK, Rhim H (2011) Gadoxetic acid-enhanced hepatobiliary phase MRI and high-b-value diffusion-weighted imaging to distinguish well-differentiated hepatocellular carcinomas from benign nodules in patients with chronic liver disease. *AJR Am J Roentgenol* 197:W868–W875
15. Orlacchio A, Chegai F, Fabiano S et al (2016) Role of MRI with hepatospecific contrast agent in the identification and characterization of focal liver lesions: pathological correlation in explanted livers. *Radiol Med* 121:588–596
16. Min JH, Kim YK, Sinn DH et al (2018) Adding ancillary features to enhancement patterns of hepatocellular carcinoma on gadoxetic acid-enhanced magnetic resonance imaging improves diagnostic performance. *Abdom Radiol (NY)* 43:2309–2320
17. Ronot M, Fouque O, Esvan M, Lebigot J, Aube C, Vilgrain V (2018) Comparison of the accuracy of AASLD and LI-RADS criteria for the non-invasive diagnosis of HCC smaller than 3cm. *J Hepatol* 68:715–723
18. Kim YY, An C, Kim S, Kim MJ (2018) Diagnostic accuracy of prospective application of the Liver Imaging Reporting and Data System (LI-RADS) in gadoxetate-enhanced MRI. *Eur Radiol* 28:2038–2046
19. Min JH, Kim JM, Kim YK et al (2018) Prospective intraindividual comparison of magnetic resonance imaging with gadoxetic acid and extracellular contrast for diagnosis of hepatocellular carcinomas using the Liver Imaging Reporting and Data System. *Hepatology* 68:2254–2266
20. Joo I, Lee JM, Lee DH, Jeon JH, Han JK (2019) Retrospective validation of a new diagnostic criterion for hepatocellular carcinoma on gadoxetic acid-enhanced MRI: can hypointensity on the hepatobiliary phase be used as an alternative to washout with the aid of ancillary features? *Eur Radiol* 29:1724–1732

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.