Original Article | Gastrointestinal Imaging

eISSN 2005-8330
https://doi.org/10.3348/kjr.2020.0846
Korean J Radiol 2021;22(7):1066-1076

Korean Journal of Radiology
KJR

Diagnostic Performance of 2018 KLCA-NCC Practice Guideline for Hepatocellular Carcinoma on Gadoxetic Acid-Enhanced MRI in Patients with Chronic Hepatitis B or Cirrhosis: Comparison with LI-RADS Version 2018

Sang Min Lee[1], Jeong Min Lee[2, 3, 4], Su Joa Ahn[2, 5], Hyo-Jin Kang[2], Hyun Kyung Yang[2, 6], Jeong Hee Yoon[2, 3]

[1]Department of Radiology, Hallym University Sacred Heart Hospital, Anyang, Korea; [2]Department of Radiology, Seoul National University Hospital, Seoul, Korea; [3]Department of Radiology, Seoul National University College of Medicine, Seoul, Korea; [4]Institute of Radiation Medicine, Seoul National University College of Medicine, Seoul, Korea; [5]Department of Radiology, Gil Medical Center of Gachon University, Incheon, Korea; [6]Department of Radiology, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

Objective: To evaluate the performance of the 2018 Korean Liver Cancer Association-National Cancer Center (KLCA-NCC) Practice Guidelines (hereafter, PG) for the diagnosis of hepatocellular carcinoma (HCC) using gadoxetic acid-enhanced MRI, compared to the Liver Imaging-Reporting and Data System (LI-RADS) version 2018 (hereafter, v2018).

Materials and Methods: From January 2013 to October 2015, treatment-naïve hepatic lesions (≥ 1 cm) on gadoxetic acid-enhanced MRI in consecutive patients with chronic hepatitis B or cirrhosis were retrospectively evaluated. For each lesion, three radiologists independently analyzed the imaging features and classified the lesions into categories according to the 2018 KLCA-NCC PG and LI-RADS v2018. The imaging features and categories were determined by consensus. Generalized estimating equation (GEE) models were used to compare the per-lesion diagnostic performance of the 2018 KLCA-NCC PG and LI-RADS v2018 using the consensus data.

Results: In total, 422 lesions (234 HCCs, 45 non-HCC malignancies, and 143 benign lesions) from 387 patients (79% male; mean age, 59 years) were included. In all lesions, the definite HCC (2018 KLCA-NCC PG) had a higher sensitivity and lower specificity than LR-5 (LI-RADS v2018) (87.2% [204/234] vs. 80.8% [189/234], $p < 0.001$; 86.2% [162/188] vs. 91.0% [171/188], $p = 0.002$). However, in lesions of size ≥ 2 cm, the definite HCC had a higher sensitivity than the LR-5 (86.8% [164/189] vs. 82.0 (155/189), $p = 0.002$) without a reduction in the specificity (80.0% [48/60] vs. 83.3% [50/60], $p = 0.15$). In all lesions, the sensitivity and specificity of the definite/probable HCC (2018 KLCA-NCC PG) and LR-5/4 did not differ significantly (89.7% [210/234] vs. 91.5% [214/234], $p = 0.204$; 83.5% [157/188] vs. 79.3% [149/188], $p = 0.071$).

Conclusion: For the diagnosis of HCC of size ≥ 2 cm, the definite HCC (2018 KLCA-NCC PG) had a higher sensitivity than LR-5, without a reduction in specificity. The definite/probable HCC (2018 KLCA-NCC PG) had a similar sensitivity and specificity to that those of the LR-5/4.

Keywords: *Diagnosis; Liver neoplasm; Hepatocellular carcinoma; Magnetic resonance imaging; Gadoxetic acid*

INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer and the second leading cause of cancer-related mortality globally [1]. HCC in high-risk patients can be pathologically or non-invasively diagnosed. According to the guidelines for the diagnosis and management of HCC released by several major scientific organizations [2], lesions

Received: June 30, 2020 **Revised:** October 3, 2020 **Accepted:** November 3, 2020
Corresponding author: Jeong Min Lee, MD, Department of Radiology, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea.
• E-mail: jmsh@snu.ac.kr

Korean Journal of Radiology
KJR

(≥ 1 cm) with arterial phase hyperenhancement (APHE) and washout on contrast-enhanced multiphasic computed tomography (CT) or magnetic resonance imaging (MRI) can be non-invasively diagnosed as HCC in high-risk patients. However, there are still some differences between the diagnostic criteria of APHE and washout in the guidelines proposed by Eastern and Western societies due to different management approaches [1]. In general, Eastern countries (particularly Korea) frequently use more aggressive surgical approaches and locoregional treatments for HCC because of the high prevalence of viral hepatitis-related liver cirrhosis and limited donors; thus, Asian diagnostic guidelines for HCC tend to focus on the sensitivity for the diagnosis of HCC [1]. In contrast, a high level of specificity in diagnosing HCC is more critical in Western countries, where transplantation is considered one of the major first-line treatments for early-stage HCC.

There are several guidelines for the diagnosis of HCC, including the Liver Imaging-Reporting and Data System (LI-RADS), the European Association of Study of the Liver, and the Korean Liver Cancer Association-National Cancer Center (KLCA-NCC) practice guidelines (hereafter, PG), updated and released in 2018 [2-5]. According to the LI-RADS version 2018 (hereafter, v2018) and the 2018 KLCA-NCC PG, a lesion that has the major features, including APHE and washout, is regarded as definite HCC. However, there are significant differences between the criteria for washout on gadoxetic acid-enhanced MRI [2,5]. The revised KLCA-NCC PG, which was designed to increase the sensitivity for the diagnosis of HCC over the previous version, applied washout not only during the portal venous phase (PVP), but also the transitional and hepatobiliary phases when using gadoxetic acid-enhanced MRI after applying the exclusion criteria [5]. Further, in diagnosing probable HCC, which is frequently treated in actual Korean practice, there are differences between the applications of ancillary features in LI-RADS v2018 and the 2018 KLCA-NCC PG; 2018 KLCA-NCC PG also adopted the definitions of the ancillary features of LI-RADS v2018.

Therefore, the purpose of this study was to evaluate the diagnostic performance of the definite HCC criteria and definite HCC or probable HCC (definite/probable HCC) criteria of the 2018 KLCA-NCC PG for the diagnosis of HCC on gadoxetic acid-enhanced MRI in patients with chronic hepatitis B or cirrhosis in comparison with the LR-5 (definitely HCC) and LR-5 or LR-4 (probably HCC) criteria of the LI-RADS v2018.

MATERIALS AND METHODS

Patients

This single-center retrospective cohort study was approved by the Institutional Review Board of Seoul National University Hospital, which waived the requirement for written informed consent (IRB No. H-1711-123-901). Between January 2013 and October 2015, 6156 gadoxetic acid-enhanced MRI examinations were obtained from our radiology database. Two radiologists reviewed the electronic medical records and MRI to identify patients who met the following criteria: 1) patients aged ≥ 18 years with clinical evidence of cirrhosis or chronic hepatitis B; 2) patients with at least one hepatic lesion (≥ 1 cm) other than hepatic cysts; 3) no previous treatment for hepatic lesions; 4) lesions with a conclusive diagnosis (histologic diagnosis or clinical diagnosis [typical imaging feature and size stability ≥ 2 years]). The following patients were excluded: 1) patients who received previous treatment for hepatic lesions; 2) patients with congestive hepatopathies (e.g., Budd-Chiari syndrome, cardiac congestion); 3) patients with insufficient data, such as inconclusive histopathological diagnosis or no follow-up imaging for confirming benignity; 4) patients who underwent locoregional treatments such as radiofrequency ablation for HCCs without pathologic confirmation; 5) patients with suboptimal MRI images (e.g., image degradation); 6) patients with duplicated data. The lesions (≤ 3 per patient) were selected and annotated by radiologists. Finally, 387 patients (305 male and 82 female; mean age ± standard deviation, 59 ± 10 years) with 422 lesions (234 HCC; 45 non-HCC malignancy; 143 benign lesions) were included in this study (Table 1). All HCC and non-HCC malignancies were pathologically confirmed. Of the 143 benign lesions, the diagnosis was pathologically confirmed for 18 (12.6%) (Supplementary Table 1) and clinically confirmed based on typical imaging features and no change in size for ≥ 2 years in 125 (87.4%).

The same study population has been included in a previous study, in which we reported the performance of the LI-RADS v2017 and LI-RADS v2018 [6]. However, as opposed to the previous study, which assessed the performance of LR-5 in diagnosing HCC without considering category adjustment (downgrade into LR-4) while applying ancillary features, this study aimed to evaluate and compare the diagnostic performances of the 2018 KLCA-NCC PG and LI-RADS v2018 using not only definite HCC criteria, but also the combination of definite and probable HCC criteria and

 https://doi.org/10.3348/kjr.2020.0846

Table 1. Clinical and Pathologic Information

Characteristic	Result
Patient (n = 387)	
Age, year*	59 ± 10
Male:female	305:82
Underlying liver disease	
Liver cirrhosis	287 (74.2)
Chronic hepatitis B	100 (25.8)
Etiology of liver disease	
Hepatitis B virus	347 (89.7)
Hepatitis C virus	12 (3.1)
Hepatitis B and C virus	9 (2.3)
Alcohol	6 (1.6)
Others	13 (3.4)
Time interval between MRI and pathology, day*	13.8 ± 13.3
Lesion (n = 422)	
Size, cm*	3.2 ± 2.1
Pathology	297 (70.4)
Size stability with typical imaging feature†	125 (29.6)
Final diagnosis	
HCC	234 (55.5)
Non-HCC malignancy	45 (10.7)
iCCA	24 (5.7)
cHCC-CCA	15 (3.6)
Metastasis	6 (1.4)
Benign lesion	143 (33.9)
RN/DN	45 (10.6)
Hemangioma	42 (10.0)
Arterioportal shunt	38 (9.0)
FNH or FNH-like nodule	12 (2.8)
Eosinophilic abscess	2 (0.5)
Inflammation	2 (0.5)
Fat necrosis	1 (0.2)
Adenoma	1 (0.2)

Unless indicated otherwise, data are the number of patients or lesions. Data in parentheses are percentages. *Data are presented as mean ± standard deviation, †Only for diagnosing benign lesions. cHCC-CCA = combined HCC-cholangiocarcinoma, FNH = focal nodular hyperplasia, HCC = hepatocellular carcinoma, iCCA = intrahepatic cholangiocarcinoma, RN/DN = regenerative nodule or dysplastic nodule

considering the major and ancillary features.

Gadoxetic Acid-Enhanced MRI

Gadoxetic acid-enhanced MRI studies were performed on either 3T (n = 252) or 1.5T (n = 135) MR machines. The details of the MRI are described in Supplementary Materials and Supplementary Table 2.

Imaging Features Analysis

All MRI features were reviewed independently by three fellowship-trained radiologists (with 1 year, 1 year, and 3 years of post-fellowship experience in evaluating abdominal imaging) who were unaware of the final diagnosis but aware that the study population consisted of high-risk patients with chronic hepatitis B or cirrhosis. All routine MRI sequences were provided to the reviewers. For each lesion, the size (cm), APHE, enhancing capsule, nonperipheral washout on PVP, transitional phase (TP) hypointensity, hepatobiliary phase (HBP) hypointensity, other ancillary features favoring malignancy, ancillary features favoring benignity (e.g., marked T2 hyperintensity), and a targetoid mass were evaluated according to the definition based on the LI-RADS v2018 (Supplementary Table 3) [3]. The ancillary features favoring malignancy included features favoring malignancy in general (e.g., mild-moderate T2 hyperintensity, restricted diffusion, TP hypointensity, HBP hypointensity, and corona enhancement) and those favoring HCC in particular (e.g., non-enhancing capsule, mosaic architecture, fat or blood products in mass) [7]. Threshold growth, subthreshold growth, and ultrasound visibility as discrete nodule were not evaluated. The major and ancillary features were determined by the three radiologists with 1 year, 1 year, and 3 years, respectively, of post-fellowship experience in evaluating abdominal imaging and used for per-reader data and per-consensus data. The following methods were used for the consensus data of each imaging feature: when the evaluations of two junior radiologists were concordant, their determined imaging features were used. When their evaluations were discordant, the imaging features were determined by a third radiologist. The categorization of the hepatic lesions was determined by the per-reader data and consensus data for each imaging feature.

The hepatic lesions were classified as an indeterminate nodule, probable HCC, and definite HCC, according to the 2018 KLCA-NCC PG (Fig. 1) [5]. The 2018 KLCA-NCC PG adopted the definitions of the major and ancillary features of LI-RADS v2018, except for the definition of washout (Supplementary Table 3). Definite HCC was defined as a lesion that met APHE with washout in the PVP, TP, or HBP, not showing either marked T2 hyperintensity or targetoid mass. The remaining lesions were regarded as probable HCC if they had ≥ 1 ancillary feature favoring malignancy in general and ≥ 1 ancillary feature favoring HCC in particular, while not showing marked T2 hyperintensity or a targetoid appearance. Finally, the remaining lesions were regarded as indeterminate nodules. Second-line imaging (e.g.,

https://doi.org/10.3348/kjr.2020.0846

Korean Journal of Radiology
KJR

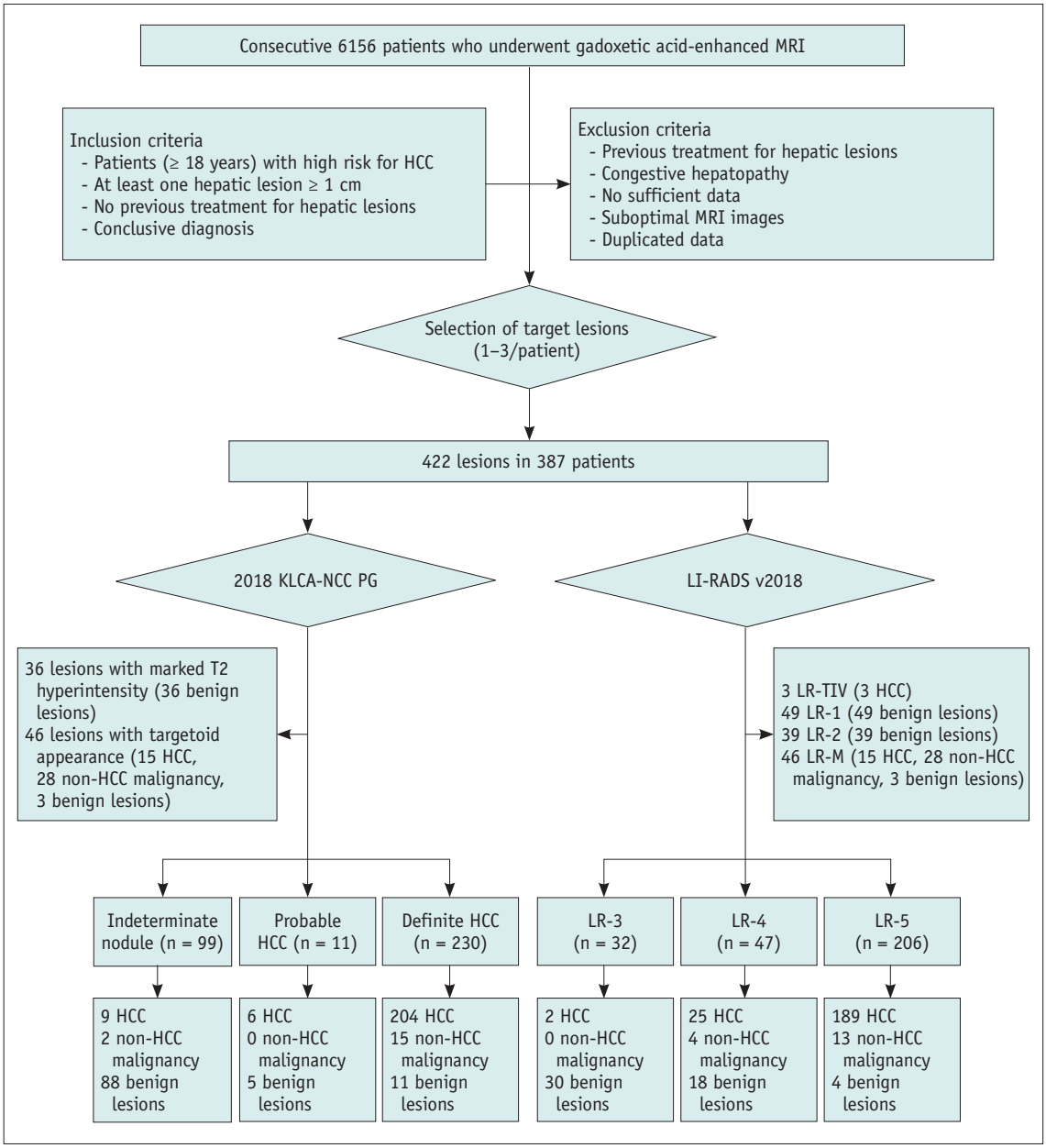

Fig. 1. Flow diagram of the study sample. HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-M = probably or definitely malignant but not HCC specific, LR-TIV = tumor in vein, LR-1 = definitely benign, LR-2 = probably benign, LR-3 = intermediate probability of malignancy, LR-4 = probably HCC, LR-5 = definitely HCC, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center Practice Guidelines

multiphase CT, contrast-enhanced ultrasound) was not considered in our study when the lesion was classified into a category. The laboratory findings (e.g., alpha-fetoprotein and peripheral eosinophil count) were not considered when the reviewers categorized the lesions.

Each lesion was also categorized as LR-TIV (tumor in vein), LR-M (probably or definitely malignant but not HCC specific), LR-1 (definitely benign) or 2 (probably benign), LR-3 (intermediate probability of malignancy), LR-4, and LR-5 based on LI-RADS v2018, and considerations of category adjustment using ancillary features and tie-breaking rules were made (Fig. 1). The category was upgraded by one up to LR-4 if the lesion had ≥ 1 ancillary feature(s) favoring malignancy, including ancillary features favoring malignancy in general and ancillary features favoring HCC in particular, without the presence of ancillary features favoring benignity. Conversely, the category was downgraded by one category when the lesion had ≥ 1

https://doi.org/10.3348/kjr.2020.0846

ancillary feature favoring benignity without the presence of ancillary features favoring malignancy. Distinctive nodules without associated major features or LR-M features were categorized differently according to size (< 2 cm or ≥ 2 cm, Supplementary Fig. 1).

Definite HCC and definite/probable HCC according to the 2018 KLCA-NCC PG and LR-5 and LR-5/4 according to the LI-RADS v2018 were used as index tests for diagnosing HCC.

Statistical Analysis

For overall lesions, the per-lesion sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), with 95% confidence intervals, for definite HCC and definite/probable HCC, as defined by 2018 KLCA-NCC PG for the diagnosis of HCC, were calculated using MedCalc's Diagnostic test evaluation calculator (https://www.medcalc.org/calc/diagnostic_test.php). The per-lesion sensitivity, specificity, PPV, and NPV of LR-5 and LR-5/4 according to LI-RADS v2018 were also calculated. Using the generalized estimating equation (GEE) model, the sensitivity and specificity of definite HCC (as per the 2018 KLCA-NCC PG) and LR-5 of the LI-RADS v2018 were compared to prevent patient cluster effects. The sensitivity and specificity of definite/probable HCC of 2018 KLCA-NCC PG and LR-5/4 of LI-RADS v2018 were also compared using GEE methods. For the subgroups of lesions of sizes < 2 cm and ≥ 2 cm,

Table 2. Per-Lesion Diagnostic Performance for 2018 KLCA-NCC PG and LI-RADS v2018 in the Diagnosis of HCC Using Consensus Data

	2018 KLCA-NCC PG Definite HCC	LI-RADS v2018 LR-5	*P*	2018 KLCA-NCC PG Definite/Probable HCC	LI-RADS v2018 LR-5/4	*P*
Overall (n = 422)						
SEN	87.2 (204/234) [82.2–91.2]	80.8 (189/234) [75.1–85.6]	< 0.001	89.7 (210/234) [83.2–91.9]	91.5 (214/234) [87.1–94.7]	0.204
SPE	86.2 (162/188) [80.4–90.8]	91.0 (171/188) [85.9–94.6]	0.002	83.5 (157/188) [77.4–88.5]	79.3 (149/188) [79.8–84.8]	0.071
PPV	88.7 (204/230) [84.6–91.8]	91.8 (189/206) [87.6–94.6]		87.1 (210/241) [83.0–90.4]	84.6 (214/253) [80.5–87.9]	
NPV	84.4 (162/192) [79.4–88.3]	79.2 (171/216) [74.4–83.2]		86.7 (157/181) [81.7–90.6]	88.2 (149/169) [83.0–91.9]	
< 2 cm (n = 173)						
SEN	88.9 (40/45) [76.0–96.3]	75.6 (34/45) [60.5–87.1]	0.009	91.1 (41/45) [78.8–97.5]	97.8 (44/45) [88.2–99.9]	0.073
SPE	89.1(114/128) [82.3–93.9]	94.5 (121/128) [89.1–97.8]	0.007	85.9 (110/128) [78.7–91.5]	81.3 (104/128) [73.4–87.6]	0.105
PPV	74.1 (40/54) [63.3–82.6]	82.9 (34/41) [69.9–91.1]		69.5 (41/59) [59.5–77.9]	64.7 (44/68) [56.0–72.5]	
NPV	95.8 (114/119) [90.9–98.1]	91.7 (121/132) [86.8–94.9]		96.5 (110/114) [91.5–98.6]	99.1 (104/105) [93.7–99.9]	
≥ 2 cm (n = 249)						
SEN	86.8 (164/189) [81.1–91.3]	82.0 (155/189) [75.8–87.2]	0.002	89.4 (169/189) [84.1–93.4]	90.0 (170/189) [84.8–93.8]	0.705
SPE	80.0 (48/60) [67.7–89.2]	83.3 (50/60) [71.5–91.7]	0.150	78.3 (47/60) [65.8–87.9]	75.0 (45/60) [62.1–85.3]	0.412
PPV	93.2 (164/176) [89.2–95.8]	93.9 (155/165) [89.8–96.5]		92.9 (169/182) [88.9–95.5]	91.9 (170/185) [87.9–94.6]	
NPV	65.8 (48/73) [56.6–73.9]	59.5 (50/84) [51.5–67.1]		70.2 (47/67) [60.3–78.4]	70.3 (45/64) [60.1–78.8]	

Data in parentheses were used to calculate percentages. Data in square brackets are 95% confidence intervals. *P* values were calculated using generalized estimating equations. HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-5 = definitely HCC, LR-5/4 = definitely HCC or probably HCC, NPV = negative predictive value, PPV = positive predictive value, SEN = sensitivity, SPE = specificity, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center Practice Guidelines

https://doi.org/10.3348/kjr.2020.0846

Korean Journal of Radiology
KJR

the diagnostic performances of 2018 KLCA-NCC PG and LI-RADS v2018 were evaluated and compared using GEE. Fleiss kappa values were used to evaluate the inter-reader agreement for each category and each imaging feature. The interpretation of Fleiss kappa was as follows: poor, < 0.00; slight, 0.00–0.20; fair, 0.21–0.40; moderate, 0.41–0.60; substantial, 0.61–0.80; almost perfect, 0.81–0.99 [8]. The statistical analyses were performed using MedCalc software version 19.1.3 (Mariakerke), SPSS version 26 (IBM Corp.), or R version 3.6.3 (R Foundation for Statistical Computing). A p value of < 0.05 was considered indicative of statistical significance.

RESULTS

Diagnostic Performance of the Definite HCC (2018 KLCA-NCC PG) and LR-5 (LI-RADS v2018)

The per-lesion diagnostic performance for each index test using the consensus data is summarized in Table 2. The sensitivity of definite HCC of 2018 KLCA-NCC PG for

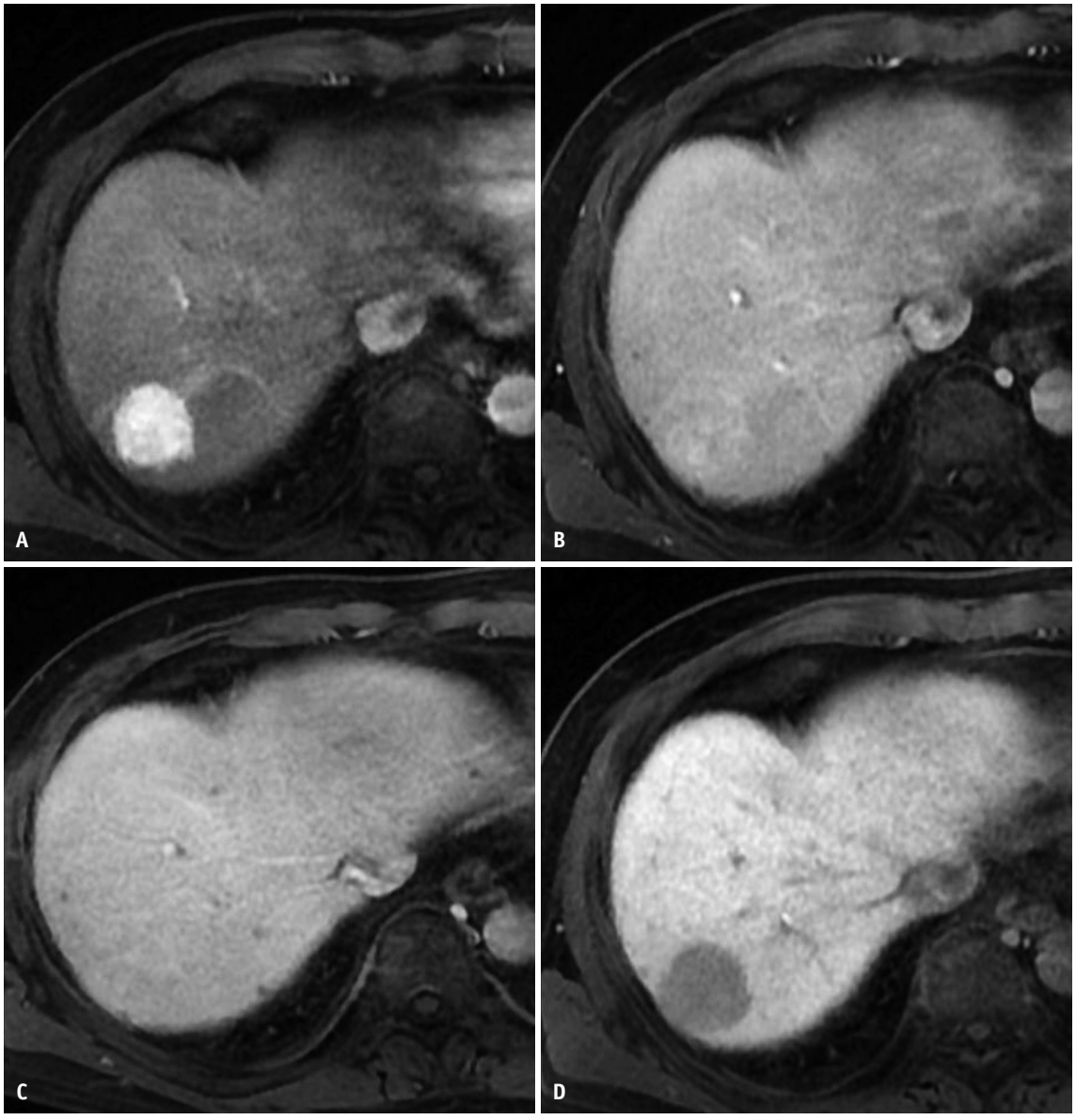

Fig. 2. A surgically-proven HCC in a 57-year-old male with chronic hepatitis B.
On gadoxetic acid-enhanced MRI, there is a 4-cm lesion with arterial phase hyperenhancement in segment 7 of the liver **(A)**, which does not show hypointensity on portal vein phase **(B)** or transitional phase **(C)** but hepatobiliary phase **(D)**. It could not be categorized as LR-5 according to the LI-RADS v2018 but it was classified as definite HCC according to the 2018 KLCA-NCC PG. HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-5 = definitely HCC, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center practice guidelines

https://doi.org/10.3348/kjr.2020.0846

diagnosing HCC was significantly higher than that of LR-5 of LI-RADS v2018 (87.2% [204/234] vs. 80.8% [189/234]; $p < 0.001$) (Fig. 2). However, definite HCC provided significantly lower specificity than LR-5 (86.2% [162/188] vs. 91.0% [171/188], $p = 0.002$). The per-reader diagnostic performance of the definite HCC and LR-5 criteria are listed in Supplementary Table 4.

For the diagnosis of HCC of size < 2 cm, definite HCC also had higher sensitivity (88.9% [40/45] vs. 75.6% [34/45], $p = 0.009$) and lower specificity (89.1% [114/128] vs. 94.5% [121/128], $p = 0.007$) than did LR-5. However, in diagnosing HCC of size ≥ 2 cm, definite HCC had a specificity comparable to that of LR-5 (80.0% [48/60] vs. 83.3% [50/60], $p = 0.150$) as well as a higher sensitivity than LR-5 (86.8% [164/189] vs. 82.0% [155/189], $p = 0.002$).

Diagnostic Performance of Definite/Probable HCC (2018 KLCA-NCC PG) and LR-5/4 (LI-RADS v2018)

For the diagnosis of HCC, there were no significant differences between the sensitivity or specificity of definite/probable HCC of 2018 KLCA-NCC PG and LR-5/4 of LI-RADS v2018 (89.7% [210/234] vs. 91.5% [214/234], $p = 0.204$; 83.5% [157/188] vs. 79.3% [149/188], $p = 0.071$). The per-reader diagnostic performance of the definite/probable HCC and LR-5/4 criteria is listed in Supplementary Table 3. For the diagnosis of HCC, definite/probable HCC did not show significant differences in sensitivity or specificity compared with LR-5/4 in both subgroups (< 2 cm and ≥ 2 cm).

Multiple comparisons of the index test for the diagnosis of HCC are shown in Figure 3 and Supplementary Table 5.

Categories and Imaging Features of HCC, Non-HCC Malignancies, and Benign Lesions

The categories of hepatic lesions according to the 2018 KLCA-NCC PG and LI-RADS v2018 are described in Figures 1, 4 and Table 3. The inter-reader agreements for index tests, major features, and ancillary features favoring malignancy were moderate to almost perfect (0.54. 0.82), except for enhancing capsule (0.28) and corona enhancement (0.09). However, the inter-reader agreements were minimal to fair for ancillary features favoring HCC in particular (0.13. 0.37) and for targetoid mass (0.10). The false-positive results of each index test for the diagnosis of HCC are listed in Supplementary Table 6. The inter-reader agreements and frequencies of major features and ancillary features of lesions are listed in Table 4. However, ≥ 1 ancillary feature(s) favoring HCC, in particular, were less frequently observed in HCCs of sizes < 2 cm than in those of sizes ≥ 2 cm (8.8% vs. 91.3%, $p = 0.002$). There was no difference in the frequency of major features and ancillary features favoring malignancy for HCCs of sizes < 2 cm and those of sizes ≥ 2 cm (Supplementary Table 7).

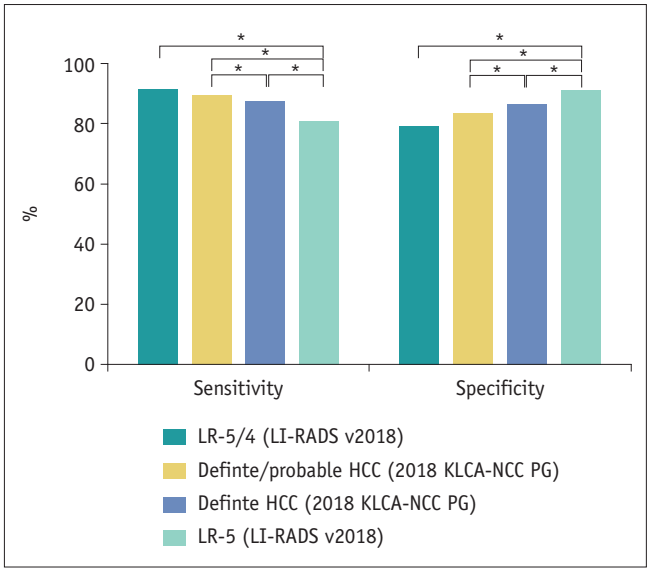

Fig. 3. Sensitivity and specificity of four index tests for the diagnosis of HCC. Of the four index tests for diagnosing HCC, LR-5/4 (91.5%) had the highest sensitivity, followed by definite/probable HCC of 2018 KLCA-NCC PG (89.7%), definite HCC of 2018 KLCA-NCC PG (87.2%), and LR-5 (80.8%). The specificity was the highest for LR-5 (91.0%), followed by definite HCC of 2018 KLCA-NCC PG (86.2%), definite/probable HCC of 2018 KLCA-NCC PG (83.5%), and LR-5/4 (79.3%). There were no statistical differences between the sensitivity and specificity for definite/probable HCC (2018 KLCA-NCC PG) and LR-5/4 (LI-RADS v2018); there were statistical differences in sensitivity and specificity for the other index tests. The multiple comparisons of the index tests for the diagnosis of HCC using the generalized estimating equation are listed in Supplementary Table 6. *Statistical difference between the two index tests. HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-5 = definitely HCC, LR-5/4 = definitely HCC or probably HCC, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center practice guidelines

DISCUSSION

We aimed to compare the diagnostic performance of the 2018 KLCA-NCC PG and LI-RADS v2018 criteria for the diagnosis of HCC on gadoxetic acid-enhanced MRI in patients with chronic hepatitis B or cirrhosis. Although there are several previous studies on the diagnostic performance of definite HCC criteria of the 2018 KLCA-NCC PG for diagnosing HCC [9,10], no studies have investigated the diagnostic performance of the 2018 KLCA-NCC PG

https://doi.org/10.3348/kjr.2020.0846

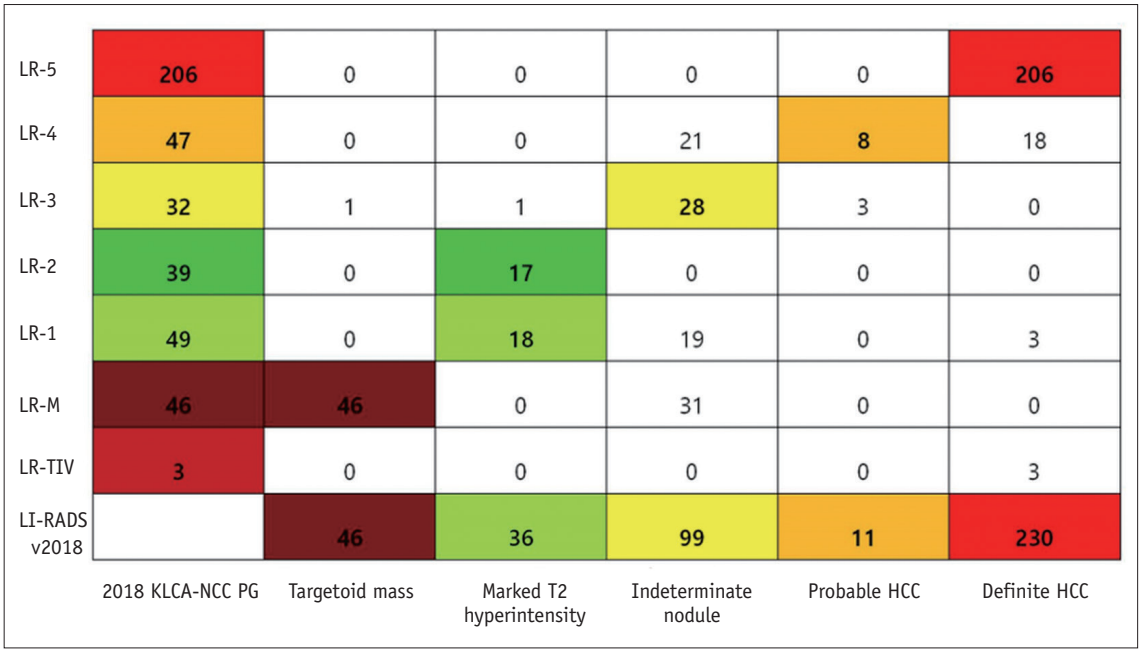

Fig. 4. Distribution of categories according to 2018 KLCA-NCC PG and LI-RADS v2018. The numbers indicate the number of lesions. HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-M = probably or definitely malignant but not HCC specific, LR-TIV= tumor in vein, LR-1 = definitely benign, LR-2 = probably benign, LR-3 = intermediate probability of malignancy, LR-4 = probably HCC, LR-5 = definitely HCC, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center Practice Guidelines

algorithm for the combination of definite HCC and probable HCC (definite/probable HCC) as well as definite HCC. Our results showed that when definite HCC or LR-5 was used as an index test for diagnosing HCC, 2018 KLCA-NCC PG had higher sensitivity and lower specificity than LI-RADS v2018. In diagnosing ≥ 2 cm HCCs, however, definite HCC (2018 KLCA-NCC PG) yielded higher sensitivity than the LR-5, without a reduction in the specificity. In addition, when using definite/probable HCC (2018 KLCA-NCC PG) or a combination of LR-5 and LR-4 (LR-5/4) as an index test for the diagnosis of HCC, 2018 KLCA-NCC PG had similar sensitivity and specificity to LI-RADS v2018.

The difference between the diagnostic performances of definite HCC (2018 KLCA-NCC PG) and LR-5 may be mainly attributed to the definition of washout for the diagnosis of HCC. Previous studies have reported increased sensitivity and decreased or comparable specificity when washout was evaluated during TP and HBP instead of during the PVP [9,10]. The 2018 KLCA-NCC PG applied washout during TP and HBP as well as during PVP of gadoxetic acid-enhanced MRI after excluding bright T2 lesions and targetoid lesions, resulting in 15 additional true positives and nine additional false-positives in our study. The 2018 KLCA-NCC PG showed significantly lower specificity than LI-RADS v2018 for the diagnosis of HCC of size < 2 cm; small hemangiomas showing APHE and HBP hypointensity but not marked T2 hyperintensity were categorized as definite HCC (2018 KLCA-NCC PG). This may lead to unnecessary treatments for benign lesions. In contrast, in diagnosing ≥ 2 cm HCCs, the definite HCC of the 2018 KLCA-NCC PG may be preferable over the LR-5 due to its increased sensitivity without a decrease in the specificity, especially in Eastern societies that more frequently use resection and locoregional treatments than transplantation for the treatment of HCC.

In our study, no statistical difference was observed between the sensitivities and specificities of the definite/probable HCC (2018 KLCA-NCC PG) and LR-5/4 (LI-RADS v2018), whereas the sensitivity of definite HCC was significantly higher than that of LR-5. A total of 11 lesions, including 6 HCCs and 5 non-HCCs, were categorized as probable HCC (LI-RADS v2018), whereas a total of 47 lesions, including 25 HCCs and 22 non-HCCs, were categorized as LR-4. This indicated that probable HCC (2018 KLCA-NCC PG) had a smaller number of false-positive results, suggesting a higher specificity than LR-4. Our results could be attributed to the stricter application of the ancillary features for the diagnosis of probable HCC (2018 KLCA-NCC PG), compared with the LI-RADS v2018. Since the probable HCC criteria of the 2018 KLCA-NCC PG

 https://doi.org/10.3348/kjr.2020.0846

Table 3. Distributions of Categories according to 2018 KLCA-NCC PG and LI-RADS v2018 Using Consensus Data

	No.	Incidence of HCC	Incidence of Malignancy
Overall (n = 422)			
2018 KLCA-NCC PG			
Definite HCC	230	88.7 (204/230)	95.2 (219/230)
Probable HCC	11	54.5 (6/11)	54.5 (6/11)
Indeterminate nodule	99	9.1 (9/99)	11.1 (11/99)
LI-RADS v2018			
LR-5	206	91.7 (189/206)	98.1 (202/206)
LR-4	47	53.2 (25/47)	61.7 (29/47)
LR-3	32	6.3 (2/32)	6.3 (2/32)
< 2 cm (n = 173)			
2018 KLCA-NCC PG			
Definite HCC	54	74.1 (40/54)	83.3 (45/54)
Probable HCC	5	20.0 (1/5)	20.0 (1/5)
Indeterminate nodule	82	4.9 (4/82)	4.9 (4/82)
LI-RADS v2018			
LR-5	41	82.9 (34/41)	90.2 (37/41)
LR-4	27	37.0 (10/27)	44.4 (12/27)
LR-3	30	3.3 (1/30)	3.3 (1/30)
≥ 2 cm (n = 249)			
2018 KLCA-NCC PG			
Definite HCC	176	93.2 (164/176)	98.9 (174/176)
Probable HCC	6	83.3 (5/6)	83.3 (5/6)
Indeterminate nodule	17	29.4 (5/17)	41.2 (7/17)
LI-RADS v2018			
LR-5	165	93.9 (155/165)	100.0 (165/165)
LR-4	20	75.0 (15/20)	85.0 (17/20)
LR-3	2	50.0 (1/2)	50.0 (1/2)

Data in parentheses were used to calculate percentages. HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-5 = definitely HCC, LR-4 = probably HCC, LR-3 = intermediate probability of malignancy, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center Practice Guidelines

require ≥ 1 ancillary features favoring HCC in particular and ≥ 1 ancillary features favoring malignancy in general, a lower sensitivity for diagnosing HCC, especially HCCs of sizes < 2 cm, might have been observed.

In actual clinical settings of Eastern societies, treatments (e.g., locoregional treatment) were performed for lesions diagnosed as probable HCC using an imaging modality without pathologic confirmation. In addition, biopsies for hepatic lesions are not always feasible due to the difficulty in accessing the area as well as concerns about complications such as bleeding and seeding risk [11,12]. Furthermore, it is not easy to differentiate HCC from DN, even by pathological examination of a biopsy specimen [13]. Therefore, the diagnostic performance of probable HCC as well as definite HCC is important. Although the diagnostic algorithm of the 2018 KLCA-NCC PG may be simpler than that of the LI-RADS v2018, definite/probable HCC of the 2018 KLCA-NCC PG had a similar diagnostic performance as that of the LR-5/4 of LI-RADS v2018. The complexity of the diagnostic algorithm of LI-RADS v2018 may have caused errors in categorization as well as moderate inter-reader agreements (intraclass correlation coefficient: 0.67–0.73) [14]. For instance, benign lesions with one or more ancillary features favoring malignancy will excessively upgrade to LR-4 without applying the rule that nodules < 2 cm without major or LR-M features (e.g., siderotic nodules, HBP hyperintense nodules) should be initially categorized as LR-2 before category adjustment. However, to further improve the diagnostic performance of the 2018 KLCA-NCC PG for the diagnosis of HCC, it may be necessary to modify the criteria for probable HCC or definite HCC depending on the size of the lesion.

Our study has several limitations. First, this was a retrospective study based on data from a single center. Our

https://doi.org/10.3348/kjr.2020.0846

Korean Journal of Radiology
KJR

study had only 45 HCCs of sizes < 2 cm since locoregional treatments are often performed for small HCCs without pathologic confirmation at our center. This may have caused selection bias. Second, our study did not include patients with non-cirrhotic chronic hepatitis C; they were included as a target population in the 2018 KLCA-NCC PG but excluded from the LI-RADS v2018. This may have led to a difference in the performance of the criteria for diagnosing HCC. Third, benign lesions were not always confirmed in our study, since those with atypical imaging features need to be pathologically confirmed. Finally, we did not consider imaging features such as threshold growth and ultrasound visibility, second-line imaging, and the peripheral eosinophil count when categorizing the lesions. According to the 2018 KLCA-NCC PG, second-line imaging methods, such as contrast-enhanced ultrasound,

Table 4. The Inter-Reader Agreements and Frequencies of Categories and Imaging Features of Lesions

	Kappa Value	HCC (n = 234)	Non-HCC Malignancies (n = 45)	Benign Lesions (n = 143)
Categories				
Definite HCC (2018 KLCA-NCC PG)	0.64	87.20 (204/234)	33.30 (15/45)	7.7 (11/143)
Definite/probable HCC (2018 KLCA-NCC PG)	0.62	89.70 (210/234)	33.30 (15/45)	11.20 (16/143)
LR-5 (LI-RADS v2018)	0.71	80.80 (189/234)	28.90 (13/45)	2.8 (4/143)
LR-5/4 (LI-RADS v2018)	0.7	91.50 (214/234)	37.80 (17/45)	15.40 (22/143)
Major features				
≥ 20 mm	0.82	73.1 (171/234)	26.7 (12/45)	2.8 (4/143)
APHE	0.55	90.2 (211/234)	40.0 (18/45)	55.9 (80/143)
Washout	0.54	88.9 (208/234)	68.9 (31/45)	13.3 (19/143)
Enhancing capsule	0.28	21.8 (51/234)	2.2 (1/45)	1.4 (2/143)
Combination of APHE and washout				
APHE + washout*	0.72	87.7 (192/219)	76.5 (13/17)	3.8 (4/104)
APHE + TP hypointensity†	0.70	91.8 (201/219)	88.2 (15/17)	4.8 (5/104)
APHE + HBP hypointensity‡	0.71	93.2 (204/219)	88.2 (15/17)	10.6 (11/104)
Ancillary features favoring malignancy				
Favoring malignancy in general				
Mild-to-moderate T2 hyperintensity	0.78	97.9 (229/234)	95.6 (43/45)	12.6 (18/143)
Restricted diffusion	0.67	93.2 (218/234)	93.3 (42/45)	9.8 (14/143)
HBP hypointensity	0.8	98.7 (231/234)	100.0 (45/45)	60.1 (86/143)
TP hypointensity	0.69	97.0 (227/234)	100.0 (45/45)	40.6 (58/143)
Corona enhancement	0.09	7.3 (17/234)	8.9 (4/45)	2.1 (3/143)
Favoring HCC in particular				
Non-enhancing capsule	0.13	7.3 (17/234)	0	0
Mosaic architecture	0.22	15.0 (35/234)	2.2 (1/45)	0
Fat in mass, more than adjacent liver	0.37	18.4 (43/234)	4.4 (2/45)	4.2 (6/143)
Blood products in mass	0.29	12.8 (30/234)	4.4 (2/45)	0
Ancillary features favoring benignity				
Marked T2 hyperintensity	0.71	0	0	25.20 (36/143)
No mass effect	< 0.01	0	0	0
HBP isointensity	0.64	2.1 (5/234)	0	28.0 (40/143)
Parallel enhancement	0.37	0	0	10.50 (15/143)
Marked T2 hypointensity	0.34	0	2.20 (1/45)	3.50 (5/143)
Targetoid mass	0.1	6.4 (15/234)	62.2 (28/45)	2.1 (3/143)

Data in parentheses were used to calculate percentages. Fleiss' Kappa value of each major feature for the three readers was calculated. *Lesions with APEH and washout on portal venous phase not showing either marked T2 hyperintensity or targetoid mass, †Lesions with APEH and TP hypointensity not showing either marked T2 hyperintensity or targetoid mass, ‡Lesions with APEH and HBP hypointensity not showing either marked T2 hyperintensity or targetoid mass. APHE = arterial phase hyperenhancement, HBP = hepatobiliary phase, HCC = hepatocellular carcinoma, LI-RADS v2018 = Liver Imaging Reporting and Data System version 2018, LR-5 = definitely HCC, LR-5/4 = definitely HCC or probably HCC, TP = transitional phase, 2018 KLCA-NCC PG = 2018 Korean Liver Cancer Association-National Cancer Center Practice Guidelines

https://doi.org/10.3348/kjr.2020.0846

are recommended when first-line imaging is inconclusive. Further, the peripheral eosinophil count should also be checked before the diagnosis of HCC can be made [5].

In conclusion, for the diagnosis of HCCs of sizes ≥ 2 cm, the definite HCC (2018 KLCA-NCC PG) had a higher sensitivity than the LR-5, without a reduction in the specificity. The definite/probable HCC (2018 KLCA-NCC PG) had similar sensitivity and specificity to those of the LR-5/4.

Supplement

The Data Supplement is available with this article at https://doi.org/10.3348/kjr.2020.0846.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Acknowledgments

The authors would like to thank Professor Junhee Han, Department of Statistics and Institute of Statistics, Hallym University, for his statistical assistance.

Author Contributions

Conceptualization: Jeong Min Lee. Data curation: Sang Min Lee, Su Joa Ahn, Hyo-Jin Kang, Hyun Kyung Yang. Formal analysis: Sang Min Lee. Investigation: Jeong Min Lee, Sang Min Lee. Methodology: Jeong Min Lee, Sang Min Lee. Project administration: Jeong Min Lee, Sang Min Lee. Visualization: Sang Min Lee. Writing—original draft: Sang Min Lee. Writing—review & editing: Jeong Min Lee, Jeong Hee Yoon.

ORCID iDs

Sang Min Lee
https://orcid.org/0000-0001-7719-3849
Jeong Min Lee
https://orcid.org/0000-0003-0561-8777
Su Joa Ahn
https://orcid.org/0000-0001-9026-0399
Hyo-Jin Kang
https://orcid.org/0000-0002-6771-2112
Hyun Kyung Yang
https://orcid.org/0000-0003-3576-9146
Jeong Hee Yoon
https://orcid.org/0000-0002-9925-9973

REFERENCES

1. Choo SP, Tan WL, Goh BKP, Tai WM, Zhu AX. Comparison of hepatocellular carcinoma in Eastern versus Western populations. *Cancer* 2016;122:3430-3446
2. Kim TH, Kim SY, Tang A, Lee JM. Comparison of international guidelines for noninvasive diagnosis of hepatocellular carcinoma: 2018 update. *Clin Mol Hepatol* 2019;25:245-263
3. Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, et al. Liver Imaging Reporting and Data System (LI-RADS) version 2018: imaging of hepatocellular carcinoma in at-risk patients. *Radiology* 2018;289:816-830
4. European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018;69:182-236
5. Korean Liver Cancer A, National Cancer Center GK. 2018 Korean Liver Cancer Association-National Cancer Center Korea Practice Guidelines for the management of hepatocellular carcinoma. *Korean J Radiol* 2019;20:1042-1113
6. Lee SM, Lee JM, Ahn SJ, Kang HJ, Yang HK, Yoon JH. LI-RADS version 2017 versus version 2018: diagnosis of hepatocellular carcinoma on gadoxetate disodium-enhanced MRI. *Radiology* 2019;292:655-663
7. Cerny M, Chernyak V, Olivié D, Billiard JS, Murphy-Lavallée J, Kielar AZ, et al. LI-RADS version 2018 ancillary features at MRI. *Radiographics* 2018;38:1973-2001
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1997;33:159-174
9. Joo I, Lee JM, Lee DH, Jeon JH, Han JK. 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* 2019;29:1724-1732
10. Kim DH, Choi SH, Kim SY, Kim MJ, Lee SS, Byun JH. Gadoxetic acid-enhanced MRI of hepatocellular carcinoma: value of washout in transitional and hepatobiliary phases. *Radiology* 2019;291:651-657
11. Silva MA, Hegab B, Hyde C, Guo B, Buckels JA, Mirza DF. Needle track seeding following biopsy of liver lesions in the diagnosis of hepatocellular cancer: a systematic review and meta-analysis. *Gut* 2008;57:1592-1596
12. Young AL, Lodge JP. Needle-track seeding following biopsy of liver lesions in the diagnosis of hepatocellular cancer: a systematic review and meta-analysis. *Gut* 2009;58:887-888
13. Roskams T, Kojiro M. Pathology of early hepatocellular carcinoma: conventional and molecular diagnosis. *Semin Liver Dis* 2010;30:17-25
14. Fowler KJ, Tang A, Santillan C, Bhargavan-Chatfield M, Heiken J, Jha RC, et al. Interreader reliability of LI-RADS version 2014 algorithm and imaging features for diagnosis of hepatocellular carcinoma: a large international multireader study. *Radiology* 2018;286:173-185