Intraindividual Comparison between Gadoxetate-Enhanced Magnetic Resonance Imaging and Dynamic Computed Tomography for Characterizing Focal Hepatic Lesions: A Multicenter, Multireader Study

Chansik An, MD1, Chang Hee Lee, MD2, Jae Ho Byun, MD3, Min Hee Lee, MD4, Woo Kyoung Jeong, MD5, Sang Hyun Choi, MD1, Do Young Kim, MD6, Young-Suk Lim, MD7, Young Seok Kim, MD8, Ji Hoon Kim, MD9, Moon Seok Choi, MD10, Myeong-Jin Kim, MD1

1Department of Radiology, Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea; Departments of 2Radiology and 3Internal Medicine, Korea University Guro Hospital, Korea University College of Medicine, Seoul, Korea; Department of Radiology and Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Korea; Departments of 4Radiology and 8Internal Medicine, Soonchunhyang University Bucheon Hospital, Bucheon, Korea; Departments of 5Radiology and 10Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea; 6Department of Internal Medicine, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea; 7Department of Internal Medicine, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Korea

Objective: To compare the diagnostic accuracy of dynamic computed tomography (CT) and gadoxetate-enhanced magnetic resonance imaging (MRI) for characterization of hepatic lesions by using the Liver Imaging Reporting and Data System (LI-RADS) in a multicenter, off-site evaluation.

Materials and Methods: In this retrospective multicenter study, we evaluated 231 hepatic lesions (114 hepatocellular carcinomas [HCCs], 58 non-HCC malignancies, and 59 benign lesions) confirmed histologically in 217 patients with chronic liver disease who underwent both gadoxetate-enhanced MRI and dynamic CT at one of five tertiary hospitals. Four radiologists at different institutes independently reviewed all MR images first and the CT images 4 weeks later. They evaluated the major and ancillary imaging features and categorized each hepatic lesion according to the LI-RADS v2014. Diagnostic performance was calculated and compared using generalized estimating equations.

Results: MRI showed higher sensitivity and accuracy than CT for diagnosing hepatic malignancies; the pooled sensitivities, specificities, and accuracies for categorizing LR-5/5V/M were 59.0% vs. 72.4% (CT vs. MRI; \( p < 0.001 \)), 83.5% vs. 83.9% (\( p = 0.906 \)), and 65.3% vs. 75.3% (\( p < 0.001 \)), respectively. CT and MRI showed comparable capabilities for differentiating between HCC and other malignancies, with pooled accuracies of 79.9% and 82.4% for categorizing LR-M, respectively (\( p = 0.139 \)).

Conclusion: Gadoxetate-enhanced MRI showed superior accuracy for categorizing LR-5/5V/M in hepatic malignancies in comparison with dynamic CT. Both modalities had comparable accuracies for distinguishing other malignancies from HCC.

Keywords: Hepatocellular carcinoma; Magnetic resonance imaging;Computed tomography; Contrast media; Data systems; Gadolinium ethoxybenzyl DTPA

Received May 29, 2019; accepted after revision August 15, 2019.

This study was supported by a grant from Bayer Korea Ltd.

Corresponding author: Myeong-Jin Kim, MD, Department of Radiology, Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.

Tel: (82) 2228-7400 • Fax: (82) 393-3035 • E-mail: kimnex@yuhs.ac

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
INTRODUCTION

Current guidelines for the management of hepatocellular carcinoma (HCC) allow the diagnosis of HCC in high-risk patients without histologic confirmation if a hepatic lesion shows typical findings on dynamic contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI) (1-5). However, non-HCC malignancies such as intrahepatic cholangiocarcinoma (CCA) or combined HCC-CCA may also display imaging findings of typical HCCs, particularly in small tumors in the cirrhotic liver (6-8). Since the treatment options and prognoses are different for non-HCC malignancies and HCC, imaging-based differentiation of HCC and other malignancies is important (9, 10).

The Liver Imaging Reporting and Data System (LI-RADS), unlike other diagnostic criteria, addresses the issue of imaging diagnosis of non-HCC malignancies (11, 12). It requires the LR-M category to be assigned to a hepatic lesion suspected to be a non-HCC malignancy, and recommends histological diagnosis of these LR-M lesions (13, 14).

Although previous studies have compared the LI-RADS diagnostic accuracy of gadoxetate-enhanced MRI and dynamic CT (15-18), a multicenter study involving off-site evaluations by multiple readers has not been performed. Therefore, the purpose of this study was to compare the diagnostic performance of gadoxetate-enhanced MRI and dynamic CT for assessments based on LI-RADS v2014 during characterization of focal hepatic lesions in patients with chronic liver disease.

MATERIALS AND METHODS

Subjects

This retrospective multicenter study was approved by the Institutional Review Boards of all five participating institutions, which are academic tertiary hospitals in Korea. The requirement for written informed consent for the retrospective review of medical records and image data was waived.

Investigators from five institutions searched the databases to identify patients aged 18 years or older who met the following inclusion criteria: 1) at risk of developing HCC (hepatitis B virus [HBV] carrier, hepatitis C virus carrier with chronic hepatitis, or cirrhosis from any causes), 2) pathologic diagnosis of focal hepatic lesions based on surgery or biopsy findings between January 2008 and July 2014, 3) both dynamic CT and gadoxetate-enhanced MRI performed within 2 months before the pathologic diagnosis, and 4) no past history of hepatic malignancy. Of the 2958 patients who met the inclusion criteria, 287 were excluded because they had undergone local ablative treatments (n = 274), had ruptured HCCs (n = 3), had active extrahepatic malignancy (n = 6), or had false-negative biopsy results (n = 4). From the remaining 2671 patients, 103 patients with 117 non-HCC lesions (59 benign lesions and 58 non-HCC malignancies) were included; we also randomly selected 114 patients with 114 HCCs. Our final study population consisted of 217 patients with 231 pathologically-confirmed hepatic lesions (Fig. 1).

Image Acquisition

MRI was performed using 1.5T or 3T scanners. The protocols included a dual-echo T1-weighted gradient-recalled echo sequence, moderately and heavily T2-weighted turbo spin-echo sequences, dynamic three-dimensional T1-weighted gradient-recalled echo sequence using gadoxetic acid disodium (Primovist, Bayer AG, Berlin, Germany) as
the contrast medium, and diffusion-weighted imaging. For
dynamic imaging, arterial-, portal venous-, transitional-, and
hepatobiliary-phase images were acquired 25–35
seconds, 60–70 seconds, 2 or 3 minutes, and 15 or 20
minutes after contrast injection, respectively. Contrast-
enhanced dynamic CT was performed using 16-, 40-, or
64-detector scanners. After acquiring unenhanced images,
arterial-, portal venous-, and delayed-phase images were
obtained approximately 30 seconds, 70 seconds or 90
seconds, and 180 seconds or 210 seconds after iodinated
contrast medium injection, respectively. Technical details
for CT and MRI acquisition are provided in Supplementary
Material 1.

Data Collection and Preparation
Clinical, pathologic, and imaging data were retrieved
from the database of each participating institution, which
included information about patient demographics, the
cause of chronic liver disease, serum alpha-fetoprotein
level, child-Pugh class, pathologic diagnosis, and CT/MR
images. They were then sent to the central site (Severance
Hospital), where the image data were anonymized,
randomized, and assigned new identification numbers.
The processed imaging data, along with screen-captured
images marking the lesions’ sites (which was performed by
an abdominal radiologist with 4 years of experience), were
then sent back for image analysis. The data collection and
preparation steps are described in detail in Supplementary
Material 2.

Image Analysis
Four board-certified abdominal radiologists from different
institutions (Reviewer 1, Reviewer 2, Reviewer 3, and
Reviewer 4, with 5, 6, 10, and 21 years of experience,
respectively, in abdominal CT and MRI) independently
analyzed the CT/MR images and submitted the results to
the central site. In the first round, they analyzed the MR
images to evaluate the LI-RADS features and determine the
initial LI-RADS category in each case without assessing
the ancillary features. Then, they evaluated the ancillary
features and assigned the final LI-RADS category. They did
not evaluate the presence or absence of threshold growth.
In the second round conducted at least 8 weeks after the
first round to minimize recall bias, they evaluated the CT
images in the same manner. In cases where a hepatic lesion
could not be delineated in the marked area, it was recorded
as “not visible” and considered benign. The washout
appearance was determined twice, first as defined in the
LI-RADS by using the portal venous phase alone (11, 13,
14), and then by using both portal- and transitional-phase
images. The reviewers were asked to characterize a feature
as absent if they could not unequivocally determine whether
the feature was present or absent due to suboptimal or poor
image quality.

A detailed description of the image analysis protocol is
provided in Supplementary Material 3.

Statistical Analysis
For sample size estimation, the expected per-lesion
sensitivities of dynamic CT and MRI for HCC were assumed
to be 68% and 80%, respectively, with the same specificity
of 94% (19). On the basis of these expected values and a
power of 80%, we calculated that we would need 209 or
more hepatic lesions to obtain a significant difference in
diagnostic performance with a two-sided type I error of 5%
(20).

The baseline characteristics were compared using
analysis of variance for continuous variables and the chi-
squared test for categorical variables. Multiple comparisons
were corrected using Bonferroni’s method. The frequency
of image features was compared using chi-squared or
Fisher’s exact tests. We used a generalized estimating
equation method to compare the diagnostic performance
between dynamic CT and MRI for distinguishing hepatic
malignancies from benign lesions categorized as LR-5, LR-
5V, or LR-M; for differentiating HCC from non-HCC lesions
categorized as LR-5 or LR-5V; for differentiating non-HCC
malignancies from other lesions categorized as LR-M; and
then in subgroups containing only two disease entities
for differentiating between HCC and benign lesions (with
other malignancies excluded) and between HCC and other
malignancies (with benign lesions excluded). In addition,
we examined the diagnostic performance of LR-5 (without
LR-5V) for HCC. We used LR-5 alone, not including LR-4, for
the calculation of diagnostic accuracy, because LR-5 alone
is considered to indicate “definite HCC” in LI-RADS (11,
13, 14). Lastly, we calculated the diagnostic performance
of gadoxetate-enhanced MRI when using the portal phase
alone for washout evaluation and compared it with the
results obtained using both the portal and transitional
phases. Pathologic diagnosis was used as a reference
standard. To examine the added value of the ancillary
features, we tabulated the LI-RADS categories before and
after applying the ancillary features and calculated the
Intraindividual Multicenter Comparison of EOB-MRI and CT for Diagnosing HCC

RESULTS

Baseline Characteristics, LI-RADS Categories, and Major Features

The baseline characteristics of the 217 patients with 231 hepatic lesions (114 HCCs, 58 other malignancies, and 59 benign lesions) are summarized in Table 1. Patients with other malignancies were significantly older compared to those with HCCs ($p = 0.017$) and benign lesions ($p = 0.006$). HBV was a more common etiology of chronic liver disease in patients with HCCs than in those with benign lesions ($p = 0.003$). Tumor size was the largest in patients with other malignancies and the smallest in patients with benign lesions ($p < 0.001$), and biopsy was more frequently used for the diagnosis of other malignancies or benign lesions than HCCs ($p < 0.001$). The interval between CT and MRI ranged from 0 to 64 days, with a median interval of 12 days. Pathologic diagnosis was performed within 2 months after the first imaging study as it was one of our inclusion net reclassification improvement. The net reclassification improvement is a statistic for assessing the improvement in performance gained by adding a new factor to a model by measuring the extent to which individual subjects with and without disease are appropriately reclassified into more appropriate categories (21). We performed this analysis only on MRI cases after excluding the LR-M and “not visible” cases. Kappa statistics were computed as indices of inter-reader agreements between the four readers. Kappa statistics were used to indicate agreement, with 0.8–1.0 indicating excellent agreement; 0.60–0.79, good agreement; 0.40–0.59, moderate agreement; 0.20–0.39, fair agreement; and 0–0.19, poor agreement. All statistical analyses were performed using SAS 9.2 (SAS Institute Inc., Cary, NC, USA) and R version 3.3.3 (The R Foundation for Statistical Computing, Vienna, Austria). $P$ values < 0.05 were considered statistically significant. A detailed description of our statistical analysis is provided in Supplementary Material 4.

Table 1. Baseline Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>HCC</th>
<th>OM*</th>
<th>Benign†</th>
<th>$P$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per-patient basis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of patients</td>
<td>114</td>
<td>53</td>
<td>50</td>
<td></td>
<td>217</td>
</tr>
<tr>
<td>Median age (range, years)</td>
<td>59 (40–85)</td>
<td>63 (42–76)</td>
<td>55 (36–78)</td>
<td>0.003</td>
<td>59 (36–85)</td>
</tr>
<tr>
<td>Sex (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.835</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>89 (78.1)</td>
<td>40 (75.5)</td>
<td>37 (74.0)</td>
<td></td>
<td>166 (76.5)</td>
</tr>
<tr>
<td>Female</td>
<td>25 (21.9)</td>
<td>13 (24.5)</td>
<td>13 (26.0)</td>
<td></td>
<td>51 (23.5)</td>
</tr>
<tr>
<td>Cause of chronic liver disease (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>HBV</td>
<td>98 (86.8)</td>
<td>38 (71.7)</td>
<td>32 (64.0)</td>
<td></td>
<td>168 (77.9)</td>
</tr>
<tr>
<td>HCV</td>
<td>7 (6.1)</td>
<td>2 (3.8)</td>
<td>3 (6.0)</td>
<td></td>
<td>12 (5.5)</td>
</tr>
<tr>
<td>HBV/HCV coinfection</td>
<td>0 (0)</td>
<td>1 (1.9)</td>
<td>0 (0)</td>
<td></td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Non-viral</td>
<td>9 (7.0)</td>
<td>12 (22.6)</td>
<td>15 (30.0)</td>
<td></td>
<td>36 (16.1)</td>
</tr>
<tr>
<td>Child-Pugh class (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.285</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>102 (89.5)</td>
<td>50 (94.3)</td>
<td>48 (96.0)</td>
<td></td>
<td>200 (92.2)</td>
</tr>
<tr>
<td>B</td>
<td>10 (8.8)</td>
<td>3 (5.7)</td>
<td>2 (4.0)</td>
<td></td>
<td>15 (6.9)</td>
</tr>
<tr>
<td>C</td>
<td>2 (1.7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Median AFP (range, ng/mL)</td>
<td>9.35 (0.6–50676.0)</td>
<td>4.4 (1.3–1771.0)</td>
<td>3.9 (0.9–55.1)</td>
<td>&lt; 0.001</td>
<td>5.5 (0.6–50676.0)</td>
</tr>
<tr>
<td>Per-lesion basis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of lesions</td>
<td>114</td>
<td>58</td>
<td>59</td>
<td></td>
<td>231</td>
</tr>
<tr>
<td>Median size (range, cm)</td>
<td>2.3 (0.5–12.2)</td>
<td>2.8 (0.5–13.2)</td>
<td>1.5 (0.7–7.1)</td>
<td>&lt; 0.001</td>
<td>2.1 (0.3–13.3)</td>
</tr>
<tr>
<td>Method of diagnosis (%)</td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Biopsy</td>
<td>3 (2.6)</td>
<td>12 (20.7)</td>
<td>22 (37.3)</td>
<td></td>
<td>37 (16.0)</td>
</tr>
<tr>
<td>Resection</td>
<td>110 (96.5)</td>
<td>46 (79.3)</td>
<td>35 (59.3)</td>
<td></td>
<td>191 (82.7)</td>
</tr>
<tr>
<td>Transplantation</td>
<td>1 (0.9)</td>
<td>0 (0)</td>
<td>2 (3.4)</td>
<td></td>
<td>3 (1.3)</td>
</tr>
</tbody>
</table>

*OM includes 30 intrahepatic cholangiocarcinomas, 15 combined hepatocellular-cholangiocarcinomas, 11 metastases, 1 neuroendocrine carcinoma, and 1 lymphoepithelioma-like carcinoma, †Benign lesion includes 29 dysplastic/regenerative nodules, 8 focal nodular hyperplasia-like nodules, 8 hemangiomas, 5 angiomylipomas, 3 bile duct adenomas, 2 adenomas, 2 eosinophilic abscesses, 1 inflammatory pseudotumor, and 1 tuberculosis. AFP = alpha-fetoprotein, HBV = hepatitis B virus, HCC = hepatocellular carcinoma, HCV = hepatitis C virus, OM = other malignancies
criteria.

The frequencies of HCCs, other malignancies, and benign lesions according to the LI-RADS category and imaging modality are presented in Figure 2 (only pooled results) and Supplementary Table 1 (including full results). Minor proportions of non-HCC malignancies were categorized as LR-5V (3.4% [8/232] for CT and 4.7% [11/232] for MRI) (Figs. 2, 3).

The frequencies of major features—arterial-phase hyperenhancement, washout appearance, and capsule appearance—according to imaging modality are presented in Supplementary Tables 2 (by disease) and 3 (by size, only including HCCs). There were seven cases in which at least one reviewer decided that evaluation of arterial-phase MR images was suboptimal because of severe motion artifacts or too early acquisition. In HCC lesions, diffuse arterial hyperenhancement and washout appearance either in the portal or transitional phase were more frequently observed on MRI than on CT (88.2% vs. 82.5% [p = 0.015] and 90.4% vs. 77.2% [p = 0.005], respectively), but the frequency of washout appearance observed on MRI became lower than that on CT when the washout appearance was determined in the portal phase alone (71.1% vs. 77.2% [p = 0.034]). Rim-like arterial-phase hyperenhancement, a feature favoring malignancies other than HCC, was more frequently observed on MRI than on CT in non-HCC malignancies (44.8% vs. 32.3%, p = 0.006), but not in HCC (4.6% vs. 6.8%, p = 0.205) or benign lesions (7.2% vs. 5.5%, p = 0.572).

### Diagnostic Performance of Dynamic CT and MRI

#### Diagnosis of Hepatic Malignancies

In the diagnosis of hepatic malignancies, MRI showed higher sensitivity and accuracy than CT, but the specificities of the two methods were not significantly different (Table 2 [only pooled results], Supplementary Table 4 [including full results], and Fig. 4). With data pooled from all the reviewers, the sensitivity, specificity, and accuracy were 72.4%, 83.9%, and 75.3%, respectively, for MRI, and 59.0%, 83.5%, and 65.3%, respectively, for CT (p < 0.001 for sensitivity and accuracy, and p = 0.906 for specificity). For differentiation of HCC from non-HCC lesions, MRI showed higher sensitivity but lower specificity than CT, and accuracy was not significantly different with or without including LR-5V in the diagnostic criteria (p > 0.576). For differentiation of other malignancies from HCC and benign lesions, no significant differences were observed in sensitivity, specificity, and accuracy between CT and MRI (p = 0.139).

---

Fig. 2. Frequencies and proportions of HCC, OM, and benign lesions according to imaging modality and Liver Imaging Reporting and Data System category. Numbers and areas of segments in each vertical bar indicate numbers and proportions of HCC, OM, and benign lesions, respectively. Pooled results from four reviewers are shown here (see Supplementary Table 1 for full results). NV = not visible, OM = other malignancies.
Analyses of Subgroups Containing Only Two Disease Entities

In a subgroup with HCC and benign lesions (other malignancies excluded), MRI showed higher sensitivity and accuracy than CT for the diagnosis of HCC (with the pooled data, 62.5% vs. 52.4% \( p = 0.003 \) and 72.0% vs. 66.2% \( p = 0.019 \), respectively) with similar specificities (90.3% vs. 92.8%, \( p = 0.352 \)). However, in a subgroup with HCC and other malignancies (benign lesions excluded), the sensitivity, specificity, and accuracy of MRI and CT for discriminating other malignancies from HCC were not significantly different \( (p > 0.150) \) (Table 3 [only pooled results] and Supplementary Table 5 [including full results]). Similar results were obtained when LR-5V was not considered diagnostic for HCC.

Comparison of the Portal and Delayed Phases for Evaluating the Washout Appearance

When the portal phase alone was used for evaluating the washout appearance, the pooled results demonstrated that the specificity significantly increased \( (p = 0.049) \) but

| Fig. 3. Non-HCC malignancy with tumor in vein. Diffuse hypervascular tumor with infiltrative margins is seen at right hemi-liver. Tumor also invades adjacent portal vein branch (P8), forming mass within vein (arrows). Two of our four reviewers categorized this mass as LR-5V, while other two assigned score of LR-M. Pathologic diagnosis obtained after biopsy was combined HCC-cholangiocarcinoma. |

<p>| Table 2. Diagnostic Performance of CT and Gadoxetate-Enhanced MRI for Hepatic Malignancy |
|--------------------------------------------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>CT MRI P</th>
<th>CT MRI P</th>
<th>CT MRI P</th>
<th>CT MRI P</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP/FN</td>
<td>39/282 38/190</td>
<td>44/217 79/171</td>
<td>35/229 66/184</td>
<td>45/115 38/89</td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>59.0 72.4 &lt; 0.001</td>
<td>52.4 62.5 0.003</td>
<td>49.8 59.7 0.004</td>
<td>50.4 57.8 0.167</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>83.5 83.9 0.906</td>
<td>90.6 83.1 0.002</td>
<td>92.5 85.9 0.003</td>
<td>93.5 94.5 0.494</td>
</tr>
<tr>
<td>Accuracy, %</td>
<td>65.3 75.3 &lt; 0.001</td>
<td>71.8 72.9 0.576</td>
<td>71.4 72.9 0.468</td>
<td>82.7 85.3 0.139</td>
</tr>
</tbody>
</table>

These are pooled results from four reviewers. Unless otherwise specified, data represent numbers of cases. CT = computed tomography, FN = false negative, FP = false positive, MRI = magnetic resonance imaging, TN = true negative, TP = true positive.
sensitivity and accuracy significantly decreased ($p < 0.002$) compared to the findings obtained using both portal and transitional phases in the diagnosis of hepatic malignancies (Table 4 [only pooled results] and Supplementary Table 6 [including full results]). However, despite the decrease in sensitivity, MRI showed significantly higher sensitivity ($67.7\% \text{ vs. } 59.0\%$, $p < 0.001$) and accuracy ($72.5\% \text{ vs. } 65.3\%$, $p < 0.001$) than CT (Table 4).

**Added Value of the Ancillary Features in Gadoxetate-Enhanced MRI**

Analysis of the pooled data revealed that the ancillary features modified the final LI-RADS category in 110 (12.0%) of 918 cases (Supplementary Table 7). None of the LR-4 cases was upgraded to LR-5, as it is not warranted by the LI-RADS (7). Of the 686 malignant cases, 22 (3.2%) were correctly upgraded into higher LI-RADS categories, while 25 (3.6%) were incorrectly downgraded into lower categories. Of the 232 benign cases, 36 (15.5%) and 28 (12.1%) were reclassified into higher and lower LI-RADS categories, respectively. The overall net reclassification improvement was estimated to be $-0.055$ ($p = 0.117$), indicating a negative impact on the overall categorization without statistical significance. However, of 32 and 19 benign cases initially categorized as LR-5/5V and LR-M, 10 (31.3%) and 4 (21.1%), respectively, were correctly downgraded into lower categories after applying ancillary features (Fig. 5).

**Inter-Reader Agreement**

When the LI-RADS categories were grouped into three categories of LR-5/5V, LR-M, and LR-1/2/3/4, inter-reader agreements were good for both CT and MRI ($\kappa = 0.626$ and 0.601, respectively). Inter-reader agreements were good for arterial hyperenhancement assessments using CT ($\kappa = 0.675$) and moderate MRI ($\kappa = 0.563$), for assessments of washout appearance by using CT and MRI ($\kappa = 0.510$ and 0.532, respectively), as well as for assessments of capsule appearance by using CT ($\kappa = 0.476$). However, inter-reader agreement was fair for assessments of capsule appearance by using MRI ($\kappa = 0.326$).

**DISCUSSION**

The results of this multicenter, off-site reader study showed that gadoxetate-enhanced MRI has higher sensitivity and accuracy than CT in discriminating between malignant and benign lesions in patients with chronic liver disease, as shown in previous studies (19, 22). However, other recent studies have shown that the diagnostic accuracy was comparable between gadoxetate-enhanced MRI and CT in the diagnosis of hepatic malignancies (23).

Table 3. Diagnostic Performance of CT and Gadoxetate-Enhanced MRI in Differentiation of HCC from Benign Lesions and OM from HCC

<table>
<thead>
<tr>
<th></th>
<th>LR-5/5V for HCC vs. Benign</th>
<th>LR-5 for HCC vs. Benign</th>
<th>LR-M for HCC vs. OM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CT</td>
<td>MRI</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TP/TP/FP/FN</td>
<td>Sensitivity, %</td>
<td>Specificity, %</td>
</tr>
<tr>
<td></td>
<td>239/219/17/23</td>
<td>52.4</td>
<td>92.8</td>
</tr>
<tr>
<td></td>
<td>285/213/23</td>
<td>62.5</td>
<td>90.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.003</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>49.8</td>
<td>49.8</td>
<td>46.6</td>
</tr>
<tr>
<td></td>
<td>59.7</td>
<td>66.4</td>
<td>70.4</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>50.4</td>
<td>95.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57.8</td>
<td>95.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.167</td>
<td>0.999</td>
</tr>
</tbody>
</table>

These are pooled results from four reviewers. Unless otherwise specified, data represent numbers of cases. These are results from subgroup analysis in differentiation of HCC from benign lesions (with OM excluded from analysis) and OM from HCC (with benign lesions excluded from analysis).

Table 4. Diagnostic Performance of Gadoxetate-Enhanced MRI Using PP Alone for Washout Appearance

<table>
<thead>
<tr>
<th></th>
<th>(a) CT</th>
<th>(b) MRI Using PP and TRP</th>
<th>(c) MRI Using PP Alone</th>
<th>$P$</th>
<th>$P^*$ (a vs. (c))</th>
<th>$P^*$ (b vs. (c))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TP/TP/FP/FN</td>
<td>Sensitivity, %</td>
<td>Specificity, %</td>
<td>Accuracy, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>406/197/39/282</td>
<td>59.0</td>
<td>83.5</td>
<td>65.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>498/198/38/190</td>
<td>72.4</td>
<td>83.9</td>
<td>75.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>466/204/32/222</td>
<td>67.7</td>
<td>86.4</td>
<td>72.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 0.001</td>
<td>0.119</td>
<td>&lt; 0.001</td>
<td>0.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 0.001</td>
<td>0.374</td>
<td>&lt; 0.001</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These are pooled results from four reviewers. These results show diagnostic performance of gadoxetate-enhanced MRI when using portal phase alone for evaluating washout appearance, in comparison with CT or MRI using both portal and transitional phases in diagnosis of hepatic malignancies. $^*p$ values from post-hoc tests. PP = portal phase, TRP = transitional phase.
In contrast, the accuracy for differentiating HCCs from other malignancies was comparable between gadoxetate-enhanced MRI and CT, as suggested by previous studies in which both gadoxetate-enhanced MRI and CT showed limitations in accurately discriminating HCC from other malignancies, especially combined HCC-CCA (27-30).

Furthermore, similar to a previous report (27), we noted that 3.4% (8/232) and 4.7% (11/232) of non-HCC malignancies showed tumors in the vein and were categorized as LR-5V for CT and MRI, respectively, which further decreased the specificity of LR-5/5V for HCC diagnosis. However, the latest version (v2018) of LI-RADS has revised LR-5V (definitely HCC with tumor in vein [TIV]) to LR-TIV, so the presence of the TIV would no longer decrease the specificity for HCC diagnosis (31).

We observed that the application of the ancillary features modified the LI-RADS categories in 12% (111/924) of cases, similar to the results of two recent studies (15.3% and 18.1%) (12, 25). In our study, no significant improvement was observed in the overall LI-RADS categorization with addition of the ancillary features. However, 10 benign lesions initially categorized as LR-5/5V were correctly downgraded (Fig. 5, Supplementary Table 7). Two recent studies using extracellular contrast media reported conflicting results on the impact of upgrading LR-3 to LR-4 based on ancillary features on the diagnostic performance (12, 32). Both studies showed that the application of ancillary features increased the sensitivity of LR-4/5/5V for HCC, while the specificity was shown to be preserved in one study (12) and decreased in another (32).

Our study also showed that multicenter, off-site evaluations could achieve comparable inter-reader agreement for LI-RADS categorization with both MRI and CT, as seen in previous single-center studies (33, 34). A recent study involving 113 readers from eight different sites reported a better inter-reader agreement (35). However, in that study, preselected screen-captured images were evaluated, in contrast to our study, in which whole series of images were provided to the reviewers.

This study had limitations. First, our study may have biases due to its retrospective design. There were differences in several baseline characteristics between patients with malignant and benign lesions in this study, which might have resulted in further biases. In addition, inclusion of only pathologically-confirmed lesions may have been a source of selection bias. For example, most of the benign lesions included were likely difficult to diagnose by enhanced MRI and CT for the diagnosis and characterization of hepatic lesions on the basis of LI-RADS (23-25). In comparison with our study, these results were not obtained in a multicenter study involving off-site evaluation by multiple readers. Furthermore, we included a substantial number of cases involving non-HCC malignancies because our purpose was to compare the performance of dynamic CT and MRI in differential diagnosis. Therefore, we believe that the multicenter and multireader study design as well as the addition of a sufficient number of non-HCC malignancies would be strengths of our study. The higher sensitivity of gadoxetate-enhanced MRI in differentiating hepatic lesions could be attributed to its superiority in demonstrating imaging features favoring malignancies. In line with a recently published study (26), we found that arterial hyperenhancement as well as washout appearance were more frequently observed in HCCs on gadoxetate-enhanced MRI than on CT.

Fig. 5. Benign lesion initially categorized as LR-5 but correctly downgraded by applying ancillary features. 52-year-old HBV carrier underwent gadoxetate-enhanced MRI after hepatic nodule was found on ultrasonography. 1.4-cm nodule in left liver shows arterial hyperenhancement (arrow) and washout appearance on portal phase (arrow). Two reviewers considered arterial enhancement as rim-like and categorized nodule as LR-M. Other reviewers initially categorized nodule as LR-5. Nodule shows signal drop from opposed-phase to in-phase of T1-weighted gradient-recalled echo sequence (arrowhead), indicating presence of intralobesional iron deposits. Note that nodule shows isointensity on T2-weighted image. These features are uncommon finding in progressed HCCs. After applying these ancillary features, reviewers downgraded their categories to LR-4. This nodule was confirmed as angiomyolipoma after hepatic resection.
imaging alone, which might have affected the diagnostic accuracy—most likely by underestimating the specificity. However, we believe that the reference standard for evaluating the performance during differential diagnosis of hepatic tumors must be histopathological findings, since some non-HCC malignancies may show the typical imaging features of HCCs (8, 23, 28) and these HCC mimickers can be treated by locoregional or systemic treatment without pathologic confirmation according to the current guidelines (28). Second, we used the earlier version of LI-RADS (v2014) because it was the latest version available at the time of our image analysis. However, except for the change from LR-5V to LR-TIV, only minor changes (i.e., mainly more elaborate definitions of features) have been made in the diagnostic table (12). Furthermore, we obtained similar results when only LR-5 was considered positive for HCC diagnosis. In addition, since the LI-RADS v2014 did not specifically define high-risk patients and advised to follow the other guidelines, we also included non-cirrhotic patients with chronic hepatitis C according to European Association for the Study of the Liver (2). However, the latest version (v2018) of LI-RADS does not consider chronic hepatitis C patients at risk. Thus, our loose inclusion criteria might have affected the results, although the proportion of such patients was small (i.e., about 5%). Third, we evaluated washout appearance using two methods; by using portal phase alone and also by using both portal and transitional phases. Using the transitional phase for evaluating washout appearance might have led to an overestimation of sensitivity of gadoxetate-enhanced MRI. However, we also obtained similar results in the supplementary analysis performed using the portal phase alone for washout evaluation.

In conclusion, our multicenter off-site evaluation study showed that gadoxetate-enhanced MRI was superior to CT in differentiating hepatic malignancies from benign lesions by using the LI-RADS v2014 in patients with chronic liver disease. Both modalities had comparable capabilities for distinguishing HCC and other malignancies. The ancillary features on gadoxetate-enhanced MRI may help avoid misdiagnosis of benign lesions as definitely malignant lesions.

Supplementary Materials

The Data Supplement is available with this article at https://doi.org/10.3348/kjr.2019.0363.

Conflicts of Interest

The funder Bayer Korea Ltd. had no role in the study design; the collection, analysis, and interpretation of data; the writing of the report; or the decision to submit the manuscript for publication.

Acknowledgments

The authors would like to thank Sungwon Kim, MD, Jun Tae Kim, and Kyunghwa Han, PhD, from the Department of Radiology, Yonsei University College of Medicine, for their assistance in data collection and preparation (S.K. and J.K.) and statistical analysis (K.H.).

ORCID iDs

Myeong-Jin Kim
https://orcid.org/0000-0001-7949-5402
Chansik An
https://orcid.org/0000-0002-0484-6658

REFERENCES

Intraindividual Multicenter Comparison of EOB-MRI and CT for Diagnosing HCC


34. Barth BK, Donati OF, Fischer MA, Ulbrich EJ, Karlo CA, Becker