



Published in final edited form as:

Abdom Radiol (NY). 2018 January ; 43(1): 169–178. doi:10.1007/s00261-017-1261-x.

Imaging features of hepatocellular carcinoma compared to intrahepatic cholangiocarcinoma and combined tumor on MRI using liver imaging and data system (LI-RADS) version 2014

Natally Horvat^{1,2}, Ines Nikolovski¹, Niamh Long¹, Scott Gerst¹, Jian Zheng³, Linda Ma Pak³, Amber Simpson³, Junting Zheng⁴, Marinela Capanu⁴, William R. Jarnagin³, Lorenzo Mannelli¹, and Richard Kinh Gian Do¹

¹Department of Radiology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

²Department of Radiology, Hospital Sírio-Libanês, Sao Paulo, Brazil

³Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA

⁴Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Abstract

Purpose: To evaluate the prevalence of major and ancillary imaging features from liver imaging reporting and data systems (LI-RADS) version 2014 and their interreader agreement when comparing hepatocellular carcinoma (HCC) to intrahepatic cholangiocarcinoma (ICC) and combined tumor (cHCC-CC).

Methods: The Institutional Review Board approved this HIPAA-compliant retrospective study and waived the requirement for patients' informed consent. Patients with resected HCC ($n = 51$), ICC ($n = 40$), and cHCC-CC ($n = 11$) and available pre-operative contrast-enhanced MRI were included from 2000 to 2015. Imaging features and final LI-RADS category were evaluated by four radiologists. Imaging features were compared by Fisher's exact test and interreader agreements were assessed by κ statistics.

Results: None of the features were unique to either HCC or non-HCC. Imaging features that were significantly more common among HCC compared to ICC and cHCC-CC included washout (76%–78% vs. 10%–35%, $p < 0.001$), capsule (55%–71% vs. 16%–49%, $p < 0.05$), and intralesional fat (27%–52% vs. 2%–12%, $p < 0.002$). Features that were more common among ICC and cHCC-CC included peripheral arterial phase hyperenhancement (40%–64% vs. 10%–14%, $p < 0.001$) and progressive central enhancement (65%–82% vs. 14%–25%, $p < 0.001$). The

Correspondence to: Richard Kinh Gian Do; dok@mskcc.org.

Author contributions Study concepts and study design, N. H., R. K. G. D.; literature search, N. H., R. K. G. D.; image review, I. N., N. L., S. G., L. M.; clinical information review, J. Z., L. M. P., A. S., W. R. J., N. H.; statistical analysis, J. Z., M. C.; manuscript drafting and edition, N. H., R. K. G. D., I. N., J. Z., J. Z.; approval of final version of submitted manuscript, all authors.

Compliance with ethical standards

Conflict of interest The authors have declared that no competing interests exist.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

interreader agreement was moderate for each of these imaging features ($\kappa = 0.41\text{--}0.55$). Moderate agreement was also achieved in the assignment of LR-M ($\kappa = 0.53$), with an overall sensitivity and specificity for non-HCC malignancy of 86.3% and 78.4%, respectively.

Conclusion: HCC and non-HCC show significant differences in the prevalence of imaging features defined by LI-RADS, and are identified by radiologists with moderate interreader agreement. Using LI-RADS, radiologists also achieved moderate interreader agreement in the assignment of the LR-M category.

Keywords

Liver neoplasms; Hepatocellular carcinoma; Magnetic resonance imaging; Cholangiocarcinoma; Intrahepatic cholangiocarcinoma

Liver cancer is a growing health care challenge with more than 850,000 new cases detected annually worldwide [1]. It is the second leading cause of cancer-related death globally [2, 3] and the sixth highest cause of cancer-related death in the United States [3]. Hepatocellular carcinoma (HCC) is the most common primary hepatic malignancy, accounting for approximately 80% of cases [1–3]. Common risk factors for HCC include chronic hepatitis B or C, alcohol abuse, and metabolic syndrome [4–6]. Intrahepatic mass-forming cholangiocarcinoma (ICC) and combined hepatocellular cholangiocarcinoma (cHCC-CC) may arise in patients with risk factors for HCC, but treatment options and prognosis are different [7–11]. It is therefore important to distinguish between HCC, ICC, and cHCC-CC radiologically, but it can be challenging due to their overlapping imaging findings [12–14]. As the diagnosis of HCC is currently based primarily on imaging findings [2, 15], accurate and reproducible radiological diagnosis is crucial for appropriate management.

The liver imaging reporting and data system (LI-RADS) is a comprehensive algorithm for evaluating and reporting liver observations on CT and MRI in patients at risk for HCC. In addition to scoring the likelihood of HCC, LI-RADS also includes a separate LR-M category for non-HCC malignancies. LR-M is selected when an observation in the liver is deemed probably malignant, but not specific for HCC. Thus, when LR-M is assigned to an observation, a tissue diagnosis is recommended before treatment [16]. The assignment of LR-M is based primarily on the recognition of ancillary features favoring non-HCC malignancies over HCC, such as peripheral arterial phase hyperenhancement (APHE), portal venous and delayed phase central enhancement, peripheral washout, and liver surface retraction [17].

The assignment of LR-M can be challenging in practice as LI-RADS requires radiologists to identify and integrate multiple major and ancillary imaging features for HCC and non-HCC malignancies. Prior studies have investigated the interreader agreement among radiologists when using LI-RADS, but have focused primarily on patients without non-HCC malignancies and on major imaging features of HCC [18–20]. The purpose of this study was to evaluate the prevalence of major and ancillary imaging features from LI-RADS v2014 and their interreader agreement when comparing HCC to ICC and cHCC-CC.

Methods

Study population

This retrospective study received approval from our Institutional Review Board, and was granted a waiver of informed consent. Patients who underwent resection of HCC, ICC, or cHCC-CC between January 2000 and June 2015 were identified from a prospectively maintained surgical database. The inclusion criterion was based on the availability of contrast-enhanced MRI within 6 months of surgery. Patients were excluded if they had multiple tumors, had pre-surgical therapy, or had poor image quality preoperatively.

Based on our selection criteria, we identified 182 patients with HCC, ICC, or cHCC-CC with MRI before surgery. We excluded 34 patients because of embolization prior to resection ($n = 24$), the presence of more than one lesion ($n = 6$), and poor MR imaging quality ($n = 4$). The remaining group included 51 patients with non-HCC malignancies (40 with ICC and 11 with cHCC-CC) and 97 patients with HCC. Among the patients with HCC, we individually matched the 51 non-HCC patients to 51 patients with HCC by tumor size (≤ 5 vs. >5 cm) and the remaining 46 patients were excluded.

The final study population consisted of 102 patients, 51 with non-HCC malignancies and 51 size-matched HCCs. The median interval between the MRI and surgery was 33 days (range 2–146 days).

MRI

All patients had contrast-enhanced MRI on 1.5T or 3T scanners with standard sequences, including T2-weighted imaging (WI), unenhanced T1-WI, and dynamic contrast-enhanced T1-weighted 3D gradient-recalled echo fat-suppressed imaging. Two patients had suboptimal arterial phase imaging. Thirty-five MRIs included diffusion-weighted imaging (DWI) and 13 MRIs included delayed hepatobiliary phase imaging with gadoxetate disodium (Eovist, Bayer HealthCare, Wayne, NJ).

Clinical data

Patient clinical information and laboratory tests were obtained from retrospective review of the medical record. We documented patient age, gender, etiology of underlying liver disease, as well as tumor markers, including alpha-fetoprotein (AFP), carcinoembryonic antigen (CEA), and carbohydrate antigen 19–9 (CA19–9).

Image review

Two senior board-certificated abdominal radiologists (S.G. and L.M. with 16 and 8 years of experience after fellowship), and two junior board-certificated radiologists (I.N. and N.L. with 1 year each of experience after fellowship) independently reviewed the MR examinations. The radiologists were blinded to the radiology report and pathologic diagnosis, but they were aware that each patient had a diagnosis of HCC, ICC, or cHCC-CC.

Before reviewing the MRIs, the readers were given an 1-hour lecture defining the imaging features of HCC and non-HCC malignancy as outlined in LI-RADS v2014 [17], followed by

1-hour of hands-on instruction with case examples not included in the study sample. In addition, the readers were provided with five separate practice cases for training purposes.

The cases were randomly ordered, and the readers reviewed the complete MRI with all acquired sequences on GE Centricity Picture Archiving and Communication System (General Electric, Milwaukee, WI). For each hepatic observation, the readers measured the maximum transverse diameter of the lesion (on the sequence of their choice, avoiding the arterial phase whenever possible), recorded the imaging features, and assigned the final diagnosis (LR-3, LR-4, LR-5, LR-5 V, or LR-M) based on LI-RADS v2014 [17].

The LI-RADS imaging features were subclassified as (a) major features (MF) of HCC, (b) ancillary features (AF) that may favor malignancy, (c) AF that favor HCC over non-HCC malignancies, and (d) AF that favor non-HCC malignancies over HCC. The LI-RADS descriptor of “threshold growth” was not applicable in this study as only one scan was evaluated.

A consensus for the final LI-RADS category was considered present if at least three of the four readers agreed on a feature or diagnosis. Otherwise, two different hepatobiliary radiologists (R.K.G.D. and N.H. with 8 and 3 years of experience after fellowship) reviewed the remaining cases to reach a consensus.

Pathologic examination

Pathology reports of liver resection were retrospectively reviewed for tumor size, final diagnosis (HCC, ICC, or cHCC-CC), and presence of underlying cirrhosis or fibrosis. The histopathologic assessment to differentiate HCC, ICC, and cHCC-CC was based on having hepatocellular, biliary, or both features, respectively. For cHCC-CC, biliary differentiation was evaluated by mucin positivity or immunoreactivity for cytokeratins, and hepatocellular differentiation was evaluated by in situ hybridization for albumin [7].

Statistical analysis

For each reader and the consensus read, we compared features between patients with HCC and non-HCC in continuous and categorical variables using the Wilcoxon rank-sum test and Fisher’s exact test, respectively. The sensitivity and the specificity were estimated for the diagnosis of HCC versus non-HCC using LI-RADS, along with 95% confidence intervals (95% CI) for each reader and the consensus read. The sensitivity and specificity were also compared using Fisher’s exact test between patients who underwent MRI with extracellular gadolinium contrast agents ($n = 89$) vs. gadoxetate disodium ($n = 13$).

Interreader agreement on tumor diameters was assessed by calculating the intraclass correlation coefficient. Interreader agreement on categorical features was assessed using *kappa* (κ) statistic or weight *kappa* statistic with squared weights (for more than two levels) between two senior radiologists (SR) and between two junior radiologists (JR). Light’s *kappa* was then used to assess interreader agreement across all 4 readers. *Kappa* (κ) values were interpreted as follows: 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement [21].

Given that the purpose of the study was hypothesis generating, multiple testing adjustment was not performed. A test with p value < 0.05 was considered statistically significant. All statistical analyses were performed with software packages SAS 9.4 (SAS Institute Inc., Cary, NC, USA), R version 3.1 (The R Foundation for Statistical Computing).

Results

Clinical characteristics

Patient characteristics are displayed in Table 1. No significant differences in patient age, tumor markers, or tumor size were found between HCC and non-HCC, although a trend towards higher AFP was found for HCC ($p = 0.078$). Significantly more patients with HCC were male compared to those with non-HCC ($p = 0.017$). Risk factors commonly associated with HCC were more often present in patients with HCC than those with non-HCC, such as Hepatitis B and C (22% and 25% vs. 6% and 8%). In 48 (47%) out of 102 patients, no risk factors for HCC was identified clinically, including 36 of the patients with non-HCC ($p < 0.001$). The presence of cirrhosis or fibrosis on pathology was also significantly more common in patients with HCC (35 patients, 69%) compared to those with non-HCC (10 patients, 20%) ($p < 0.001$).

Imaging features and interreader agreement

The percentage of imaging features identified for HCC and non-HCC among all 4 readers as well as their interreader agreement and κ statistics for each imaging feature are shown in Table 2. For each of the four readers, significant differences were found in the prevalence of several imaging features between HCC and non-HCC malignancies. For HCC, imaging features that were seen significantly more frequently included washout (76%–78% vs. 10%–35%, $p < 0.001$), capsule (55%–71% vs. 16%–49%, $p < 0.05$), and intralesional fat (27%–52% vs. 2%–12%, $p < 0.002$). In contrast, several features were more common in non-HCC: peripheral APHE (40%–64% vs. 10%–14%, $p < 0.001$) and progressive central enhancement (65%–82% vs. 14%–25%, $p < 0.001$). The interreader agreement for each of these features were moderate ($\kappa = 0.41$ –0.55).

Mosaic architecture was more common in HCC compared to non-HCC for three out of four readers with frequencies of 37%–65% for HCC, and 0%–33% for non-HCC. The interreader agreement for mosaic architecture was poor in JR compared to moderate in SR ($\kappa = 0.15$ vs. 0.46). For SR, mosaic architecture was significantly more common in HCC (47%–65%) versus non-HCC (14%–20%) ($p < 0.001$). Liver surface retraction was significantly more common in non-HCC versus HCC for three of four readers, with frequencies of 4%–16% in HCC, and 24%–35% for non-HCC ($p < 0.05$). The interreader agreement for liver surface retraction was moderate for all readers ($\kappa = 0.47$). Biliary obstruction out of proportion to the mass was infrequent overall, but significantly different for two of the four readers, seen in HCC in 0%–6% of cases, compared to 6%–16% of non-HCC cases ($p < 0.03$). The lack of significant differences for this imaging finding in the other two readers was possibly due to the low frequency of this imaging finding.

While restricted diffusion achieved almost perfect interreader agreement ($\kappa = 0.97$), agreement for target appearance on DWI or marked restricted diffusion were only slight and fair, respectively ($\kappa = 0.14$ and 0.37). Target appearance on DWI was significantly different between HCC and non-HCC for one reader alone. Given the small number of exams ($n = 13$) performed with a delayed hepatobiliary phase, the LI-RADS features related to the use of a hepatobiliary contrast agent were not evaluated.

Among the other imaging features, most had only fair interreader agreement, including APHE ($\kappa = 0.35$), tumor in vein ($\kappa = 0.33$), fat sparing ($\kappa = 0.30$), blood products ($\kappa = 0.21$), mild to moderate T2 hyperintensity ($\kappa = 0.27$), nodule-in-nodule architecture ($\kappa = 0.38$), and a few had no interreader agreement: corona enhancement ($\kappa = -0.00$), and distinctive rim ($\kappa = -0.02$).

The interreader agreement in the assignment of a final LI-RADS category was fair for JR ($\kappa = 0.36$) and moderate for SR ($\kappa = 0.45$), and fair overall for all readers ($\kappa = 0.39$). When limiting the final LI-RADS category as either LR-M versus any other LI-RADS category, the agreement was moderate for SR ($\kappa = 0.49$) and JR ($\kappa = 0.57$). Figs. 1 and 2 are examples of two cases that lacked consensus regarding LI-RADS score and Fig. 3 shows an example of one case with a consensus LI-RADS score of LR-5.

Miscategorized non-HCC malignancies

Based on the consensus read, 7 non-HCC malignancies were not assigned as LR-M (Table 3). Of these, 4 (57%) were cHCC-CC, with a median size of 3.5 cm (range 2–8 cm) and 3 (43%) of which measured less than 3 cm. The consensus read misclassified one cHCC-CC as LR-5 V.

Sensitivity and specificity of LR-M for non-HCC malignancies

The sensitivity and specificity of LR-M for the diagnosis of a non-HCC in our patient population are summarized in Table 4. The sensitivity ranged from 74.5% to 88.2% across SR and JR, with specificity ranging from 68.6% to 82.4%. Sensitivity for the diagnosis of non-HCC malignancies was higher when radiologists analyzed the 49 tumors that were larger than 5 cm (80.0%–93.3%) compared to the 53 tumors smaller than 5 cm (61.9%–81.0%). For the 19 tumors less than 3 cm, the sensitivities were low among all four readers (25.0%–50.0%), with higher specificities (66.7%–86.7%). There was no significant difference in sensitivity and specificity across all readers (all p value > 0.05) between patients who underwent MRI with extracellular gadolinium contrast agents ($n = 89$) versus gadoxetate disodium ($n = 13$).

Discussion

In our study population, HCC and non-HCC had significant differences in several major and ancillary imaging features, which were recognized with moderate interreader agreement. Washout, capsule, and intralesional fat were more frequently seen in HCC, while peripheral APHE and progressive central enhancement were more common in ICC and cHCC-CC. For a majority of the readers, mosaic architecture and liver surface retraction were also significantly different between the two groups, with mosaic architecture more common in

HCC and liver surface retraction more common in non-HCC. None of the features were unique to either HCC or non-HCC, including intralesional fat and tumor in vein.

Overall, the interreader agreements for the major features APHE and washout in our study were fair ($\kappa = 0.35$) and moderate ($\kappa = 0.50$), respectively, compared with substantial and moderate as reported by Davenport et al. [18], and moderate for both as reported by Barth et al. [19]. The lower agreement for APHE may be explained by the inclusion of non-HCC malignancies in our patient population, where APHE may be limited to the rim and distinguished from APHE as a major feature of HCC. Regarding AF favoring non-HCC malignancies, our moderate interreader agreement in assessing enhancement patterns was consistent with previous report [14] and is comparable to reported agreements for the major features of HCC. We also obtained a moderate interreader agreement for liver surface retraction and biliary obstruction, in contrast to only slight and fair agreement as reported by Fowler et al. [14]. In our patient population, the most common AF that favored non-HCC malignancies were peripheral APHE and delayed progressive enhancement, as noted by other studies [14, 22].

Interreader agreement for the choice of LR-M as a final LI-RADS category was moderate among our readers, with a few non-HCC still categorized as LR4, LR5, or LR-5 V based on consensus. These included seven out of 51 (14%) non-HCC malignancies, three of which were less than 3 cm. All of those patients had at least one MF of HCC, with 3 patients having AF favoring non-HCC tumors. Thus, in these three patients, the AF favoring non-HCC was not considered significant in the final interpretation of the LI-RADS category. Joo et al. [23] also showed miscategorization of ICC as not LR-M in 11.4% to 25.7% of their cases and Potretzke et al. [24] demonstrated the same in 6.5% of cHCC-CC cases. Without using LI-RADS, Fowler et al. [14] demonstrated that 34%–41% of cHCC-CC were misdiagnosed as ICC whereas 24% were misdiagnosed as HCC. These results illustrate the current challenges in distinguishing HCC from cHCC-CC. Furthermore, presence of tumor in vein was not exclusive to HCC, which is consistent with previous studies [23]. This result raises the potential utility for a new “LR-M with tumor in vein” category in order to avoid labeling all lesions with “tumor in vein” as HCC by default.

While our study was not designed to evaluate the sensitivity and specificity of LR-M in an HCC screening population, our results are comparable to Joo et al. [23], who recently evaluated LI-RADS v2014 for ICC on gadoteric acid-enhanced MRI. The sensitivity and specificity for non-HCC in tumors measuring less than 3 cm were low in our patient population, similar to previous results [25]. Given the low frequency of ICC in an actual screening HCC population, the performance of LI-RADS and LR-M, in particular, may change in clinical practice and should be evaluated prospectively.

There were some limitations in our retrospective study. First, 48 (47.1%) patients included in our study population did not have any documented risk factors for HCC, so they were not representative of a LI-RADS population. Furthermore, 58 (56.9%) patients had fibrosis or cirrhosis on pathological analysis and it is known that the imaging patterns of ICC differs between patients with and without cirrhosis [12, 26]. In our study population, many patients also had tumors larger than 5 cm, a size threshold that excludes the possibility of liver

transplantation in most centers. Furthermore, all lesions were either HCC or non-HCC malignancies, which probably increased our LR-M sensitivities. Only a third of the patients had MRI studies with DWI, limiting the evaluation for target appearance which has previously been described in non-HCC malignancies [23, 27]. We also had a very small number of patients who underwent gadoxetate disodium-enhanced MRI, which precluded us from adequately evaluating the utility of imaging features of non-HCC on hepatobiliary phase. Given the retrospective design and the fact that some scans were performed at an outside hospital, slight variations in arterial phase timing may have accounted for lower interobserver agreement for arterial phase hyperenhancement. Finally, our study did not evaluate other non-HCC malignancies, such as liver metastases, although these presumably occur less frequently than ICC and cHCC-CC in a LI-RADS population.

Our study had a number of strengths including the inclusion of a large number of consecutive patients with pathologically proven HCC, ICC, and cHCC-CC after surgical resection, all with contrast-enhanced MRIs. We also had enough patients to compare size-matched HCC and non-HCC malignancies, given the known similar radiographic appearances in these tumors when they are small size. Our study is also the first to specifically evaluate the interreader agreement in the use of LR-M and the first in comparing HCC with both ICC and cHCC-CC, which are the dominant non-HCC malignancies encountered in a LI-RADS patient population.

In conclusion, HCC and non-HCC malignancies showed significant differences in both major and ancillary imaging features as defined by LI-RADS v2014, features that were identified by radiologists with moderate interreader agreement. Peripheral arterial phase hyperenhancement and progressive central enhancement were the most frequent AF in non-HCC malignancies whereas washout, capsule, and intralesional fat were more frequent in HCC. Challenges remain in the diagnosis of ICC and cHCC-CC in smaller tumors because of overlapping appearances with HCC. No single imaging feature was found exclusively in HCC, including tumor in vein, intralesional fat, and mosaic architecture, although they were rare in ICC and cHCC-CC. Further prospective studies in a screening population are needed to overcome the limitations of this study and to provide better estimates of sensitivity and specificity of LI-RADS in the use of LR-M.

Acknowledgments

Funding This work was supported in part by NIH/NCI P30 CA008748 Cancer Center Support Grant.

References

1. Torre LA, Bray F, Siegel RL, et al. (2015) Global cancer statistics, 2012. *CA: A Cancer J Clin* 65(2): 87–108
2. Beyene TJ, Hoek H, Zhang Y, Vos T (2015) Causes of death C. Global, regional, and national age-sex specific all-cause and causespecific mortality for 240 causes of death, 1990–2013: a systematic analysis for the global burden of disease study 2013. *Lancet* 385(9963):117–171 [PubMed: 25530442]
3. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. *CA: A Cancer J Clin* 66(1):7–30
4. Llovet JM, Zucman-Rossi J, Pikarsky E, et al. (2016) Hepatocellular carcinoma. *Nat Rev Dis Prim* 2:16018 [PubMed: 27158749]

5. Bruix J, Boix L, Sala M, Llovet JM (2004) Focus on hepatocellular carcinoma. *Cancer cell* 5(3): 215–219 [PubMed: 15050913]
6. Zheng J, Kuk D, Gonen M, et al. (2016) Actual 10-year survivors after resection of hepatocellular carcinoma. *Ann Surg Oncol* 24:1358–1366 [PubMed: 27921192]
7. Jarnagin WR, Weber S, Tickoo SK, et al. (2002) Combined hepatocellular and cholangiocarcinoma: demographic, clinical, and prognostic factors. *Cancer* 94(7):2040–2046 [PubMed: 11932907]
8. Chantajitr S, Wilasrusmee C, Lertsitichai P, Phromsopha N (2006) Combined hepatocellular and cholangiocarcinoma: clinical features and prognostic study in a Thai population. *J Hepato-biliary-pancreatic Surg* 13(6):537–542
9. Yin X, Zhang BH, Qiu SJ, et al. (2012) Combined hepatocellular carcinoma and cholangiocarcinoma: clinical features, treatment modalities, and prognosis. *Ann Surg Oncol* 19(9): 2869–2876 [PubMed: 22451237]
10. Sapisochin G, Fidelman N, Roberts JP, Yao FY (2011) Mixed hepatocellular cholangiocarcinoma and intrahepatic cholangiocarcinoma in patients undergoing transplantation for hepatocellular carcinoma. *Liver Transplant* 17(8):934–942
11. Valverde A, Bonhomme N, Farges O, et al. (1999) Resection of intrahepatic cholangiocarcinoma: a Western experience. *J Hepatobiliary-pancreatic Surg* 6(2):122–127
12. Kim SA, Lee JM, Lee KB, et al. (2011) Intrahepatic mass-forming cholangiocarcinomas: enhancement patterns at multiphasic CT, with special emphasis on arterial enhancement pattern–correlation with clinicopathologic findings. *Radiology* 260(1):148–157 [PubMed: 21474703]
13. Chung YE, Kim MJ, Park YN, et al. (2009) Varying appearances of cholangiocarcinoma: radiologic-pathologic correlation. *Radiographics* 29(3):683–700 [PubMed: 19448110]
14. Fowler KJ, Sheybani A, Parker RA 3rd, et al. (2013) Combined hepatocellular and cholangiocarcinoma (biphenotypic) tumors: imaging features and diagnostic accuracy of contrast-enhanced CT and MRI. *AJR Am J Roentgenol* 201(2):332–339 [PubMed: 23883213]
15. Bruix J, Sherman M (2011) American association for the study of liver D. Management of hepatocellular carcinoma: an update. *Hepatology* 53(3):1020–1022 [PubMed: 21374666]
16. 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(3):1056–1065 [PubMed: 25041904]
17. Singh AK, Nachiappan AC, Verma HA, et al. (2010) Postoperative imaging in liver transplantation: what radiologists should know. *Radiographics* 30(2):339–351 [PubMed: 20228321]
18. Davenport MS, Khalatbari S, Liu PS, et al. (2014) Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging. *Radiology* 272(1):132–142 [PubMed: 24555636]
19. Barth BK, Donati OF, Fischer MA, et al. (2016) Reliability, validity, and reader acceptance of LI-RADS—an in-depth analysis. *Acad Radiol* 23(9):1145–1153 [PubMed: 27174029]
20. Bashir MR, Huang R, Mayes N, et al. (2015) Concordance of hypervascular liver nodule characterization between the organ procurement and transplant network and liver imaging reporting and data system classifications. *J Magn Reson Imaging* 42(2):305–314 [PubMed: 25371354]
21. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159–174 [PubMed: 843571]
22. Park HJ, Jang KM, Kang TW, et al. (2016) Identification of imaging predictors discriminating different primary liver tumours in patients with chronic liver disease on gadoxetic acid-enhanced MRI: a classification tree analysis. *Eur Radiol* 26(9):3102–3111 [PubMed: 26634931]
23. Joo I, Lee JM, Lee SM, et al. (2016) Diagnostic accuracy of liver imaging reporting and data system (LI-RADS) v2014 for intrahepatic mass-forming cholangiocarcinomas in patients with chronic liver disease on gadoxetic acid-enhanced MRI. *J Magn Reson Imaging* 44(5):1330–1338 [PubMed: 27087012]
24. Potretzke TA, Tan BR, Doyle MB, et al. (2016) Imaging features of biphenotypic primary liver carcinoma (hepatocholangiocarcinoma) and the potential to mimic hepatocellular carcinoma: LI-RADS analysis of CT and MRI features in 61 cases. *AJR Am J Roentgenol* 207(1):25–31 [PubMed: 26866746]

25. Huang B, Wu L, Lu XY, et al. (2016) Small intrahepatic cholangiocarcinoma and hepatocellular carcinoma in cirrhotic livers may share similar enhancement patterns at multiphase dynamic MR imaging. *Radiology* 281(1):150–157 [PubMed: 27077381]
26. Xu J, Igarashi S, Sasaki M, et al. (2012) Intrahepatic cholangiocarcinomas in cirrhosis are hypervascular in comparison with those in normal livers. *Liver Int* 32(7):1156–1164 [PubMed: 22417172]
27. Hwang J, Kim YK, Park MJ, et al. (2012) Differentiating combined hepatocellular and cholangiocarcinoma from mass-forming intrahepatic cholangiocarcinoma using gadoxetic acid-enhanced MRI. *J Magn Reson Imaging* 36(4):881–889 [PubMed: 22730271]

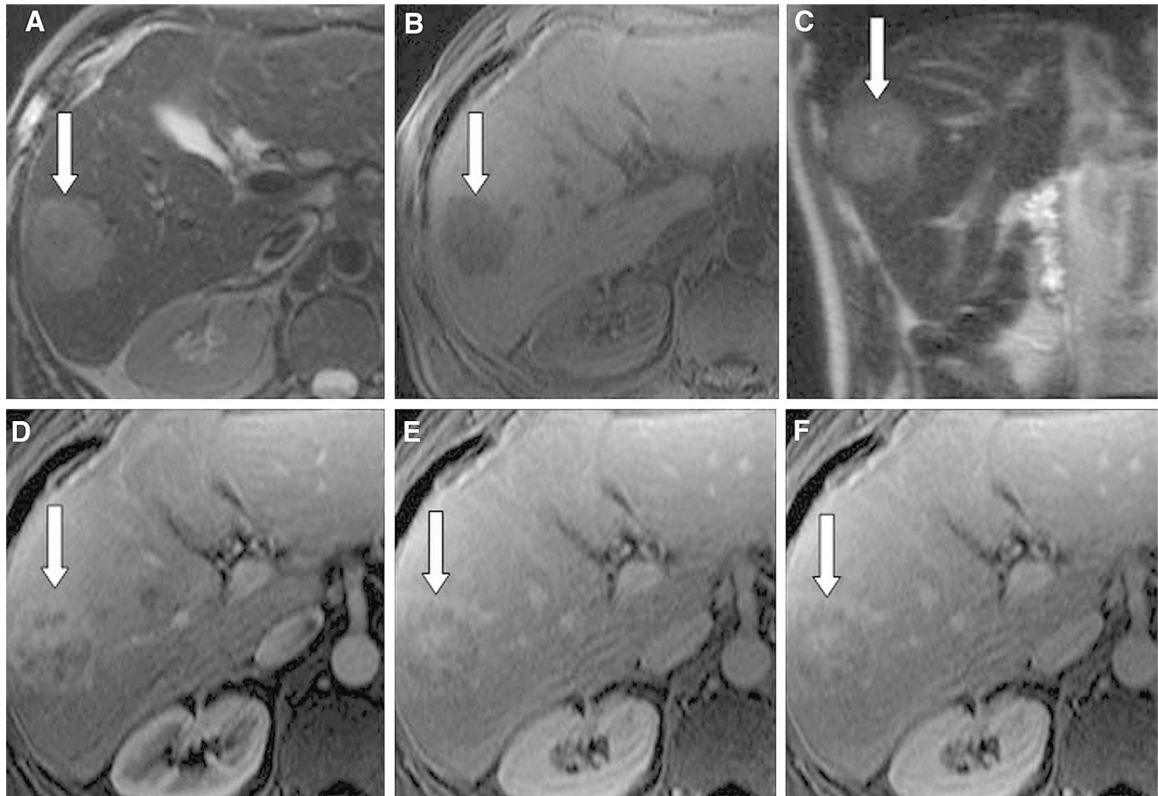


Fig. 1. 54-year-old man with ICC. There was a consensus between readers regarding high signal intensity (SI) on T2- weighted imaging (**A, C**) and low SI on T1-weighted imaging (**B**). However, there was no consensus regarding peripheral arterial phase hyperenhancement (**D**), peripheral washout (**E, F**), and LI-RADS score (2 scores of LR-5 and 2 scores of LR-M) for this liver observation (*arrow*). The mass was resected and found to be ICC.

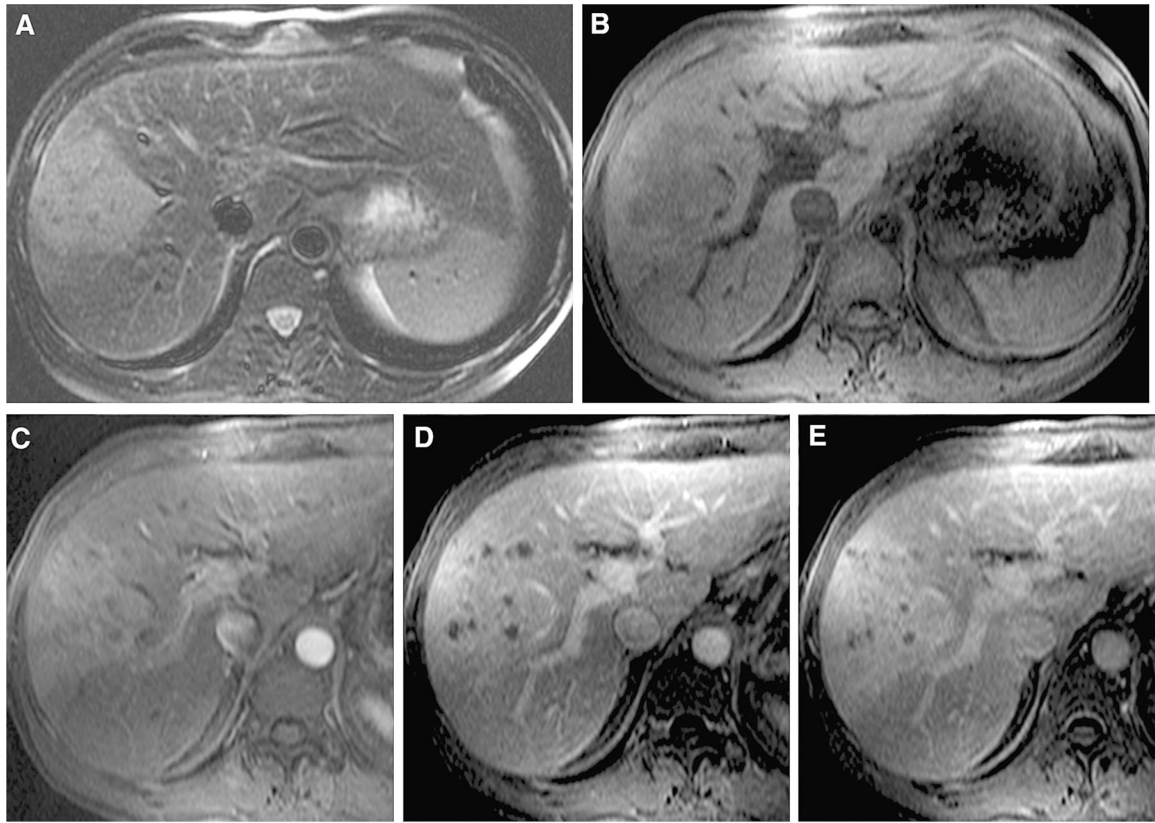


Fig. 2. 50-year-old male with chHCC-CC. There was a consensus achieved among readers regarding moderate high signal intensity on T2-weighted imaging (A) and arterial phase hyperenhancement (C). However, there was no consensus regarding liver surface retraction, presence of progressive enhancement (D, E), tumor in vein, or LI-RADS score (2 scores of LR-M, 1 score as LR-4, and 1 as LR-5 V). The patient underwent partial hepatectomy and the final diagnosis was chHCC-ICC.

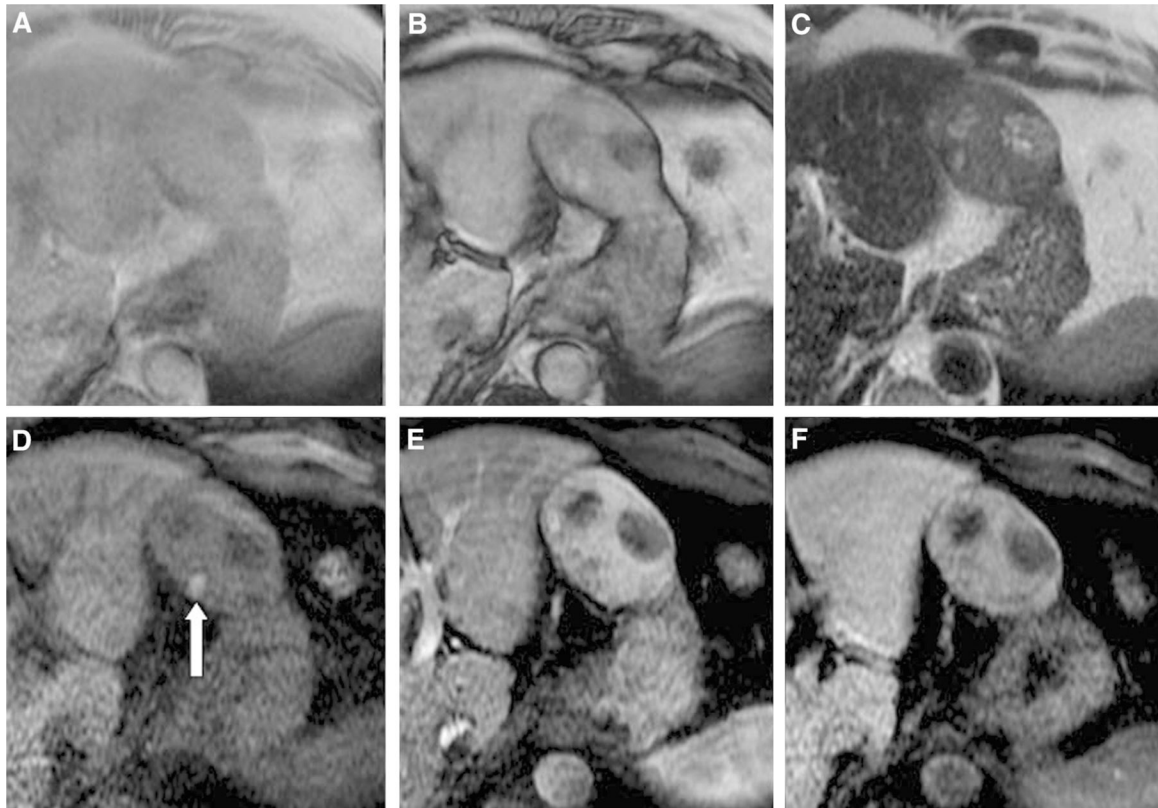


Fig. 3. 69-year-old man with HCC. There was a consensus achieved among readers regarding intralesional fat on T1- weighted in and out of phase imaging (**A**, **B**), moderate hyperintense signal intensity on T2-weighted imaging (**C**), blood products (*arrow*) within the observation on fat saturated T1-weighted imaging (**D**), arterial phase hyperenhancement (**E**), washout (**F**), mosaic architecture (**B–F**), and LI-RADS score (LR-5).

Table 1.

Characteristics of the patients

Characteristic	HCC (51)	Non-HCC (51)	p value*
Age (years)			
Mean (\pm SD)	65.6 (13.7)	63.1 (13.3)	0.295
Sex (M/F)	34/17	21/30	0.017
Chronic liver disease			
Hepatitis B	11 (22%)	3 (6%)	<0.001
Hepatitis C	13 (25%)	4 (8%)	
Hepatitis B and C	1 (2%)	0 (0%)	
Alcoholic hepatitis	3 (6%)	1 (2%)	
Non-alcoholic steatohepatitis	11 (22%)	4 (8%)	
Primary sclerosing cholangitis	0 (0%)	3 (6%)	
No known risk factor for HCC	12 (24%)	36 (71%)	
Tumor markers			
AFP level (ng/mL)			
Mean, median (range)	69.8, 7.2 (1, 798)	2183, 4.7 (0, 35121)	0.22
Elevated (>5 ng/mL)	31 (67%)	18 (47%)	0.078
CA 19-9 (ng/mL)			
Mean, median (range)	8.6, 2.6 (0, 62)	2.7, 2 (0, 11.1)	0.139
Elevated (>5 ng/mL)	6 (27%)	6 (15%)	0.314
CEA (ng/mL)			
Mean, median (range)	53.6, 34 (2, 197)	294.3, 41.5 (0, 6142)	0.556
Elevated (>40 ng/mL)	6 (40%)	19 (53%)	0.541
Pathological data			
Tumor size (cm)			
Mean (\pm SD)	6.2 (4.7)	6.2 (2.7)	0.102
Cirrhosis	13 (25%)	3 (6%)	0.012
Fibrosis	35 (69%)	7 (14%)	<0.001

Table 2.

Imaging features and interreader agreement

Imaging features	SR 1			SR 2			JR 1			JR 2			k SR	k JR	k AR
	HCC	Non-HCC	p	HCC	Non-HCC	p	HCC	Non-HCC	p	HCC	Non-HCC	p			
Arterial phase hyperenhancement (APHE)	46 (90%)	43 (86%)	0.439	43 (84%)	35 (70%)	0.134	47 (92%)	42 (84%)	0.159	50 (98%)	48 (96%)	0.745	0.45	0.29	0.35
Washout	39 (76%)	5 (10%)	<0.001	39 (76%)	13 (25%)	<0.001	40 (78%)	11 (22%)	<0.001	39 (76%)	18 (35%)	<0.001	0.65	0.29	0.50
Capsule	36 (71%)	25 (49%)	0.043	31 (61%)	10 (20%)	<0.001	33 (65%)	10 (20%)	<0.001	28 (55%)	8 (16%)	<0.001	0.43	0.53	0.41
Tumor in vein	5 (10%)	8 (16%)	0.374	16 (32%)	13 (26%)	0.738	3 (6%)	2 (4%)	0.617	1 (2%)	0 (0%)	0.999	0.35	0.37	0.33
Diameter Mean (±SD)	5.9 (4)	5.3 (2.3)	0.357	5.9 (4.2)	5.3 (2.4)	0.385	5.9 (4)	5.3 (2.4)	0.334	5.6 (3.9)	5.1 (2.2)	0.436	0.98	0.92	0.95
LI-RADS AF that favor Malignancy															
Fat sparing	4 (8%)	6 (12%)	0.741	8 (17%)	12 (24%)	0.455	2 (4%)	4 (8%)	0.678	1 (2%)	4 (8%)	0.362	0.38	0.13	0.30
Iron sparing	1 (2%)	2 (4%)	0.999	2 (4%)	4 (8%)	0.678	0 (0%)	1 (2%)	0.999	1 (2%)	2 (4%)	0.999	0.42	-0.02	0.05
Blood products	2 (4%)	0 (0%)	0.237	15 (31%)	6 (12%)	0.027	3 (6%)	1 (2%)	0.352	7 (15%)	1 (2%)	0.028	0.14	0.47	0.21
Mild-moderate T2 hyperintensity	40 (80%)	46 (90%)	0.172	43 (86%)	45 (88%)	0.775	31 (62%)	45 (88%)	0.003	42 (84%)	43 (84%)	0.999	0.25	0.37	0.27
Restricted DWI	38 (75%)	29 (57%)	0.095	38 (75%)	29 (57%)	0.095	38 (75%)	29 (57%)	0.095	38 (75%)	29 (57%)	0.095	1.00	0.94	0.97
LI-RADS AF that favor HCC over non-HCC															
Intralesional fat	25 (52%)	6 (12%)	<0.001	18 (38%)	5 (10%)	0.002	19 (40%)	2 (4%)	<0.001	13 (27%)	1 (2%)	<0.001	0.70	0.53	0.55
Mosaic architecture	33 (65%)	10 (20%)	<0.001	24 (47%)	7 (14%)	<0.001	19 (37%)	0 (0%)	<0.001	20 (39%)	17 (33%)	0.681	0.46	0.15	0.32
Nodule-in-nodule architecture	12 (36%)	3 (27%)	0.722	3 (12%)	0 (0%)	0.999	2 (11%)	0 (0%)	NA	6 (30%)	1 (6%)	0.104	0.36	0.41	0.38
Corona enhancement	1 (2%)	5 (10%)	0.205	4 (8%)	10 (20%)	0.148	1 (2%)	1 (2%)	0.999	9 (18%)	9 (18%)	0.999	-0.09	0.17	-0.00
Distinctive rim	5 (10%)	1 (2%)	0.205	2 (4%)	0 (0%)	0.495	2 (4%)	2 (4%)	0.999	5 (10%)	15 (29%)	0.023	-0.03	0.02	-0.02
LI-RADS AF that favor non-HCC over HCC															
Marked restricted DWI	4 (31%)	14 (64%)	0.086	8 (62%)	18 (82%)	0.243	0 (0%)	4 (18%)	0.274	2 (15%)	1 (5%)	0.541	0.31	0.21	0.14
Target appearance on DWI	1 (8%)	5 (23%)	0.377	1 (8%)	4 (18%)	0.630	1 (8%)	12 (55%)	0.010	1 (8%)	3 (14%)	0.999	0.46	0.22	0.37
Liver surface retraction	8 (16%)	18 (35%)	0.04	8 (16%)	14 (27%)	0.228	5 (10%)	15 (29%)	0.023	2 (4%)	12 (24%)	0.008	0.35	0.58	0.47
Biliary obstruction	0 (0%)	7 (14%)	0.013	0 (0%)	3 (6%)	0.243	1 (2%)	8 (16%)	0.051	3 (6%)	8 (16%)	0.200	0.58	0.34	0.46
Peripheral APHE	7 (14%)	31 (62%)	<0.001	5 (10%)	21 (42%)	<0.001	7 (14%)	32 (64%)	<0.001	6 (12%)	20 (40%)	0.001	0.51	0.44	0.52
Progressive central enhancement	12 (24%)	42 (82%)	<0.001	13 (25%)	34 (67%)	<0.001	7 (14%)	38 (75%)	<0.001	12 (24%)	33 (65%)	<0.001	0.47	0.52	0.51
Peripheral washout	2 (4%)	1 (2%)	0.999	0 (0%)	5 (10%)	0.056	3 (6%)	7 (14%)	0.318	0 (0%)	2 (4%)	0.495	-0.04	-0.04	0.06
Final Diagnosis													0.45	0.36	0.39

Imaging features	SR 1			SR 2			JR 1			JR 2			k SR	k JR	k AR			
	HCC	Non-HCC	p	HCC	Non-HCC	p	HCC	Non-HCC	p	HCC	Non-HCC	p						
LR-3	0 (0%)	0 (0%)	<0.001	0 (0%)	3 (6%)	<0.001	0 (0%)	0 (0%)	<0.001	1 (2%)	0 (0%)							
LR-4	2 (4%)	2 (4%)		3 (6%)	2 (4%)		0 (0%)	1 (2%)		8 (16%)	5 (10%)							
LR-5	30 (59%)	2 (4%)		22 (43%)	5 (10%)		36 (71%)	5 (10%)		32 (63%)	8 (16%)							
LR-5V	3 (6%)	2 (4%)		14 (27%)	2 (4%)		2 (4%)	1 (2%)		1 (2%)	0 (0%)							
LR-M	16 (31%)	45 (88%)		12 (24%)	39 (76%)		13 (25%)	44 (86%)		9 (18%)	38 (75%)							
LR-M vs. LR-non-M															<0.001	0.49	0.57	0.53

AR all readers, AF ancillary features, JR junior reader, SR senior reader, κ kappa statistic

Table 3.

Characteristics of misclassified non-HCC malignancies after consensus read

Case no	Liver disease	Age	Sex	Size	MF of HCC	HCC AF	Non-HCC AF	LI-RADS	Pathological diagnosis
1	HBV	44	M	8	APHE, capsule	None	Progressive enhancement	5	cHCC-CC
2	Alcoholic	54	M	4	APHE	Mosaic architecture, distinctive rim	Progressive enhancement	4	ICC
3	HBV	50	M	8	APHE, tumor in vein	None	None	5 V	cHCC-CC
4	PSC	53	F	2	APHE	Intralesional fat	Markedly restricted diffusion	4	cHCC-CC
5	No	61	F	2.3	APHE, capsule	None	None	5	ICC
6	No	76	F	2.8	APHE, capsule	None	None	5	ICC
7	NASH	74	M	3.5	APHE, washout, capsule	None	None	5	cHCC-CC

AF ancillary features, APHE arterial phase hyperenhancement, HBV hepatitis B virus, MF major features, PSC primary sclerosing cholangitis, NASH non-alcoholic steatohepatitis

Table 4.

Sensitivity and specificity of LR-M according to the size of the observation

LR-M vs. non LR-M	SR 1	SR 2	JR 1	JR 2	Consensus read
All patients (<i>n</i> = 102)					
Sensitivity (95% CI)	0.882 (0.761,0.956)	0.765 (0.625,0.872)	0.863 (0.737,0.943)	0.745 (0.604,0.857)	0.863 (0.737,0.943)
Specificity (95% CI)	0.686 (0.541,0.809)	0.765 (0.625,0.872)	0.745 (0.604,0.857)	0.824 (0.691,0.916)	0.784 (0.647,0.887)
Tumor size >5 cm (<i>n</i> = 49)					
Sensitivity (95% CI)	0.933 (0.779,0.992)	0.867 (0.693,0.962)	0.933 (0.779,0.992)	0.800 (0.614,0.923)	0.933 (0.779,0.992)
Specificity (95% CI)	0.684 (0.434,0.874)	0.684 (0.434,0.874)	0.737 (0.488,0.909)	0.737 (0.488,0.909)	0.842 (0.604,0.966)
Tumor size 5 cm (<i>n</i> = 53)					
Sensitivity (95% CI)	0.810 (0.581,0.946)	0.619 (0.384,0.819)	0.762 (0.528,0.918)	0.667 (0.430,0.854)	0.762 (0.528,0.918)
Specificity (95% CI)	0.688 (0.500,0.839)	0.812 (0.636,0.928)	0.750 (0.566,0.885)	0.875 (0.710,0.965)	0.750 (0.566,0.885)
Tumor size 3.1–5 cm (<i>n</i> = 34)					
Sensitivity (95% CI)	0.882 (0.636,0.985)	0.647 (0.383,0.858)	0.882 (0.636,0.985)	0.765 (0.501,0.932)	0.882 (0.636,0.985)
Specificity (95% CI)	0.588 (0.329,0.816)	0.824 (0.566,0.962)	0.824 (0.566,0.962)	0.882 (0.636,0.985)	0.706 (0.44,0.897)
Tumor size 3 cm (<i>n</i> = 19)					
Sensitivity (95% CI)	0.5 (0.068,0.932)	0.5 (0.068,0.932)	0.25 (0.006,0.806)	0.25 (0.006,0.806)	0.25 (0.006,0.806)
Specificity (95% CI)	0.8 (0.519,0.957)	0.8 (0.519,0.957)	0.667 (0.384,0.882)	0.867 (0.595,0.983)	0.8 (0.519,0.957)

JR junior reader, SR senior reader