POLISH
JOURNAL *of* **RADIOLOGY**

Original paper

How high is the inter-observer reproducibility in the LIRADS reporting system?

Sezgin Sevim, Oğuz Dicle, Naciye S. Gezer, Mustafa M. Barış, Canan Altay, Işıl Başara Akın

Department of Radiology, School of Medicine, Dokuz Eylül University, İnciraltı, İzmir, Turkey

Abstract

Purpose: To investigate the reproducibility of LIRADS v2014 and contribute to its widespread use in clinical practice.

Material and methods: This retrospective, single-centre study was conducted between January 2010 and October 2015. A total of 132 patients who had dynamic magnetic resonance imaging (MRI)/computed tomography (CT) images in the Picture Archiving and Communication Systems (PACS) with liver nodule were included in the study, 37 of whom had histopathology results. Five radiologists who participated in the study, interpreted liver nodules independently on different PACS stations according to the LIRADS reporting system and its main parameters.

Results: We determined that level of inter-observer agreement in the LR-1, LR-5, and LR-5V categories was higher than in the LR-2, LR-3, and LR-4 categories ($\kappa = 0.522$, 0.442, and 0.600 in the LR-1, LR-5, and LR-5V categories, respectively; $\kappa = 0.082$, 0.298, and 0.143 in the LR-2, LR-3, and LR-4 categories, respectively). The parameter that we observed to have the highest level of inter-observer agreement was venous thrombus ($\kappa = 0.600$).

Conclusions: Our study showed that LIRADS achieves an acceptable inter-observer reproducibility in terms of clinical practice although it is insufficient at intermediate risk levels. We think that the prevalence of its use will be further increased with training related to the subject and the assignment of numerical values that express the probability of malignancy for each category and including the ancillary features in the algorithm according to clearer rules.

Key words: HCC, cirrhosis, LIRADS, liver nodule.

Introduction

Hepatocellular carcinoma (HCC) is increasing in frequency all over the world [1]. It is the sixth most commonly occurring cancer in the world and the third greatest cause of cancer mortality [2]. Risk factors have been demonstrated well enough, and early diagnosis is possible with suitable surveillance. Early diagnosis allows curative treatment opportunities.

Recently published diagnostic guidelines utilise the imaging modalities more than ever. Due to biopsy-related complications, including haemorrhage, inadequate sampling in small lesions, and seeding along the biopsy tract, the American Liver Diseases Association (AASLD) recommends biopsy only for lesions greater than 1 cm in diameter and suspicious radiological findings [3-6]. The result is the increasing responsibility of radiologists in HCC diagnosis.

For several years great effort has been devoted to the study of HCC diagnosis and treatment. The decision to be made between diagnosis and treatment, as in all diseases, is vital to the prognosis of the disease. One of the parameters that make these decisions accurate is the correct communication. The communication expressed in this sense should ensure that the diagnostic information could be transferred to the clinician with an effective method. Diagnostic classifications have been developed for this purpose.

Although there is some conflict on the pathognomonic imaging features of HCC, there is no single algorithm accepted in the physician-radiologist axis. The Liver Imaging Reporting and Data System (LIRADS), introduced

Correspondence address:
Dr. Sezgin Sevim, Department of Radiology, School of Medicine, Dokuz Eylül University, 35340, İnciraltı, İzmir, Turkey, e-mail: ssezgin.sevim@gmail.com

by American College of Radiology (ACR) in 2011, aims to fill this gap in the field. However, it has not gained as wide currency as the Breast Imaging Reporting and Data Systems (BIRADS) developed for reporting breast lesions. Nonetheless, several centres around the world have begun to utilise LIRADS as a reporting tool. This trend has been increasing as the reliability rises and the practicability of the method is accepted more.

Studies that investigate the effectiveness of LIRADS in clinical practice are limited [7-10]. One of the questions that we think is important in this regard is: What is the level of alignment between observers in evaluations made with LIRADS? Indication of a high level of reproducibility will lead to widespread use of these classifications.

The purpose of this study was to investigate the reproducibility of LIRADS v2014 and contribute to its widespread use in clinical practice.

Material and methods

This retrospective study was approved by the Institutional Review Board (IRB; 2609-GOA/09-10/2016) and was conducted in accordance with the ethical standards of the Institutional and National Research Committee and with the 1964 Helsinki Declaration and its later amendments.

Patients admitted to our hospital between January 2010 and October 2015 were involved in the study. We considered only those patients who had cirrhosis due to any cause and had a dynamic contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI) examination for further downstream analysis.

A total of 2398 cirrhosis patients with International Classification of Disease code 10 (ICD-10) were extracted using hospital information system. One of the authors of the study reviewed all the cases and eliminated 2077 patients who had no contrast-enhanced dynamic series. Also 151 patients lacking liver lesions and 38 patients with suboptimal examination quality were removed from the study. The most identifiable lesion in different series and sequences was selected. A total of 132 lesions were finally listed with their identification numbers, demographic data, lesion localisation by liver segments, and the examination dates, to be review by the observers.

Lesions smaller than 1 cm in diameter are not recommended for dynamic contrast imaging; therefore, patients under 18 years old and patients with suboptimal image quality were excluded from the study. In addition, the patients received any kind of treatment previously, were excluded from the study.

Five observers with more than five years of radiology experience evaluated the lesions. Images were interpreted independently on different PACS stations. Observers were blind to the clinical and pathological findings.

The presence of threshold growth, arterial contrast enhancement, washout, pseudocapsule appearance and vascular thrombus were evaluated for each lesion. Measurement of the nodules was done in the best visualised sequence and phase concerning the longest diameter. Minor parameters were also taken into consideration in the LIRADS class allocation.

Classification criteria were downloaded from the official ACR website. The LIRADS v2014 reporting algorithm was used in nodule reporting.

All of the MRI examinations were performed in the Department of Radiology of our hospital with the use of Achieva (Philips Medical Systems, Best, The Netherlands) and Intera (Philips Medical Systems, Best, The Netherlands) 1.5-Tesla MRI equipment. Precontrast images included the DWI series, fast spin echo (FSE) fat-suppressed axial T2-weighted, axial T2-weighted images with single-shot FSE technique, and GRE T1-weighted out-phase/in-phase axial images. Post-contrast dynamic images were acquired in late arterial, portal, and late portal phases with breath-hold spoiled-GRE 3D technique in axial and coronal planes. Subtracted images were obtained from dynamic series in order to make a more price interpretation. All the cases have DWI axial images and ADC maps obtained from the same data sets.

MRI acquisition parameters included FOV: 385-415 mm, matrix: 256×256, slice thickness: 8 mm, fat-sat T2-weighted axial images (TSE SPIR, TR: 1500-2350 ms, TE: 70 ms, ETL: 24), T2 star axial images (TSE SSH, TR: 1300-1500 ms, TE: 325 ms, ETL: 144), T1-weighted dual phase axial images (GRE T1 DUAL, TR: 96-138, TE: 4.6/2.3, ETL: 2, FA: 10-15^0), DWI axial images (SE EPI SENSE DWI B = 0 sn/mm^2, B = 5000 sn/mm^2, B = 10,000 sn/mm^2), pre- and post-contrast fat-sat axial and coronal images (Breath-Hold 3D GRE sTHRIVE/TFE-IP/WATS T1, TR: 230-251 ms, TE: 4.6-6.9 ms, ETL: 1, FA: 10-15^0), and subtracted images obtained from these data.

CT images of the patients consisted of arterial and portal phase images acquired with "Philips Brilliance" (16 and 64 slice) equipment using parameters of KVP: 120, mA: 280-400, FOV: 385-410 mm, matrix: 512×512, and pitch: 1-1.5 with 2 mm slice thickness. Precontrast imaging was skipped in order not to increase the patient radiation dose.

Gadolinium-based contrast materials (Dotarem, Gadovist, Multihance, Omniscan, Magnevist, Primovist) and iodine containing contrast agents (Omnipaque, Optiray, Pamiray, Ultravist) were used in MRI and CT examinations, respectively. Contrast material admission was given via peripheral veins at 3-5 cc/s. Late arterial, portal, and venous phase images were taken after 35-40, 60-70, and 180-350 seconds, respectively.

100 ml of iodinated contrast medium (density of 300 mg/ml) for dynamic CT and 10 ml of gadolinium contrast agent (density of 0.1 mmol/ml) for dynamic MR were injected intravenously, followed by normal saline flush (20 ml). Using the bolus-tracking method, the late arterial phase was scanned when the contrast media passed

the portal vein. Late arterial, portal, and delayed hepatic phases were obtained according to recommended technical specifications in the LI-RADS guidelines [11].

Statistical evaluation

Statistical analyses were done by means of SPSS v20.0 (IBM, Chicago, USA) software. The significance of the statistical analysis results was evaluated in 95% CIP-values less than 0.05 were considered as significant.

Maximum and minimum values, arithmetic mean, standard deviation, and frequencies were calculated in descriptive analysis.

Intraclass correlation analysis was performed in order to compare the observer's liver nodule diameter measurements. χ^2 test was performed to compare different proportions.

κ-analysis was used to evaluate the inter-observer reproducibility between major specifications of LIRADS and final LIRADS classification. Cohen's κ for two-fold and Fleiss's κ for more than two inter-observer conformity was calculated. The Landis & Koch scale; 0-0.20 (harmonious compliance), 0.21-0.40 (poor compliance), 0.41-0.60 (moderate compliance), 0.61-0.80 (good compliance), and 0.81-1.00 (perfect fit) was utilised to evaluate κ-values [12].

Results

In total 132 cases who had dynamic contrast CT/MRI examinations were included in the study. Ninety-two (69.79%) and 40 (30.3%) of the cases were men and women, respectively. Mean age was 58.77 ± 11.09 years, and no statistical significance was found between genders according to the age averages ($p = 0.822$). Thirty-seven (28.0%) of the cases were evaluated with CT, and 95 (72.0%) of the study group were evaluated with MRI.

ICC was found (95% CI: 0.756 [0.696-0.811]) for the nodule diameter measurements of all observers.

Fleiss's κ value was used to evaluate inter-observer reproducibility for arterial contrast enhancement, washout, pseudocapsule appearance, threshold growth, and venous thrombus, which were the major parameters in the LIRADS reporting system. Fleiss's κ was calculated for arterial contrast enhancement, washout, pseudocapsule appearance, threshold growth, and thrombus as 0.414, 0.471, 0.494, 0.575, and 0.600, respectively.

Contrast enhancement patterns in 62 of the lesions were concluded after an inter-observer consensus. The number of the nodules in which the observers were in consensus according to washout, pseudocapsule appearance, threshold growth, and venous thrombus were 61, 77, 18, and 109, respectively. Threshold growth assessment was not possible with cases who had no previous examinations. The best agreed parameter between observers was venous thrombus.

When the final LIRADS classes were evaluated by the observers, Fleiss's κ values ranged from 0.442 to 0.600 in LR-5, LR-5V, and LR-1 classes. These results were acceptable for clinical practise. However, it was noted that the inter-observer agreement was lower in the intermediate categories LR-2, LR-3, and LR-4 for the HCC risk grade. Only 10 of all nodules were classified as LR-1 consensus by the observers, while the consensus nodule number of the observers was 12 in the LR-5 class and 6 in the LR-5V class. However, there were no consensus-reported nodules in the LR-2, LR-3, and LR-4 classes among the observers. When all the LIRADS classes were evaluated together, it was found that the inter-observer reproducibility was weaker (κ = 0.392). Only 28 (21.2%) of all nodules had consensus among the observers in the determination of final LIRADS class of liver nodules (Table 1).

When the major parameters and categories of LIRADS were evaluated according to the modality, the level of reproducibility in CT for arterial contrast enhancement was higher than in MRI. A higher compliance score was found on MRI compared to CT in assessing washout, pseudocapsule findings, venous thrombosis, and threshold growth.

Table 1. The analysis evaluation results of liver nodules according to the LIRADS reporting system by observers and inter-observer reproducibility levels in LIRADS categories

LIRADS	Ob-1 n (%)	Ob-2 n (%)	Ob-3 n (%)	Ob-4 n (%)	Ob-5 n (%)	Con n (%)	κ (95% CI)
LR-1	30 (22.73%)	21 (15.91%)	40 (30.30%)	32 (24.24%)	37 (28.03%)	10 (35.71%)	0.522 (0.468-0.575)
LR-2	12 (9.09%)	23 (17.42%)	8 (6.06%)	1 (0.76%)	8 (6.06%)	0 (0%)	0.082 (0.028-0.135)
LR-3	17 (12.88%)	20 (15.15%)	8 (6.06%)	12 (9.09%)	14 (10.61%)	0 (0%)	0.298 (0.244-0.352)
LR-4	15 (11.36%)	11 (8.33%)	13 (9.85%)	24 (18.18%)	9 (6.82%)	0 (0%)	0.143 (0.089-0.197)
LR-5	32 (24.24%)	34 (25.76%)	48 (36.36%)	48 (36.36%)	51 (38.64%)	12 (42.86%)	0.442 (0.388-0.496)
LR-5V	24 (18.18%)	22 (16.67%)	15 (11.36%)	14 (10.61%)	8 (6.06%)	6 (21.43%)	0.600 (0.546-0.654)
LR-M	2 (1.52%)	1 (0.76%)	0 (0%)	1 (0.76%)	5 (3.79%)	0 (0%)	0.268 (0.214-0.322)
Total	132 (100%)	132 (100%)	132 (100%)	132 (100%)	132 (100%)	28 (100%)	0.392 (0.366-0.418)

Ob – observer, κ – Fleiss's kappa, Con – consensus, n – number of nodules, LR – LIRADS

Table 2. Inter-observer reproducibility values according to radiological examination modality in LIRADS classes and major parameters. The kappa value (κ) for categorical variables and the intraclass correlation coefficient (ICC) for the nodule diameter were calculated to determine the level of interobserver compliance

Factor		Fleiss's κ (95% CI)	ICC (95% CI)
Arterial contrast enhancement	MRI	0.387 (0.322-0.451)	–
	CT	0.483 (0.381-0.584)	–
Washout	MRI	0.525 (0.460-0.590)	–
	CT	0.336 (0.233-0.440)	–
Pseudocapsule appearance	MRI	0.487 (0.422-0.552)	–
	CT	0.412 (0.306-0.518)	–
Threshold growth	MRI	0.600 (0.454-0.746)	–
	CT	0.452 (0.175-0.729)	–
Venous thrombus	MRI	0.607 (0.544-0.671)	–
	CT	0.570 (0.468-0.671)	–
LIRADS	MRI	0.368 (0.337-0.399)	–
	CT	0.441 (0.389-0.494)	–
Diameter	MRI	–	0.676 (0.593-0.754)
	CT	–	0.833 (0.737-0.908)

Observational correlation was found to be stronger in CT of the nodule diameter measurement (ICC = 0.833 and 0.676). The level of inter-observer reproducibility in all LIRADS classes was found to be higher in CT (Table 2).

Discussion

The best inter-observer reproducibility was found in LR-5, LR-5V, and LR-1; lesions in these classes can be categorised with little doubt of their benign or malign probability. However, it was found to be low in LR-2, LR-3, and LR-4 lesions, which are more important in the surveillance of cirrhotic patients. Unfortunately, we think that the reason was poor level of reproducibility in all LIRADS classes we observed in our study.

In the study of Davenport *et al.* a poor reproducibility level (κ = 0.35) was found for LIRADS when they compared some similar reporting systems. They found the highest level in LR-5/5V and LR-1 classes (κ = 0.62 and 0.54, respectively). The reproducibility level was low in LR-2, LR-3, and LR-4 classes, similarly to our study (κ = 0.11, 0.26, and 0.28, respectively) [13]. Schellhaas *et al.* [14] observed good reproducibility for all LIRADS categories in the study they conducted using only MRI (κ = 0.609). However, they performed the study with only two observers [14]. In the multicentre prospective study by Basha *et al.* 296 liver lesions were followed-up clinically and radiologically every six months, classified using LIRADS v2014 by six observers. They observed κ-values

for venous thrombus, arterial hyperenhancement, washout appearance, and capsule appearance 0.983, 0.621, 0.546, and 0.549 by CT; 0.991, 0.649, 0.674, and 0.742 by MRI, respectively. They found a good level of reproducibility for all LIRADS classes (κ = 0.895 by CT and 0.926 by MRI) [15].

The LIRADS guide published by ACR recommends radiological follow-up for LR-2 and LR-3 nodules. LR-4 nodules require a different approach, and the guide recommends further clinical and radiological evaluation in multidisciplinary case discussion sessions. Biopsy should only be performed as a result of multidisciplinary discussion [11].

Although there are a limited number of studies that question the malignancy probability in LIRADS classes, in a study by Burke *et al.* [16] it was shown that 30.9% of LR-4 nodules progressed to LR-5 in a median 163-day follow-up [16]. In another study by Tanabe *et al.* [17], nodules were followed 614 days on average, and they found that 4% of LR-3 and 38% of LR-4 nodules progressed to LR-5 [17]. Darnell *et al.* [18] designed to evaluate the sensitivity and specificity of LIRADS, and suggested that 96% of LR-4 nodules, which were smaller than 20 mm in diameter and followed up within a six-month period, were HCC histopathologically. This percentage was 98 for LR-5 nodules [18]. We believe that LR-4 nodules have to be evaluated and managed as LR-5 nodules due to the high risk of malignancy and low level of inter-observer reproducibility for this category of nodules.

In our study, we found a medium level of reproducibility (κ = 0.414-0.600) for the major parameters of the LIRADS reporting system. These were in accord with the results of similar research [13,19].

We found a higher correlation for CT than for MRI (0.483 and 0.387, respectively) in the evaluation of arterial contrast enhancement. However, inter-observer correlation was higher with MR than with CT in the evaluation of washout (0.525 and 0.336). Subtracted images in the MRI series could be responsible for this result.

It is generally expected that MRI is superior to CT in pseudocapsule interpretation [20]. Ehman *et al.* [21] found a higher sensitivity in the observation of pseudocapsule and higher inter-observer correlation with MRI (44% and 0.62) compared to CT (31% and κ = 0.56) [21]. Kappa values for pseudocapsule were very close to each other for MRI and CT in our study (0.487 and 0.412, respectively). However, this may be deceptive because of the low number of cases detectable with CT, which influences the sample size and changes the Cohen's κ values.

In our study, we did not address the effect of the ancillary features because use of the ancillary features is subjective and difficult to reproduce but can affect nodule categorisation. We suggest that depicting the role of ancillary features more explicitly in the LIRADS algorithm could reduce this effect.

In contrast to similar reporting systems, LIRADS has an illustrative atlas with detailed criteria of categories.

This allows a standard instructional curriculum. We believe that such an educational course will improve the accuracy of reporting and the inter-observer reliability level.

LIRADS, which was developed for similar reasons, offers a numeric probability range for the malignancy of the lesions. LIRADS currently does not accept this quantitative approach. Multi-centric statistical studies with large groups using the LIRADS classification will make use this quantitative assessment, which will improve the clinical use of this reporting system.

Our study showed that LIRADS achieves an acceptable inter-observer reproducibility in terms of clinical practice, although the consistency at intermediate-risk levels is insufficient. This result is trustworthy in terms of further use of classification. We think that the prevalence of its use will be further increased with training related to the subject and the assignment of the numerical values, which express the probability of malignancy for each category, and including the ancillary features in the algorithm according to clearer rules.

Limitations

Our study has some limitations. These are the retrospective nature, the relatively small sample size, and the single-centre design. In some multicentric studies higher reproducibility levels for LI-RADS were found [15,22,23]. Also, all MRI examinations were performed with an extracellular gadolinium contrast agent, and we did not evaluate the liver-specific contrast media.

In our department, LIRADS has not been routinely used by the observers; therefore, the observers needed a special training session prior to the study to become familiar with the use of LIRADS. Davenport *et al.* [13] showed that interobserver agreement for all LIRADS classes and its main parameters was better in experienced readers [13].

Conclusions

In our study, there were small differences in MR scanning parameters and in the rate of injection of contrast media. This may reflect selection biases and differences between the imaging protocols. However, all imaging examinations satisfied the recommended technical specifications in the LI-RADS system. Also, the trends observed in this study may be considered as more generalisable based on the heterogeneous study population.

In our study, the observers evaluated the cases blind to clinical and laboratory data. Having these data could improve the precision of decision and increase the inter-observer reproducibility.

Conflict of interest

The authors report no conflict of interest.

References

1. Gomes MA, Priolli DG, Tralhao JG, et al. Hepatocellular carcinoma: epidemiology, biology, diagnosis, and therapies. Rev Assoc Med Bras 2013; 59: 514-524.

2. Ferlay J, Shin H, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010; 127: 2893-2917.

3. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatol 2011; 53: 1020-1022.

4. European Association for the Study of the Liver, European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2012; 943: 908-943.

5. Kokudo N, Hasegawa K, Akahane M, et al. Evidence-based Clinical Practice Guidelines for Hepatocellular Carcinoma: The Japan Society of Hepatology 2013 update (3rd JSH-HCC Guidelines). Hepatol Res 2015; 45: 123-127.

6. Ryder SD. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (HCC) in adults. Gut 2003; 52: iii1-iii8.

7. Becker AS, Barth BK, Marquez PH, et al. Increased interreader agreement in diagnosis of hepatocellular carcinoma using an adapted LI-RADS algorithm. Eur J Radiol 2017; 86: 33-40.

8. Hope TA, Fowler KJ, Sirlin CB, et al. Hepatobiliary agents and their role in LI-RADS. Abdom Imaging 2015; 40: 613-625.

9. Flusberg M, Ganeles J, Ekinci T, et al. Impact of a structured report template on the quality of CT and MRI reports for hepatocellular carcinoma diagnosis. J Am Coll Radiol 2017; 14: 1206-1211.

10. Cruite I, Santillan C, Mamidipalli A, et al. Liver imaging reporting and data system: review of ancillary imaging features. Semin Roentgenol 2016; 51: 301-307.

11. LIRADS v2013.1. ACR. Available from: https://www.acr.org/-/media/ACR/Files/RADS/LI-RADS/LIRADSv2013.pdf?la=en

12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-174.

13. Davenport MS, Khalatbari S, Liu PSC, et al. Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging. Radiology 2014; 272: 132-142.

14. Schellhaas B, Hammon M, Strobel D, et al. Interobserver and intermodality agreement of standardized algorithms for non-invasive diagnosis of hepatocellular carcinoma in high-risk patients: CEUS-LI-RADS versus MRI-LI-RADS. Eur Radiol 2018; 28: 4254-4264.

15. Basha MAA, AlAzzazy MZ, Ahmed AF, et al. Does a combined CT and MRI protocol enhance the diagnostic efficacy of LI-RADS in the categorization of hepatic observations? A prospective comparative study. Eur Radiol 2018; 28: 2592-2603.

16. Burke LM, Sofue K, Alagiyawanna M, et al. Natural history of liver imaging reporting and data system category 4 nodules in MRI. Abdom Radiol 2016; 41: 1758-1766.

17. Tanabe M, Kanki A, Wolfson T, et al. Imaging outcomes of liver imagingreporting and data system version 2014 category 2, 3, and 4 observations detected at CT and MR imaging. Radiol 2016; 26: 152-173.

18. Darnell A, Forner A, Rimola J, et al. Liver imaging reporting and data system with MR imaging: evaluation in nodules 20 mm or smaller detected in cirrhosis at screening U. Radiology 2015; 275: 698-707.

19. Barth BK, Donati OF, Fischer MA, et al. Reliability, validity, and reader acceptance of LI-RADS-an in-depth analysis. Acad Radiol 2016; 23: 1145-1153.

20. Ros PR, Taylor HM. Malignant tumors of the liver. In: Gore RM, Levine MS (eds.). Textbook of gastrointestinal radiology. 2nd ed. W.B Saunders, Philadelphia 2000; 1523-1568.

21. Ehman EC, Behr SC, Umetsu SE, et al. Rate of observation and inter-observer agreement for LI-RADS major features at CT and MRI in 184 pathology proven hepatocellular carcinomas. Abdom Radiol 2016; 41: 963-969.

22. Hwang SH, Park S, Han K, et al. Optimal lexicon of gadoxetic acid-enhanced magnetic resonance imaging for the diagnosis of hepatocellular carcinoma modifed from LI-RADS. Abdom Radiol 2019; 44: 3078-3088.

23. Fowler KJ, Tang A, Santillan C, at 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.