

Original Article

Liver Imaging Reporting and Data System Categorization of Benign and Malignant Tumors on Contrast-Enhanced Computed Tomography

Muhammad Wajid¹, Farah Javed², Muhammad Naveed Babur³

¹ Faculty of Allied Health Sciences, The Superior University, Lahore, Pakistan

² Faculty of Allied Health Sciences, The Superior University, Lahore, Pakistan

³ Professor, Faculty of Allied Health Sciences, The Superior University, Lahore, Pakistan

✉ Correspondence: Farah Javed, farahjaved@superior.edu.pk

Received: 12/17/2025

Accepted: 1/21/2026

Published: 1/31/2026

HOW TO CITE

Muhammad Wajid et al. 2026. Liver Imaging Reporting and Data System Categorization of Benign and Malignant Tumors on Contrast-Enhanced Computed Tomography. Journal of Health, Wellness and Community Research. 4(2). DOI: <https://doi.org/10.61919/k4vszd62>

ARTICLE INFORMATION

Author Contributions	Concept: MW, MNB; Design: MW, FJ, MNB; Data Collection: MW, FJ; Analysis: MW, FJ, MNB; Drafting: MW, FJ, MNB.
Ethical Approval	The study was approved by research ethical board of The Superior University, Lahore, Pakistan
Informed Consent	Written informed consent was obtained from all participants
Conflict of Interest	The authors declare no conflict of interest.
Funding	No external funding.
Data Availability	Available from the corresponding author on reasonable request.
AI Declaration	No generative AI was used for conceptualization, analysis, or content generation. AI use was limited to language refinement, and the authors remain fully responsible for the accuracy, originality, validity, and integrity of the manuscript's content.

ABSTRACT

Background: The Liver Imaging Reporting and Data System (LI-RADS) standardizes CT and MRI interpretation of hepatic observations in patients at risk for hepatocellular carcinoma (HCC). LI-RADS categories communicate progressively increasing probability of HCC but are intended for defined at-risk populations rather than unrestricted assessment of liver lesions. **Objective:** To evaluate the distribution of LI-RADS categories and their association with a study-defined final benign or malignant diagnosis on contrast-enhanced CT. **Methods:** A prospective cross-sectional study included 132 adults undergoing contrast-enhanced CT assessment of liver disease at THQ-level facilities in Punjab. CT observations were categorized from LR-1 through LR-5. Arterial enhancement, washout, biliary dilation, portal-vein findings, ascites, and lesion characteristics were recorded. **Results:** LI-RADS distribution was LR-1 in 15/132 cases (11.4%), LR-2 in 27 (20.5%), LR-3 in 34 (25.8%), LR-4 in 23 (17.4%), and LR-5 in 33 (25.0%). Final study classification was benign in 43 cases and malignant in 89. All LR-1 and LR-2 observations were classified as benign; LR-3 included one benign and 33 malignant observations, whereas all LR-4 and LR-5 observations were classified as malignant. The association between LI-RADS category and diagnosis was very strong, $\chi^2(4)=127.58$, $p<0.001$, Cramér's $V=0.983$. Using the study-defined threshold of $LR\geq 3$ as positive yielded sensitivity of 100%, specificity of 97.7%, positive predictive value of 98.9%, negative predictive value of 100%, and accuracy of 99.2%. **Conclusion:** LI-RADS category was strongly associated with the study-defined benign/malignant classification. However, these very high diagnostic estimates should be interpreted cautiously because the independence of the diagnostic reference standard, LI-RADS eligibility of all participants, and conflicting focal-lesion counts remain important methodological limitations. **Keywords:** LI-RADS; Hepatic Lesion; Hepatocellular Carcinoma; Contrast-Enhanced CT; Liver Tumor; Diagnostic Imaging.

INTRODUCTION

Hepatocellular carcinoma is the principal primary malignancy of the liver and most commonly develops in the setting of chronic liver disease, particularly cirrhosis and selected high-risk chronic hepatitis B populations. Contemporary clinical guidance emphasizes surveillance and diagnostic assessment in populations with sufficiently elevated HCC risk because tumor stage at diagnosis strongly influences therapeutic options and prognosis (1). In appropriately selected high-risk patients, characteristic

enhancement patterns on multiphasic CT or MRI can establish a non-invasive diagnosis of HCC without requiring biopsy in every case (1,2).

The Liver Imaging Reporting and Data System was developed to standardize terminology, image-acquisition expectations, interpretation, and reporting of liver observations in populations at risk for HCC (2,3). CT/MRI LI-RADS incorporates major imaging features such as non-rim arterial-phase hyperenhancement, non-peripheral washout, enhancing capsule, observation size, and threshold growth to assign observations to clinically meaningful diagnostic categories (2,3). Standardization is particularly important because interpretation of individual imaging features and final categories may vary among readers, although multicenter and multireader investigations have demonstrated clinically useful reproducibility when the system is applied systematically (4,5).

LI-RADS categories represent graded probabilities rather than a simple benign-versus-malignant dichotomy. LR-1 and LR-2 denote definitely and probably benign observations, respectively; LR-3 represents intermediate probability; LR-4 indicates probable HCC; and LR-5 represents definite HCC when the criteria are met in an appropriate at-risk population (2,3). LR-M is reserved for observations that are probably or definitely malignant but are not specifically diagnostic of HCC. Systematic reviews and meta-analyses confirm a progressive increase in the proportion of HCC across LI-RADS categories, while also demonstrating that LR-3 and LR-4 remain probabilistic rather than definitive categories (6-9). Diagnostic performance therefore depends materially on which category is used as the positive threshold. LR-5 generally provides high specificity for HCC, while broader thresholds can improve sensitivity at the cost of diagnostic specificity (9-14). Comparative meta-analyses have also shown that performance varies according to CT versus MRI, contrast agent, LI-RADS version, and the diagnostic criteria against which LI-RADS is compared (10-13). More recent individual-participant-data analyses further demonstrate that the diagnostic contribution of major-feature combinations and ancillary imaging features is not uniform, emphasizing that the meaning of a LI-RADS category cannot be reduced to the presence of a single imaging sign (15-17).

The present study evaluated the distribution of contrast-enhanced CT observations assigned to LR-1 through LR-5 and examined the association of these categories with a study-defined final benign or malignant diagnosis. A secondary analysis evaluated the performance of the study-defined binary threshold $LR \geq 3$ for differentiating malignant from benign cases.

MATERIALS AND METHODS

Study Design and Setting

A prospective cross-sectional imaging study included 132 adults undergoing contrast-enhanced CT evaluation for known or suspected liver disease. The study was conducted in the Khanpur/Sadiqabad setting, including THQ Hospital Sadiqabad, Punjab, Pakistan. Data collection was conducted from February to October 2025.

Participants were recruited through convenience sampling. Eligibility criteria included diagnosed liver disease, abnormal liver-function testing, abdominal pain, suspected or known hepatic mass, fatty liver disease, HCC, and benign or malignant hepatic tumors. Pregnant patients, individuals with significant renal dysfunction precluding contrast administration, and patients with a known contrast allergy were excluded.

CT Assessment and LI-RADS Categorization

CT examinations were performed on a Canon 16-slice scanner. Hepatic observations were classified into LR-1, LR-2, LR-3, LR-4, or LR-5 according to the CT/MRI LI-RADS categorization framework (2,3). Recorded imaging variables included arterial-phase enhancement, venous washout, delayed washout, liver morphology, common-bile-duct dilation, intrahepatic and extrahepatic biliary dilation, portal-vein status, ascites, and lesion location. A final lesion outcome was recorded as benign or malignant.

The analytical cross-tabulation was reconciled against the reported LI-RADS marginal totals because one category label in the original tabulation was duplicated. The category totals uniquely supported reconstruction of LR-3 as $n=34$ and LR-4 as $n=23$.

Statistical Analysis

Categorical variables were summarized using frequencies and percentages. The association between LI-RADS category and final benign/malignant classification was evaluated using the chi-square test, and effect size was expressed using Cramér's V.

For the study-defined binary analysis, LR-3, LR-4, and LR-5 were collectively treated as positive, whereas LR-1 and LR-2 were treated as negative. Sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and corresponding 95% confidence intervals were calculated against the recorded final benign/malignant classification. This analytical threshold was used specifically for the present dataset and should not be interpreted as equivalent to the standard clinical meaning of the individual LI-RADS categories.

Ethical Approval

Ethical approval was obtained from the Institutional Review Board of Superior University, Lahore, Pakistan.

RESULTS

Clinical and Imaging Characteristics

All 132 participants were recorded as presenting with abdominal pain. Hepatitis C virus-related disease was the most frequently recorded liver-disease association, followed by hepatitis B virus-related disease and fatty liver disease. Biliary dilation, portal-vein abnormalities, ascites, arterial enhancement, and washout characteristics were also recorded.

Table 1. Selected Clinical and Imaging Characteristics

Variable	n (%)
Total sample	132
HCV-related disease	83 (62.9)
HBV-related disease	33 (25.0)
Fatty liver disease	16 (12.1)
Common-bile-duct dilation	74 (56.1)
Intrahepatic biliary dilation	76 (57.6)
Extrahepatic biliary dilation	63 (47.7)
Portal vein normal	48 (36.4)
Portal vein compressed	50 (37.9)
Portal vein thrombosed	34 (25.8)
Ascites	63 (47.7)
Early arterial enhancement	69 (52.3)
Venous washout	72 (54.5)
Delayed washout	56 (42.4)

LI-RADS Distribution

LR-3 was the most frequently assigned category, accounting for 34 observations (25.8%), followed by LR-5 in 33 observations (25.0%), LR-2 in 27 (20.5%), LR-4 in 23 (17.4%), and LR-1 in 15 (11.4%).

Table 2. Distribution of LI-RADS Categories

LI-RADS Category	n	%
LR-1	15	11.4
LR-2	27	20.5
LR-3	34	25.8
LR-4	23	17.4
LR-5	33	25.0
Total	132	100.0

Final Study Classification

Forty-three observations were classified as benign and 89 as malignant. All LR-1 and LR-2 observations were classified as benign. Among the 34 LR-3 observations, one was classified as benign and 33 as malignant. All observations assigned to LR-4 and LR-5 were classified as malignant.

Table 3. LI-RADS Category Versus Study-Defined Final Diagnosis

LI-RADS Category	Benign	Malignant	Total
LR-1	15	0	15

LR-2	27	0	27
LR-3	1	33	34
LR-4	0	23	23
LR-5	0	33	33
Total	43	89	132

LI-RADS category was strongly associated with the study-defined final diagnosis, with $\chi^2(4)=127.58$, $p<0.001$ and Cramér's $V=0.983$.

Study-Defined Binary Diagnostic Performance

When LR-3, LR-4, and LR-5 were collectively treated as positive and LR-1 and LR-2 as negative, there were 89 true-positive, one false-positive, 42 true-negative, and no false-negative classifications.

Table 4. Performance of the Study-Defined $LR \geq 3$ Threshold

Measure	Estimate	95% CI
Sensitivity	100.0%	95.9–100
Specificity	97.7%	87.9–99.6
Positive predictive value	98.9%	94.0–99.8
Negative predictive value	100.0%	91.6–100
Accuracy	99.2%	95.8–99.9

These estimates describe agreement between the study-defined $LR \geq 3$ threshold and the recorded final benign/malignant classification.

DISCUSSION

The present study demonstrated a very strong association between assigned LI-RADS category and the study-defined final benign or malignant classification. All observations categorized as LR-1 or LR-2 were classified as benign, whereas all LR-4 and LR-5 observations were classified as malignant. LR-3 showed a markedly different pattern, with 33 of 34 observations classified as malignant. Under the study-defined binary threshold of $LR \geq 3$, sensitivity was 100%, specificity was 97.7%, and overall accuracy was 99.2%. The overall ordering of the categories is consistent with the intended probabilistic architecture of LI-RADS, in which increasing category number generally corresponds to increasing probability of HCC or malignancy (2,3,6). Systematic reviews have consistently shown low malignancy proportions among LR-1/LR-2 observations, intermediate but heterogeneous probabilities in LR-3, substantially greater probabilities in LR-4, and high specificity for HCC in LR-5 (6-9). The complete separation of LR-1/LR-2 from LR-4/LR-5 in the present dataset therefore follows the expected directional gradient, although the magnitude of separation is stronger than generally observed in external cohorts.

The LR-3 findings require particular caution. Thirty-three of 34 LR-3 observations in this study were classified as malignant, representing a much higher proportion than would ordinarily be expected for an intermediate-probability category. Meta-analytic evidence demonstrates considerable heterogeneity in the proportion of HCC within LR-3, partly because estimates are influenced by population selection, modality, study design, and the reference standard used (6,7). Longitudinal studies of LR-2 and LR-3 observations have also demonstrated that only a subset progress to higher LI-RADS categories during follow-up, reinforcing that LR-3 should not automatically be interpreted as malignant at the time of initial categorization (18).

The study-defined $LR \geq 3$ binary threshold therefore requires a different interpretation from the ordinary clinical use of LI-RADS. LR-3 denotes an indeterminate observation rather than a definitive malignant diagnosis, while LR-4 denotes probable HCC rather than the diagnostic certainty associated with LR-5 (1-3). Meta-analysis of $LR \geq 3$ thresholds has demonstrated that inclusion of lower categories can increase sensitivity but necessarily changes specificity and the clinical meaning of a positive result (14). Similarly, analyses comparing LR-5 with broader LR-4/5 thresholds demonstrate the expected sensitivity-specificity trade-off as the diagnostic threshold is lowered (12,13). Consequently, the 100% sensitivity and 99.2% accuracy observed in the present dataset describe the performance of the study's chosen binary rule against its recorded final classification rather than the external diagnostic accuracy of standard LI-RADS categories.

The validity of any diagnostic-performance estimate also depends critically on the reference standard against which the index test is evaluated. Diagnostic-accuracy reporting standards emphasize clear specification of the index test, positivity threshold, reference standard, and whether interpretation of each was independent (19,20). Methodological quality frameworks similarly identify inappropriate

reference standards and incorporation of the index test into the reference diagnosis as important potential sources of bias (21). In the present study, the procedure used to establish the final benign/malignant classification was not documented in sufficient detail to determine whether it was independent of the CT features used for LI-RADS categorization. If the same enhancement or washout findings contributed substantially to both LI-RADS assignment and the final classification, incorporation bias could increase apparent agreement. The reported sensitivity, specificity, and accuracy should therefore be interpreted primarily as internal concordance measures unless an independent pathological, longitudinal imaging, or multidisciplinary reference standard can be established.

Eligibility of the study population represents a second important consideration. CT/MRI LI-RADS is designed for populations with a sufficiently high pre-test probability of HCC, particularly adults with cirrhosis and selected patients with chronic hepatitis B infection or other specifically defined risk states (1,2). The inclusion criteria in the present study were broader and included abdominal pain, abnormal liver-function tests, fatty liver disease, suspected hepatic masses, and broadly defined liver disease. These characteristics do not independently establish eligibility for application of the LI-RADS diagnostic algorithm. Recent investigation of LI-RADS outside conventional cirrhotic populations, including patients with non-cirrhotic steatotic liver disease, illustrates that diagnostic performance and applicability can differ when the underlying population departs from the standard LI-RADS target group (22). Verification of the underlying risk status of each included participant would therefore strengthen interpretation of the present findings.

Reproducibility of category assignment is another relevant methodological consideration. Multicenter and international multireader studies have demonstrated that LI-RADS can achieve useful agreement but that both individual imaging features and final categories remain subject to reader variability (4,5). Reporting the number and experience of interpreting radiologists, whether assessments were performed independently, and whether disagreement was resolved through consensus would therefore improve reproducibility in future work using this dataset.

The study also contains an internal inconsistency regarding focal-lesion prevalence. One recorded table indicated that 66 of 132 participants had a focal lesion, whereas lesion location, LI-RADS category, and final lesion classification were recorded for all 132 cases. Because LI-RADS categorization is observation based, clarification of this discrepancy is necessary to determine whether all 132 records represent discrete hepatic observations or whether some variables were coded at the patient rather than lesion level.

Despite these limitations, the study demonstrates a strong ordered relationship between assigned LI-RADS category and final benign/malignant classification. It also highlights an important methodological distinction between category-based risk stratification and converting a multilevel diagnostic framework into a binary threshold. The findings are most defensible as a descriptive and concordance analysis unless the eligibility of the study population and independence of the final diagnostic reference standard can be established.

CONCLUSION

LI-RADS categories were strongly associated with the study-defined benign or malignant classification in this cohort. LR-1 and LR-2 observations were exclusively classified as benign, whereas LR-4 and LR-5 observations were exclusively classified as malignant. The study-defined $LR \geq 3$ threshold produced very high sensitivity, specificity, predictive values, and overall accuracy. These estimates should not be interpreted as external diagnostic-accuracy estimates for standard LI-RADS until the underlying LI-RADS eligibility of participants, the independence and definition of the final diagnostic reference standard, and the conflicting focal-lesion counts are clarified.

LIMITATIONS

The study has several important limitations. The reference standard used to classify observations as benign or malignant was not defined in sufficient detail to determine its independence from the CT findings used for LI-RADS categorization. The study population may also have included participants outside conventional CT/MRI LI-RADS eligibility criteria. An internal inconsistency existed between the reported number of participants with focal lesions and the universal assignment of lesion location, LI-RADS category, and final lesion diagnosis. CT acquisition details and image-reading procedures were incompletely characterized, limiting assessment of technical and reader-related reproducibility. Finally,

no LR-M or LR-TIV observations were represented, limiting evaluation across the full LI-RADS diagnostic spectrum (2,3).

REFERENCES

1. Singal AG, Llovet JM, Yarchoan M, Mehta N, Heimbach JK, Dawson LA, et al. AASLD Practice Guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. *Hepatology*. 2023;78(6):1922-1965. doi:10.1097/HEP.0000000000000466.
2. Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, et al. Liver Imaging Reporting and Data System (LI-RADS) Version 2018: imaging of hepatocellular carcinoma in at-risk patients. *Radiology*. 2018;289(3):816-830. doi:10.1148/radiol.2018181494.
3. Santillan C, Chernyak V, Sirlin C. LI-RADS categories: concepts, definitions, and criteria. *Abdom Radiol (NY)*. 2018;43(1):101-110. doi:10.1007/s00261-017-1334-x.
4. Hong CW, Chernyak V, Choi JY, Lee S, Potu C, Delgado T, et al. A multicenter assessment of interreader reliability of LI-RADS Version 2018 for MRI and CT. *Radiology*. 2023;307(5). doi:10.1148/radiol.222855.
5. Fowler KJ, Tang A, Santillan C, Bhargavan-Chatfield M, Heiken J, Jha RC, et al. Interreader reliability of LI-RADS Version 2014 algorithm and imaging features for diagnosis of hepatocellular carcinoma: a large international multireader study. *Radiology*. 2018;286(1):173-185. doi:10.1148/radiol.2017170376.
6. Lee S, Kim YY, Shin J, Son WJ, Roh YH, Choi JY, et al. Percentages of hepatocellular carcinoma in LI-RADS categories with CT and MRI: a systematic review and meta-analysis. *Radiology*. 2023;307(1). doi:10.1148/radiol.220646.
7. van der Pol CB, Lim CS, Sirlin CB, McGrath TA, Salameh JP, Bashir MR, et al. Accuracy of the Liver Imaging Reporting and Data System in computed tomography and magnetic resonance image analysis of hepatocellular carcinoma or overall malignancy: a systematic review. *Gastroenterology*. 2019;156(4):976-986. doi:10.1053/j.gastro.2018.11.020.
8. Lee S, Kim SS, Roh YH, Choi JY, Park MS, Kim MJ. Diagnostic performance of CT/MRI Liver Imaging Reporting and Data System v2017 for hepatocellular carcinoma: a systematic review and meta-analysis. *Liver Int*. 2020;40(6):1488-1497. doi:10.1111/liv.14424.
9. Lee S, Kim YY, Shin J, Hwang SH, Roh YH, Chung YE, et al. CT and MRI Liver Imaging Reporting and Data System Version 2018 for hepatocellular carcinoma: a systematic review with meta-analysis. *J Am Coll Radiol*. 2020;17(10):1199-1206. doi:10.1016/j.jacr.2020.06.005.
10. Kim YY, Lee S, Shin J, Son WJ, Roh YH, Hwang JA, et al. Diagnostic performance of CT versus MRI Liver Imaging Reporting and Data System category 5 for hepatocellular carcinoma: a systematic review and meta-analysis of comparative studies. *Eur Radiol*. 2022;32(10):6723-6729. doi:10.1007/s00330-022-08985-z.
11. Shin J, Lee S, Yoon JK, Roh YH. Diagnostic performance of the 2018 EASL vs. LI-RADS for hepatocellular carcinoma using CT and MRI: a systematic review and meta-analysis of comparative studies. *J Magn Reson Imaging*. 2023;58(6):1942-1950. doi:10.1002/jmri.28716.
12. Lee S, Kim YY, Shin J, Roh YH, Choi JY, Chernyak V, et al. Liver Imaging Reporting and Data System Version 2018 category 5 for diagnosing hepatocellular carcinoma: an updated meta-analysis. *Eur Radiol*. 2024;34(3):1502-1514. doi:10.1007/s00330-023-10134-z.
13. Lee S, Kim YY, Shin J, Shin H, Sirlin CB, Chernyak V. Performance of LI-RADS category 5 vs combined categories 4 and 5: a systemic review and meta-analysis. *Eur Radiol*. 2024;34(11):7025-7040. doi:10.1007/s00330-024-10813-5.
14. Liang Y, Xu F, Guo Y, Lai L, Jiang X, Wei X, et al. Diagnostic performance of LI-RADS for MRI and CT detection of HCC: a systematic review and diagnostic meta-analysis. *Eur J Radiol*. 2021;134:109404. doi:10.1016/j.ejrad.2020.109404.
15. Goins SM, Jiang H, van der Pol CB, Salameh JP, Lam E, Adamo RG, et al. Individual participant data meta-analysis of LR-5 in LI-RADS Version 2018 versus revised LI-RADS for hepatocellular carcinoma diagnosis. *Radiology*. 2023;309(3). doi:10.1148/radiol.231656.
16. Adamo RG, van der Pol CB, Alabousi M, Lam E, Salameh JP, Abedrabbo N, et al. Diagnostic performance of CT/MRI LI-RADS Version 2018 major feature combinations: individual participant data meta-analysis. *Radiology*. 2025;315(3). doi:10.1148/radiol.243450.
17. Abedrabbo N, Lam ET, McInnes MDF, Dawit H, Kadi D, van der Pol CB, et al. Impact of LI-RADS CT and MRI ancillary features on diagnostic performance: an individual participant data meta-analysis. *Radiology*. 2025;316(1). doi:10.1148/radiol.242278.

18. Ranathunga D, Osman H, Islam N, McInnes MDF, Munir J, van der Pol CB, et al. Progression rates of LR-2 and LR-3 observations on MRI to higher LI-RADS categories in patients at high risk of hepatocellular carcinoma: a retrospective study. *AJR Am J Roentgenol.* 2022;218(3):462-470. doi:10.2214/AJR.21.26376.
19. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ.* 2015;351. doi:10.1136/bmj.h5527.
20. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open.* 2016;6(11). doi:10.1136/bmjopen-2016-012799.
21. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009.
22. Cao J, Shon A, Yoon L, Kamaya A, Tse JR. Diagnostic performance of CT/MRI LI-RADS v2018 in non-cirrhotic steatotic liver disease. *Eur Radiol.* 2024;34(12):7622-7631. doi:10.1007/s00330-024-10846-w.