Impact of Reference Standard on CT/MRI and CEUS LI-RADS diagnosis of hepatocellular carcinoma: a meta-analysis

Summary statement

In CT/MRI and CEUS LI-RADS studies, pathology-based reference standards were four times more common than clinical reference standards with observations confirmed to be hepatocellular carcinoma more than twice as likely.

Authors :

Christian B. van der Pol, MD

Matthew D. F. McInnes, MD, PhD

Jean-Paul Salameh, MSc

Victoria Chernyak, MD, MS

An Tang, MD, MSc

Mustafa R. Bashir, MD

Introduction

Varied reference standards are used in Liver Imaging Reporting and Data System (LI-RADS) diagnostic accuracy studies (1). The effect of reference standard on the diagnostic performance of each LI-RADS category is unclear. The purpose of this study was to perform a meta-analysis to estimate the impact of *pathology-based* reference standards versus *clinical reference standards* that exclude pathology on the percentage of hepatocellular carcinoma (HCC) confirmed for each LI-RADS category.

Materials and Methods

This HIPAA compliant study was approved by The Ottawa Hospital Research Ethics Board and registered on PROSPERO: CRD42020164486 (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=164486). The need for informed consent was waived due to retrospective design. The data from CT, MRI, and contrast-enhanced US studies collected by the LI-RADS individual patient data (IPD) group were reported per PRISMA-DTA (2).

Data provided by the IPD group were used with clustering accounted for as previously described to evaluate the impact of reference standard on the percentage of HCC confirmed for each LI-RADS category (3). The reference standard was classified as either 1) pathology-based if established using explant, surgical resection, or needle biopsy, or 2) a clinical reference standard if established without pathology, including using follow-up imaging. The pooled percentage of HCC was determined by fitting a random-effects model with 95% CIs, and with variability quantified using the I^2 statistic,

using R (R Foundation for Statistical Computing) (4). I^2 >50% marked substantial variability with p<0.05 defining significance (5).

Results

A total of 5 882 observations were included from 32 studies including 4 003 (68%) HCC observations. 4 730 (80%) observations were diagnosed using pathology, and 1 152 (20%) using a clinical reference standard. The percentage of HCC was higher for observations confirmed using a pathology-based versus clinical reference standard (odds ratio, 2.6; 95% CI: 2.1, 3.2; p<0.001). We found no evidence of a difference in the percentage of HCC within each LI-RADS category for pathology-based versus clinical reference standards (p>0.05), **Table 1**. Substantial variability was observed with I^2 ranging from 35–99%.

Discussion

In LI-RADS diagnostic accuracy studies, pathology-based reference standards were approximately four times more common than clinical reference standards. We also found that observations confirmed as HCC were more than twice as likely to have a pathology-based reference standard than a clinical reference standard. Requiring pathology confirmation may introduce selection bias by capturing a larger proportion of observations confirmed as HCC. This is especially relevant for the LR-2 “probably benign” and LR-3 “intermediate probability of malignancy” categories, which are not recommended for pathology confirmation (6).

While the point estimate percentages of HCC were higher when using a pathology-based reference standard than a clinical reference standard that excluded pathology for LR-2 to LR-4 observations, the 95% CIs overlapped between the two categories. Despite relatively large sample sizes, we suspect the wide CIs are the result of substantial variability across pooled studies. While increasing sample size normally improves statistical power and narrows CIs, adding data from studies that increase the variability of a random-effects model can result in less precise estimates with wider CIs (5). The I^2 values were >50% for all observations confirmed using a clinical reference standard except LR-2, which had a low point estimate for percentage of HCC (3%; 95% CI: 0%, 17%; I^2: 35%).

Future studies evaluating LI-RADS may reduce selection bias and generate measures of diagnostic accuracy closer to the truth by including reference standards that are not pathology-based, particularly for observations seldom confirmed on pathology such as LR-2, LR-3, and LR-5. The ideal non-pathology-based reference standard remains uncertain but may include a combination of follow-up imaging, imaging on other modalities, and treatment response (3). Reporting a study's population as a surveillance vs surgical cohort would also clarify potential selection bias and generalizability.

References

1. van der Pol CB, Lim CS, Sirlin CB, 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):976-86.

2. Frank RA, Bossuyt PM, McInnes MDF. Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy: The PRISMA-DTA Statement. Radiology. 2018;289(2):313-4.

3. van der Pol CB, McInnes MDF, Salameh JP, et al. CT/MRI and CEUS LI-RADS Major Features Association with Hepatocellular Carcinoma: Individual Patient Data Meta-analysis. Radiology. 2021:211244.

4. R Core Team (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; Available from: https://www.R-project.org/.

5. Ryan R. Cochrane Consumers and Communication Review Group. Heterogeneity and subgroup analyses in Cochrane Consumers and Communication Group reviews: planning the analysis at protocol stage. 2016 [cited 2021 June 2]; Available from: http://cccrg.cochrane.org.

6. American College of Radiology. Liver Reporting & Data System. 2020 [cited 2020 October 30]; Available from: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS.

Tables

Table 1. Percentage of hepatocellular carcinoma (HCC) confirmed per Liver Imaging Reporting and Data System (LI-RADS) category based on type of reference standard. CRS = clinical reference standard. I^2 is a measure of variability.

Reference standard type by LI-RADS category[1]	Observations, n	HCC (95%CI), %	I^2, %
LR-1 (definitely benign)			
Pathology-based	95	0 (0 – 1)	99.0
CRS[2]	109	0	-
LR-2 (probably benign)			
Pathology-based	222	13.2 (5.8, 27.1)	64.9
CRS	100	2.6 (0.4, 16.9)	34.5
LR-3 (intermediate malignancy probability)			
Pathology-based	576	50.7 (32.9, 68.3)	89.2
CRS	251	16.8 (5.8, 39.8)	85.1
LR-4 (probably HCC)			
Pathology-based	723	84.1 (74.4, 90.6)	83.9
CRS	299	72.1 (41.6, 90.3)	91.9
LR-5 (definitely HCC)			
Pathology-based	2351	96.8 (92.9, 98.6)	92.6
CRS	361	98.5 (91.7, 99.7)	76.8
All categories			
Pathology-based	4730	83.1 (71.2, 90.)	98.7
CRS	1152	55.3 (31.7, 76.6)	97.6

[1]Insufficient data were available to analyze the LR-NC (not categorizable), LR-M (probably or definitely malignant but not HCC specific), and LR-TIV (definite tumor in vein) categories.

[2]No HCC were diagnosed in the LR-1 category using a clinical reference standard therefore this could not be included in the model. A single HCC was reported using the pathology-based reference standard.