Journal of Hepatocellular Carcinoma

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Enhancement Pattern Mapping for Early Detection of Hepatocellular Carcinoma in Patients with Cirrhosis

Newsha Nikzad[1–3,*], David Thomas Fuentes[4,*], Millicent Roach[2], Tasadduk Chowdhury[2], Matthew Cagley[2], Mohamed Badawy[4], Ahmed Elkhesen[5], Manal Hassan[6], Khaled M Elsayes[7], Laura Beretta[8], Eugene Jon Koay[2], Prasun Kumar Jalal[1]

[1]Department of Medicine and Surgery, Baylor College of Medicine, Houston, TX, USA; [2]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [3]Department of Internal Medicine, The University of Chicago Medical Center, Chicago, IL, USA; [4]Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [5]Department of Internal Medicine, Texas Tech University Health Sciences Center, Lubbock, TX, USA; [6]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [7]Department of Abdominal Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [8]Department of Molecular and Cellular Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

*These authors contributed equally to this work

Correspondence: Newsha Nikzad; Prasun Kumar Jalal, Email nnikzad@bcm.edu; jalal@bcm.edu

Background and Aims: Limited methods exist to accurately characterize the risk of malignant progression of liver lesions. Enhancement pattern mapping (EPM) measures voxel-based root mean square deviation (RMSD) of parenchyma and the contrast-to-noise (CNR) ratio enhances in malignant lesions. This study investigates the utilization of EPM to differentiate between HCC versus cirrhotic parenchyma with and without benign lesions.

Methods: Patients with cirrhosis undergoing MRI surveillance were studied prospectively. Cases (n=48) were defined as patients with LI-RADS 3 and 4 lesions who developed HCC during surveillance. Controls (n=99) were patients with and without LI-RADS 3 and 4 lesions who did not develop HCC. Manual and automated EPM signals of liver parenchyma between cases and controls were quantitatively validated on an independent patient set using cross validation with manual methods avoiding parenchyma with artifacts or blood vessels.

Results: With manual EPM, RMSD of 0.37 was identified as a cutoff for distinguishing lesions that progress to HCC from background parenchyma with and without lesions on pre-diagnostic scans (median time interval 6.8 months) with an area under the curve (AUC) of 0.83 (CI: 0.73–0.94) and a sensitivity, specificity, and accuracy of 0.65, 0.97, and 0.89, respectively. At the time of diagnostic scans, a sensitivity, specificity, and accuracy of 0.79, 0.93, and 0.88 were achieved with manual EPM with an AUC of 0.89 (CI: 0.82–0.96). EPM RMSD signals of background parenchyma that did not progress to HCC in cases and controls were similar (case EPM: 0.22 ± 0.08, control EPM: 0.22 ± 0.09, p=0.8). Automated EPM produced similar quantitative results and performance.

Conclusion: With manual EPM, a cutoff of 0.37 identifies quantifiable differences between HCC cases and controls approximately six months prior to diagnosis of HCC with an accuracy of 89%.

Plain Language Summary: Current surveillance and diagnostic methods in hepatocellular carcinoma are suboptimal. Enhancement pattern mapping is an imaging technique that quantifies lesion signals and may be useful in diagnostic and surveillance methods. Enhancement pattern mapping describes quantifiable differences between malignant and benign liver tissue on contrast-enhanced MRI. It amplifies lesion signal and distinguishes malignancy in a surveillance population. The novel imaging technique was investigated at single institution and analyzed lesions compared to cirrhotic parenchyma. Future efforts will include further risk stratification across LI-RADS group categories. The results provide evidence that enhancement pattern mapping uses available imaging data to distinguish hepatocellular carcinoma from non-cancerous parenchyma with and without benign lesions on scans six months prior to diagnosis with standard MRI. The technique introduces a prospective modality to improve diagnostic accuracy and early detection with the goal of improving clinical outcomes.

Received: 15 November 2023
Accepted: 7 March 2024
Published: 20 March 2024

Keywords: LI-RADS, MRI, radiomics, artificial intelligence, liver cancer, machine learning

Introduction

The rapidly rising incidence of hepatocellular carcinoma (HCC) can be attributed to several risk factors, including metabolic syndrome, alcohol, viral hepatitis related to HBV or HCV, and other genetic and environmental etiologies. Morbidity and mortality related to HCC remain major challenges for the healthcare system throughout the world. Guidelines recommend surveillance for patients with cirrhosis for early detection of HCC to improve clinical outcomes. In patients with cirrhosis, the American Association for the Study of Liver Diseases (AASLD) recommends surveillance with abdominal ultrasound (US) every 6 months, with or without serum α-fetoprotein (AFP).[1] Similarly, the European Association for the Study of the Liver (EASL) recommends surveillance with abdominal US every 6 months and notes the suboptimal cost-effectiveness of biomarkers such as AFP.[2] US is an affordable, safe, and accessible imaging method; however, Tzartzeva et al found that the sensitivity of US alone or with AFP for early-stage HCC is only 47% and 63%, respectively.[3,4] A major reason for detection of HCC at a late stage is suboptimal surveillance.[5–7]

Suboptimal performance of US for small lesions motivates the use of contrast-enhanced magnetic resonance imaging (MRI),[8] which detects smaller malignant lesions with more sensitivity than US (84% versus 27%, respectively, according to Kim et al).[9] Although MRI is an expensive imaging method, the higher sensitivity of semi-annual surveillance using MRI in a high-risk population may be more cost-effective than using US in HCC surveillance.[9] A limitation of MRI includes the frequent detection of vascular benign lesions, such as arterio-venous shunts, leading to unnecessary surveillance. On the other hand, MRI surveillance can also result in under-categorization of small or early malignant lesions, leading to delayed diagnoses, and under-staging 24.9% of HCC cases.[10–12] Non-contrast MRI sequences are further limited by decreased sensitivity for HCC.

There is a need for new minimally invasive tools that better risk stratify patients under surveillance and detect HCC at earlier stages with higher sensitivity and specificity, given the limitations of US and MRI.[1] While most current surveillance methods utilize clinical, demographic, and blood-based biomarkers, diagnostic methods utilize imaging, specifically computed tomography (CT) and MRI.[10,13] The Liver Imaging Reporting and Data System (LI-RADS) represents an attempt to classify liver nodules for probability of malignancy with CT or MRI using a standardized method to minimize the discrepancies among radiologists.[14] LI-RADs use tumor size and arterial-phase hyperenhancement as defining features for risk stratification, while major features of an enhancing capsule, washout, and threshold growth can further increase the confidence in a malignant or benign diagnosis. However, there is significant heterogeneity within the LI-RADS groups, especially for LR-3 and LR-4 lesions.[1,14] Limited diagnostic performance is especially evident in these categories as 38% of LR-3 and 74% of LR-4 lesions are HCC.[15] The wide range of probability of HCC in LR-3 and LR-4 observations indicates a current limitation in diagnostic accuracy and risk stratification. Furthermore, robust recommendations for diagnostic follow-up imaging or even biopsy in patients with these indeterminate lesions do not exist.[1] Due to the integral role of vascular endothelial growth factor (VEGF) signaling in the oncogenesis of HCC, tumor enhancement is a key consideration in LI-RADs, but no risk stratification method quantifies enhancement on CT or MRI.[16] Tumor enhancement, therefore, poses a promising avenue for the optimization of imaging-based diagnosis of HCC in surveillance populations.

Recent studies have investigated machine learning and radiomic approaches for HCC detection, such as enhancement pattern mapping (EPM).[10,11] EPM is a novel voxel-based signal analysis technique that quantifies the difference in enhancement over time of a given voxel in the liver compared to either a patient-specific or population-based normal liver model, providing a measurement and visualization of how different the signal is over an entire volume of interest. EPM expands on the available set of imaging features, and it provides an interpretable value that is based on angiogenesis and tumor perfusion that are fundamental to HCC pathophysiology improving its diagnostic performance. Previous medical literature has indicated that EPM algorithm improves the contrast-to-noise (CNR) ratio enhancement for lesion detection in hepatobiliary malignancy.[12] Therefore, we embarked on this study to test the hypothesis that EPM

can differentiate between HCC and cirrhotic parenchyma with and without lesions on pre-diagnostic and diagnostic MRI scans, with a future view of applications of EPM for early detection of HCC in patients undergoing surveillance.

Materials/Patients and Methods

Patient Selection and Characteristics

With approval from the Baylor College of Medicine Institutional Review Board (H-47711 and H-45208) and MD Anderson Cancer Center (PA14-0646), all consecutive patients presenting with cirrhosis at the Hepatology Clinic between 2012 and 2023 at a single tertiary care hospital (Baylor St. Luke's Medical Center) were prospectively followed. The institutional practice for HCC surveillance is with contrast-enhanced MRI and surveillance is performed every six months. Patients with cirrhosis were included in the study if they had at least two consecutive contrast-enhanced MRIs, utilizing liver protocol, for HCC surveillance. Imaging protocol details are further described in Supplementary Data 1. Patients with cirrhosis presenting with HCC in the initial scan were excluded.

Cases were defined as patients with a LI-RADS 3 or 4 lesion identified in a pre-diagnostic scan that progressed to HCC in the subsequent diagnostic scan. HCC diagnosis was defined as meeting diagnostic criteria through liver biopsy or LI-RADS 5 categorization on MRI imaging. Criteria *Controls* were defined as patients with and without liver lesions that did not progress to HCC. Controls were age and sex-matched with cases to minimize confounding variables in patient characteristics. Patients were excluded after initial review if follow-up imaging was unavailable (1/166 patients, 0.6%) or poor-quality such as significant motion artifact (18/166 patients, 10.8%). A final total of 48 cases and 99 controls were identified, and their medical records were reviewed. Variables extracted via chart review included demographics (sex, age at first scan, race, ethnicity), tumor marker (AFP), concurrent risk factors (etiology of cirrhosis, type II diabetes mellitus, and body mass index). A consort diagram is provided in Figure 1.

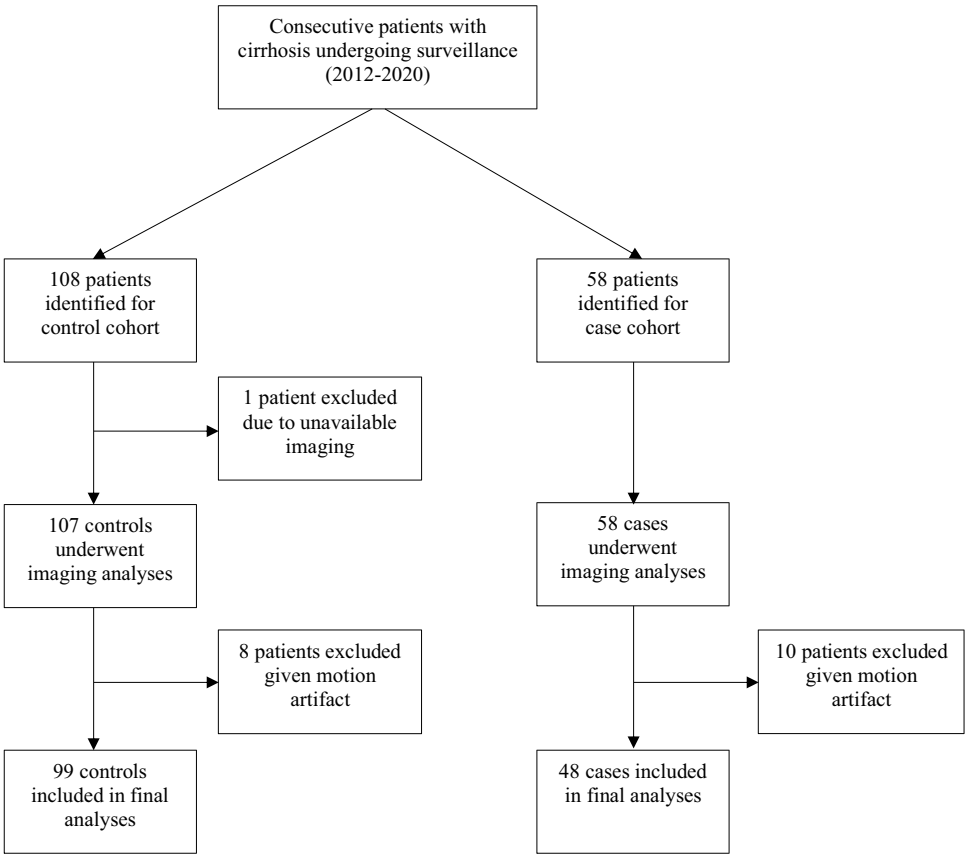

Figure 1 Patient selection process.

Statistical analysis to describe patient demographics was performed using SPSS Statistics (Version 26, SPSS Ltd, Chicago, IL). Data were tested for normality and homogeneity of variance using a Shapiro–Wilk test. Based on this outcome ($P<0.05$) and after visual examination of each variable's histogram and QQ plot, data were reported as mean (SD) for normally distributed variables and as median (IQR 25th to 75th percentiles) for asymmetrical distributed data. A Student's *t*-test was used to compare unpaired symmetrical continuous variables, and the non-parametric Mann–Whitney test was used for unpaired asymmetrical continuous variables. The Chi-square test or Fisher's exact test was used to compare binary variables. A $P<0.05$ and a confidence interval (CI) of 95% were considered statistically significant.

Region of Interest (ROI) Placement

Previous studies have demonstrated decreasing segmentation accuracy as a function of lesion size.[17] Thus, one region of interest was manually placed on each observed HCC lesion to avoid potential confounding factors from lesion segmentation inaccuracies. Similarly, to avoid any potential confounding factors from auto segmentation inaccuracies ROIs in background liver parenchyma were manually and randomly sampled. The normal liver ROIs were selected by visually analyzing the parenchyma of the liver to avoid medium-to-large blood vessels, cysts, and bile ducts. The software application ITK-SNAP[12] was used to perform segmentation of ROIs on the arterial phase scan for all cases. A total of 8 ROIs were selected per image slice, with 3 slices per case, giving a total of 24 ROIs per case. The diameter of each ROI was selected as 6 millimeters. ROIs of lesions were delineated on the arterial phase of the contrast-enhanced scan. Arterial phase hyperenhancement of the lesion, venous and delayed phases washout, in addition to the lesion's size and growth pattern were the criteria used to assign category codes based on LI-RADS version 2018 guidelines.[14]

Additional analysis evaluating the sensitivity of the EPM to the ROI placement was completed (see Supplementary Data 1). The EPM threshold analysis for classifying malignant and non-malignant tissue based on an automated liver ROI method was also conducted. Similar results and interpretation were seen when comparing the manual and automated methods. This data overall represents two ends of the spectrum for EPM analyses: fully automated background liver selection versus a manual ROI placement approach. The automatically generated liver mask has a mean segmentation accuracy of 0.92; however, it occasionally was seen to include non-liver tissue or exclude tissue. Segmentation inaccuracies are expected to bias the EPM thresholds, thus the manual ROI placement approach was chosen to minimize the impact of potential segmentation inaccuracies as a true representation of the EPM cutoff (see Table A1 in Supplementary Material).

Enhancement Pattern Mapping

A three-dimensional, voxel-based method called the EPM algorithm was used for quantitative image analysis. Previous implementation of the EPM algorithm[11] in multi-phase CT was modified for multi-phase MRI data. Briefly, the generalized enhancement pattern, such as the change in intensity values over the period of multi-phase MRI due to uptake and washout of contrast materials, of the liver was acquired from the registered multi-phase MRI scans. The normal liver enhancement curve was obtained by fitting the MRI intensity values over time within user-selected ROIs, as described in the previous section, sampled uniformly across normal liver parenchyma from the given patient. Second, the root-mean-square deviation (RMSD) for each voxel was computed by taking the average across all time points of the squares of the differences between the generalized normal liver intensity and the voxel intensity and then taking the square root of the average. Finally, the calculated RMSD values of all voxels were mapped to the original MRI coordinates.

The normal liver enhancement curve was obtained by fitting the MRI intensity values sampled from the normal liver ROIs over the period of contrast phases by a piece-wise smooth function, where each piece was a second-order polynomial. The EPM algorithm was implemented numerically using MATLAB (MathWorks, Inc.). Contrast to noise (CNR) measurements within the EPM image as well as the original multiphase MRI data were calculated as the average intensity value of the lesion minus the average intensity of healthy tissue divided by the standard deviation of the intensity of the health tissue. Here, the intensity average and standard deviation were calculated over the ROI within the lesion and cirrhotic background, as described in the previous section. A Wilcoxon rank-sum test is used to evaluate the

statistical significance of the difference in the EPM RMSD measurements between cases and controls. RMSD differences between LI-RADS categories of the cases were also evaluated.

ROC analysis was applied to study the EPM RMSD threshold to discriminate cases and control. The optimal cut point for the ROC analysis was defined as the point with the closest Euclidean distance to the perfect classifier (sensitivity=-specificity=1). To validate the cut point, the variability of the optimal EPM cut point across 5-folds of the case and control dataset was evaluated. For each fold, the cut point was obtained from the training data independent from the validation hold-out fold. On average, within 5-fold cross validation, the data is split into 80% training data (n=79 controls, n=38 cases) and 20% validation data (n=20 controls, n=10 cases). The sensitivity, specificity, and accuracy were aggregated across all hold-out folds as an estimate of the expected performance of the EPM RMSD threshold. Further, a multivariable logistic regression model was constructed to examine the association between case/control status, EPM, and clinical factors, including the risk factors of BMI, age, sex, and diabetes status. EPM signals of liver parenchyma between cases and controls were quantitatively validated on an independent patient set using cross validation.

Results

Data Curation

We identified consecutive patients with cirrhosis who underwent surveillance at our high-volume Hepatology clinic from 2012 to 2023. Fifty-eight cases developed HCC on surveillance and fulfilled the selection criteria, and 48 cases included in the study after exclusion of 10 cases given motion artifact (Figure 1). Ninety-nine demographically matched patients were designated as controls. Twenty-five of the 99 controls had lesions (n = 4 with LR-4; n = 21 with LR-3) and that did not progress to HCC with a mean follow-up of 6.9 years (IQR 5.0–8.4). Cases and controls were similar in baseline characteristics, as outlined in Table 1. The median age was 60 years (IQR 55–64) for combined cohort, 59 years (IQR 55–64) for controls, and 60 years (IQR 55–64) for cases. Most patients from both cohorts were male (n=89, 60.5%) compared to female patients (n=58, 39.5%). The median age (range) was 60 (55–64) for combined cohort, 59 years (55–64) for controls, and 60 years (55–64) for cases without significant difference between cases and controls. Most patients racially identified as White/Caucasian (n=129, 87.8%) and

Table 1 Baseline Characteristics of Cases and Controls

Characteristics	Total No. of Patients (N = 147)	No. of Controls (n = 99)	No. of Cases (n = 48)	P value
Age at First Scan, y (%)*				0.89
≤ 40	4 (2.7)	2 (2.0)	2 (4.2)	
41–50	15 (10.2)	10 (10.1)	5 (10.4)	
51–60	65 (44.2)	46 (46.5)	19 (39.6)	
61–70	58 (39.5)	38 (38.4)	20 (41.7)	
>70	5 (3.4)	3 (3.0)	2 (4.2)	
Sex, n (%)				0.74
Male	89 (60.5)	59 (59.6)	30 (62.5)	
Female	58 (39.5)	40 (40.4)	18 (37.5)	
Race, n (%)				
Asian	5 (3.4)	1 (1.0)	4 (8.3)	0.06
Black/African American	8 (5.4)	6 (6.1)	2 (4.2)	

(*Continued*)

Table 1 (Continued).

Characteristics	Total No. of Patients (N = 147)	No. of Controls (n = 99)	No. of Cases (n = 48)	P value
White/Caucasian	129 (87.8)	90 (90.9)	39 (81.3)	
Other	5 (3.4)	2 (1.4)	3 (2.0)	
Ethnicity, n (%)				0.40
Hispanic	37 (25.2)	27 (27.3)	10 (20.8)	
Non-Hispanic	110 (74.8)	72 (72.7)	38 (79.2)	
Body Mass Index, n (%)				0.89
<18.5	0 (0.0)	0 (0.0)	0 (0.0)	
18.5–24.9	26 (17.7)	18 (18.2)	8 (16.7)	
25.0–29.9	40 (27.2)	27 (27.3)	13 (27.1)	
30.0–34.9	52 (35.4)	33 (39.6)	19 (39.6)	
35.0–39.9	23 (15.6)	16 (16.2)	7 (14.6)	
≥40 6	6 (4.1)	5 (5.1)	1 (2.1)	
Cirrhosis Etiology, n (%)				0.82
Autoimmune Hepatitis	5 (3.4)	4 (4.0)	1 (2.1)	
Alcoholic Liver Disease	13 (8.8)	9 (9.1)	4 (8.3)	
Budd-Chiari Syndrome	3 (2.0)	2 (2.0)	1 (2.1)	
Hemochromatosis	1 (0.7)	0 (0.0)	1 (2.1)	
Hepatitis B Virus	6 (4.1)	4 (4.0)	2 (4.2)	
Hepatitis C Virus	72 (49.0)	45 (45.5)	27 (56.3)	
Metabolic Dysfunction-Associated Steatotic Liver Disease	20 (13.6)	15 (15.2)	5 (10.4)	
Nonalcoholic Steatohepatitis	24 (16.3)	18 (18.2)	6 (12.5)	
Primary Biliary Cholangitis	3 (2.0)	2 (2.0)	1 (2.1)	
Diagnosis of Type II Diabetes Mellitus at First Scan, n (%)				0.70
No	101 (68.7)	67 (67.7)	34 (70.8)	
Yes	46 (31.3)	32 (32.3)	14 (29.2)	
		Controls (n = 99)	Cases (n = 48)	
Median AFP, ng/mL (IQR)		3.7 (2.6–5.8)	6.5 (3.8–9.5)	0.003

Notes: *Median age was 60 (IQR 55–64) years for combined cohort, 59 (IQR 55–64) years for controls, and 60 years for cases (IQR 55–64).

ethnically identified as non-Hispanic (n=110, 74.8%). The median AFP during the time of baseline scans was 3.7 (IQR 2.6–5.8) for controls and 6.5 (IQR 3.8–9.5) for cases; median AFP values were significantly different between case and control groups (P=0.003). Hepatitis C virus was the most common etiology (n=72, 49.0%) of cirrhosis, and 31.3% (n=46) of patient had type II diabetes mellitus at the time of their first surveillance scan. Average BMI of controls (30.8, SD 6.2) and cases (30.5, SD 5.3) were not significantly different between groups; P=0.81. The median time between pre-diagnostic and diagnostic scans for cases was 6.8 months (IQR 5.5–11.4).

EPM Analysis

On pre-diagnostic scans for cases, the mean CNR was as follows: 3.62 on EPM, 2.39 on arterial phase, and 1.05 on PV phase. Similarly, on diagnostic scans, the mean CNR was as follows: 3.58 on EPM, 2.35 on arterial phase, and 0.89 on PV phase. An example of the ROI including a lesion that was used to calculate the CNR of the image intensity on the arterial-phase image and the EPM image is shown in Figure 2.

Boxplots of the EPM RMSD within pre-malignant lesions versus background liver parenchyma (with or without benign lesions) are shown in Figure 3. Pre-malignant lesions in cases demonstrated a greater median EPM RMSD on pre-diagnostic scans (LI-RADS 3 and 4) and diagnostic scans (HCC) compared to benign lesions and parenchyma ($p < 0.05$). The average EPM RMSD of the background liver parenchyma in pre-diagnostic and diagnostic scans of the cases and the

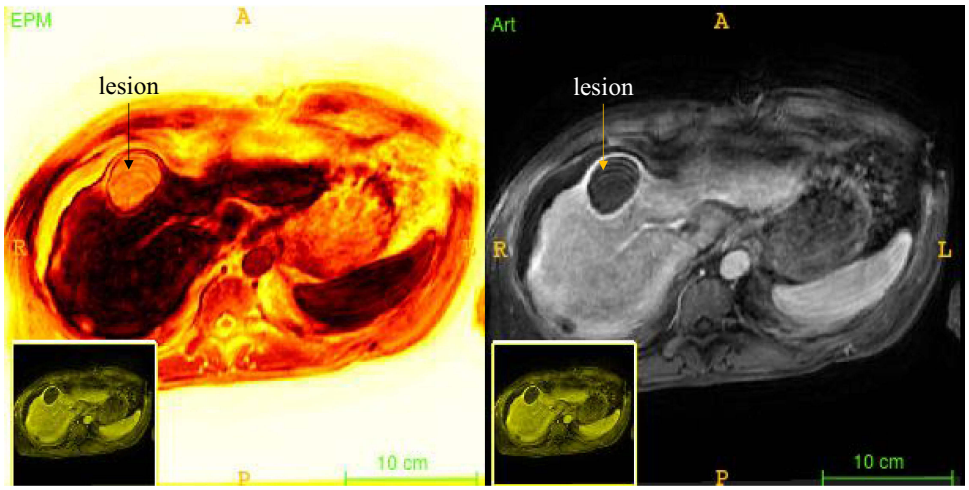

Figure 2 Representative ROI used for CNR analysis of the image intensity on EPM (left) and on arterial phase MRI (right) with lesion indicated.

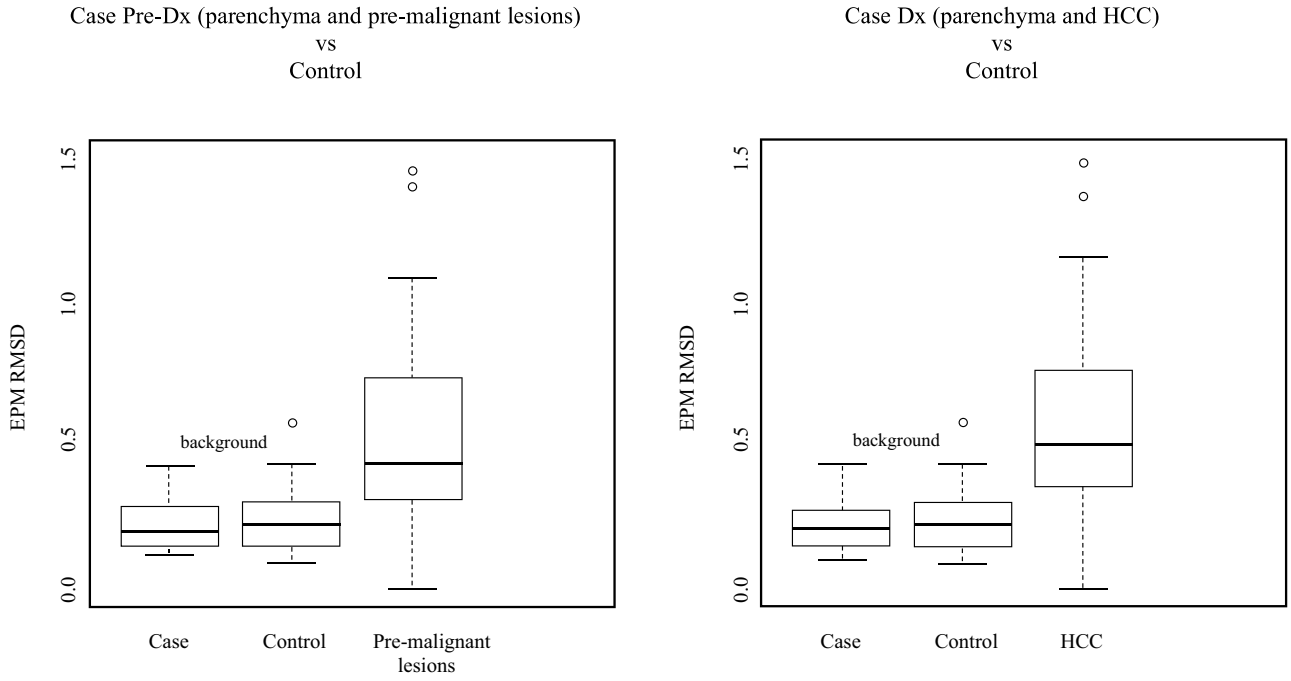

Figure 3 Box plots of EPM RMSD between observed lesions for pre-diagnostic and diagnostic scans in cases. (Left) On pre-diagnostic scans in cases, the median EPM RMSD observed was 0.44 for pre-malignant lesions and 0.22 for parenchyma. (Right) On diagnostic scans, the median EPM RMSD observed was 0.50 for HCC lesions and 0.22 for parenchyma. (Left and right).

single timepoint scan of controls were not statistically different (control parenchymal ROI = 0.22 ± 0.09, case parenchymal ROI = 0.22 ± 0.08, p = 0.8).

ROC analysis for EPM RMSD in differentiating cases and controls on pre-diagnostic and single timepoint scans is shown in Figure 4a. A sensitivity, specificity, and accuracy of 0.65, 0.97, and 0.89 are achieved at the optimal threshold of EPM RMSD 0.37. The area under the curve (AUC) of EPM RMSD at the pre-diagnostic time point was 0.83 (CI: 0.73–0.94). In a multivariable logistic regression model, adjusting for BMI, age, sex, and diabetes status, the association

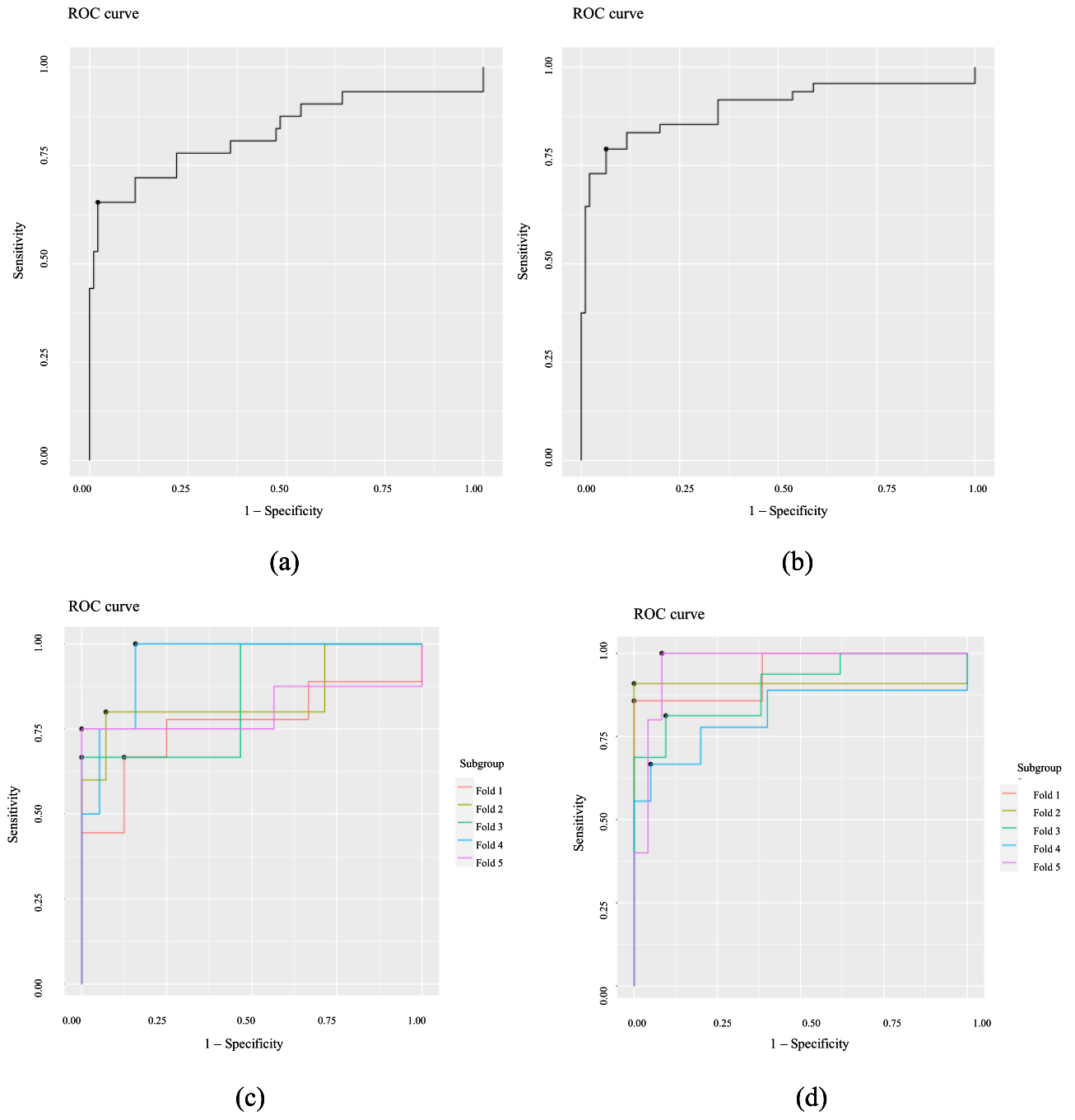

Figure 4 ROC analysis of optimal threshold for discriminating case and control for pre-diagnostic and diagnostic time points. (**a**) In-sample ROC analysis for EPM RMSD in differentiating cases and controls on pre-diagnostic and single timepoint scans is shown. Similarly, (**b**) in-sample ROC analysis for EPM RMSD in differentiating cases and controls on diagnostic and single timepoint scans is shown. (**c**) Five-fold cross validation ROC analysis for EPM RMSD in differentiating cases and controls on pre-diagnostic and single timepoint scans is shown. Similarly, (**d**) Five-fold cross validation ROC analysis for EPM RMSD in differentiating cases and controls on diagnostic and single timepoint scans is shown.

between EPM and eventual HCC status, as all pre-diagnostic timepoints progressed to HCC, is significant (OR=8.08e3; 95% CI: 1.41e2–4.61e5).

ROC analysis for EPM RMSD in differentiating cases and controls on diagnostic and single timepoint scans is shown in Figure 4b. A sensitivity, specificity, and accuracy of 0.79, 0.93, and 0.88 were achieved at the optimal threshold of EPM RMSD 0.35. The AUC of EPM RMSD at the diagnostic time point was 0.89 (CI: 0.82–0.96). In a multivariable logistic regression model, adjusting for BMI, age, sex, and diabetes status, the association between EPM and HCC status is significant (OR=2.91e5; 95% CI: 2.81e2–3.02e7).

Further permutations in ROC analysis for EPM RMSD in differentiating cases and controls on single timepoint and pre-diagnostic scans are shown in Figure 4c. Five-fold cross validation was performed to estimate the out of sample performance on an independent test set. The range of optimum EPM RMSD threshold for discriminating case and control included 0.43, 0.32, 0.030, 0.39, and 0.37. Five-fold analysis achieved an aggregate sensitivity, specificity, and accuracy of 0.75, 0.92, and 0.88; respectively. ROC analysis for EPM RMSD in differentiating cases and controls on single timepoint and diagnostic scans is shown in Figure 4d. Five-fold cross validation was performed to estimate the out of sample performance. The range of optimum EPM RMSD threshold for discriminating case and control included 0.45, 0.32, 0.35, 0.33, and 0.36. Five-fold analysis achieved an aggregate sensitivity, specificity, and accuracy of 0.83, 0.94, and 0.90, respectively.

Discussion

This study investigated the use of a novel EPM technique for distinguishing HCC from background cirrhotic parenchyma with and without benign lesions. Our results suggested that EPM successfully differentiates between cancerous lesions and non-cancerous lesions and parenchyma in a surveillance population approximately six months before they were diagnosed with standard MRI protocol. EPM results show a significant increase in the CNR compared to the arterial and PV phase imaging on both pre-diagnostic and diagnostic scans. The CNR improvement on EPM is due exclusively to lesion signal amplification from calculated intensity differences between multiple-phase scans. With EPM, a RMSD of 0.37 was a quantitative value to characterize lesions that progress to HCC on pre-diagnostic imaging (AUC of 0.83, CI: 0.73–0.94). The median time between pre-diagnostic and diagnostic scans for cases was 6.8 months (IQR 5.5–11.4). Further, the sensitivity (0.79), specificity (0.93), and accuracy (0.88) of EPM show an improvement during the diagnostics scan for HCC. The improvement in EPM performance agrees with intuition that lesions with higher LI-RADs score would be expected to have a stronger EPM signal.

Machine learning methods are a running theme to novel approaches for HCC detection and direct prediction of the LI-RAD score of a liver lesion. An overall accuracy of 60% was previously achieved in automatically assigning a score of LR-1/2, LR-3, LR-4, or LR-5 to a lesion.[18] Paralleling the goal of this study to differentiate HCC from background cirrhotic parenchyma (with and without benign lesions), another study indicated that radiomic features on CT imaging achieve an AUC of the receiver operating characteristic (ROC) curve of 0.7 in diagnosis of HCC from primarily LR-3 and LR-4 lesions.[19] The proposed machine learning models that incorporate MRI imaging instead are expected to further improve the accuracy for differentiating cases and controls. For example, a high accuracy of 90% was achieved in classification when utilizing multiphase liver MRI data to classify LR-3 versus LR-4/5 lesions.[20] Other neural network approaches for HCC lesions have also been developed.[21,22] While HCC was identified with high accuracy (84–92%), a limitation of those studies was the focus on distinguishing HCC from other large liver masses such as cyst, hemangioma, nodular hyperplasia, intrahepatic cholangiocarcinoma, and colorectal metastasis. The considered lesions fall within the categories LR-1, LR-5, and LR-M.[21,22] High classification accuracy is expected for differentiation between LR-1, LR-5, and LR-M lesions, given each category has strong imaging characterization.

In contrast to these prior deep learning studies, our study investigated a neural network approach to differentiate pre-malignant lesions within the LR-3 and LR-4 categories, which are subject to greater heterogeneity and indeterminacy, from benign cirrhotic parenchyma. The consistency in background signal between cases and controls suggests that the EPM RMSD value in background liver parenchyma is unlikely to predict the future location of a new lesion. The current results indicate that an EPM RMSD cutoff of 0.37 would identify LR-3 and LR-4 lesions that progress to HCC. The range of cutoff values in five-fold cross validation agrees with this in-sample cutoff value. Sensitivity, specificity, and

accuracy are comparable between in-sample and cross-validation analyses. The same analysis at the diagnostic time point provides a quantitative reference for the patients with known HCC. Parallels between pre-diagnostic and diagnostic values indicate that the EPM signal is likely to detect HCC earlier in the surveillance period.

Dilation of the liver mask was important in achieving robust results. Generally, over-segmentation of the liver did not decrease the registration accuracy. However, under-segmentation of the liver was prone to more registration errors due to the registration algorithm not directly visualizing the entire liver for guidance. Adding the dilation effectively ensured that the liver was included in the mask used to guide the image registration (see Figure A2 in Supplementary Material). This approach further had the effect of reducing the sensitivity of the approach to segmentation errors. An additional limitation of this study is that the image analysis pipeline was not fully automated. Manual ROI were placed on lesions and background liver parenchyma to facilitate EPM analysis (see Figure A1 in Supplementary Material). As previously referenced and included in the Supplementary Material, the manual and automated ROI methodology produced comparable results; however, the automated approach is limited by potential segmentation inaccuracies. As we improve our fully automated segmentation approach in future efforts, we expect the automated approach to converge to the same quantitative thresholds of the EPM cutoff calculated using our manual ROI placements. Our data did not include comparative analyses between different LI-RADS category lesions, as the primary goal of the study was to assess the foundational feasibility of using EPM to define malignant potential of indeterminant lesions. Additionally, we did not investigate lesions in LI-RADS M category, metastatic lesions, or cholangiocarcinoma. We also did not study the risk of progression to HCC based on cirrhosis etiology.

Conclusions

EPM identifies quantifiable differences between HCC cases and controls in a population with cirrhosis under surveillance, and a threshold cutoff of 0.37 was found to be predictive approximately six months prior to diagnosis of HCC with an accuracy of 89%. Future studies will include cohorts with a larger sample size diversified to have sex, race, ethnicity, and cirrhosis etiology distributions representative of the general population. This may subsequently allow for assessment of EPM in risk stratification for patients under surveillance for HCC, distinguishing which lesions are likely to transform into cancer. Our study introduces EPM as a prospective qualitative means of predicting lesion progression to malignancy, achieving early curative interventions, and individualizing care for patients at risk for development of HCC.

Abbreviations

AASLD, American Association for the Study of Liver Diseases; AFP, alpha fetoprotein; AUC, area under the curve; CNR, contrast to noise ratio; CT, computed tomography; EASL, European Association for the Study of the Liver; EPM, enhancement pattern mapping; HCC, hepatocellular carcinoma; IQR, interquartile range; LI-RADS, Liver Imaging Reporting and Data System; LR-1, LI-RADS-1; LR-2, LI-RADS-2; LR-3, LI-RADS-3; LR-4, LI-RADS-4; LR-5, LI-RADS-5; LR-M, LI-RADS-M; MRI, magnetic resonance imaging; PV, portovenous; RMSD, root mean square difference; ROC, receiver operating characteristic; ROI, region of interest; SD, standard deviation; US, ultrasound; VEGF, vascular endothelial growth factors.

Data Sharing Statement

The methodology for imaging acquisition, normalization, registration, and more extensive ROC analyses are included as Supplementary Materials to the Methods (Supplementary Data 1). Additional data can be made electronically, individually, and appropriately available upon request to Dr. Prasun K. Jalal (jalal@bcm.edu).

Ethics Approval and Informed Consent

Ethics approval was obtained through Baylor College of Medicine Institutional Review Board (H-47711 and H-45208) and MD Anderson Cancer Center (PA14-0646). Participants recruited prospectively provided written informed consent. All participant data was anonymized and maintained with confidentiality in compliance with the Declaration of Helsinki.

Acknowledgments

We thank Dr. Lauren Pinzás (Otolaryngology – Head and Neck Surgery, The University of Chicago Medical Center) for her edits and guidance on statistical methods. Eugene J. Koay also gratefully acknowledges the support of Mr. Greg Goodman, Jennifer and Wil vanLoh, KWS, and Michael Sinclair.

Furthermore, an earlier version of the abstract of this paper was presented at the *European Association for the Study of the Liver Congress 2023* as an abstract and poster presentation with interim findings. The poster's abstract was published in "Poster Presentations" in *Journal of Hepatology*: https://doi.org/10.1016/S0168-8278(23)01269-2.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Prasun K. Jalal would like to gratefully acknowledge the generous support of the Dora Roberts Foundation. Eugene J. Koay would also like to acknowledge generous support from Jennifer and Wil VanLoh, KWS Foundation, The Mount Brilliant Foundation, and Michael C. Linn Family Foundation. Eugene Koay, David Fuentes, and Laura Beretta would like to acknowledge research support from the NIH (R01CA195524).

Disclosure

Prasun K. Jalal reports research support from the Dora Roberts Foundation. Eugene J. Koay reports research grants from DOD, CPRIT and NIH. In addition, he also reports personal fees from Taylor and Francis, LLC, AstraZeneca, RenovoRx, Aptitude Health, Apollo Health, and Bayer Healthcare; sponsored research from GE Healthcare, Philips Healthcare, Varian, Artidis, and AstraZeneca; licensing of patent from Kallisio, during the conduct of the study. All other authors have no competing interests to disclose for this work.

References

1. Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. *Hepatology*. 2018;68(2):723–750. doi:10.1002/hep.29913
2. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatocellular carcinoma. *J Hepatol*. 2018;69(1):182–236. doi:10.1016/j.jhep.2018.03.019
3. Singal AG, Lampertico P, Nahon P. Epidemiology and surveillance for hepatocellular carcinoma: new trends. *J Hepatol*. 2020;72(2):250–261. doi:10.1016/j.jhep.2019.08.025
4. Tzartzeva K, Obi J, Rich NE, et al. Surveillance imaging and alpha fetoprotein for early detection of hepatocellular carcinoma in patients with cirrhosis: a meta-analysis. *Gastroenterology*. 2018;154(6):1706–1718 e1. doi:10.1053/j.gastro.2018.01.064
5. Giannini EG, Cucchetti A, Erroi V, Garuti F, Odaldi F, Trevisani F. Surveillance for early diagnosis of hepatocellular carcinoma: how best to do it? *World J Gastroenterol*. 2013;19(47):8808–8821. doi:10.3748/wjg.v19.i47.8808
6. Hassett M, Yopp AC, Singal AG, Palmer BF. Surveillance for hepatocellular carcinoma: how can we do better? *Am J Med Sci*. 2013;346(4):308–313. doi:10.1097/MAJ.0b013e31828318ff
7. Singal AG, Nehra M, Adams-Huet B, et al. Detection of hepatocellular carcinoma at advanced stages among patients in the HALT-C trial: where did surveillance fail? *Am J Gastroenterol*. 2013;108(3):425–432. doi:10.1038/ajg.2012.449
8. Simmons O, Fetzer DT, Yokoo T, et al. Predictors of adequate ultrasound quality for hepatocellular carcinoma surveillance in patients with cirrhosis. *Aliment Pharmacol Ther*. 2017;45(1):169–177. doi:10.1111/apt.13841
9. Kim HL, An J, Park JA, Park SH, Lim YS, Lee EK. Magnetic resonance imaging is cost-effective for hepatocellular carcinoma surveillance in high-risk patients with cirrhosis. *Hepatology*. 2019;69(4):1599–1613. doi:10.1002/hep.30330
10. Mahmud N, Hoteit MA, Goldberg DS. Risk factors and center-level variation in hepatocellular carcinoma under-staging for liver transplantation. *Liver Transpl*. 2020;26(8):977–988. doi:10.1002/lt.25787
11. Park PC, Choi GW, M. Zaid M, et al. Enhancement pattern mapping technique for improving contrast-to-noise ratios and detectability of hepatobiliary tumors on multiphase computed tomography. *Med Phys*. 2020;47(1):64–74. doi:10.1002/mp.13769
12. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006;31(3):1116–1128. doi:10.1016/j.neuroimage.2006.01.015
13. Goldberg DS, Taddei TH, Serper M, et al. Identifying barriers to hepatocellular carcinoma surveillance in a national sample of patients with cirrhosis. *Hepatology*. 2017;65(3):864–874. doi:10.1002/hep.28765

14. Chernyak V, Fowler KJ, Kamaya A, 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
15. Lee YT, Wang JJ, Zhu Y, Agopian VG, Tseng HR, Yang JD. Diagnostic Criteria and LI-RADS for Hepatocellular Carcinoma. *Clin Liver Dis*. 2021;17(6):409–413. doi:10.1002/cld.1075
16. Zhang L, Wang JN, Tang JM, et al. VEGF is essential for the growth and migration of human hepatocellular carcinoma cells. *Mol Biol Rep*. 2012;39(5):5085–5093. doi:10.1007/s11033-011-1304-2
17. Morshid A, Elsayes KM, Khalaf AM, et al. A machine learning model to predict hepatocellular carcinoma response to transcatheter arterial chemoembolization. *Radiol Artif Intell*. 2019;1(5):e180021. doi:10.1148/ryai.2019180021
18. Yamashita R, Mittendorf A, Zhu Z, et al. Deep convolutional neural network applied to the liver imaging reporting and data system (LI-RADS) version 2014 category classification: a pilot study. *Abdom Radiol*. 2020;45(1):24–35. doi:10.1007/s00261-019-02306-7
19. Mokrane FZ, Lu L, Vavasseur A, et al. Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules. *Eur Radiol*. 2020;30(1):558–570. doi:10.1007/s00330-019-06347-w
20. Wu Y, White GM, Cornelius T, et al. Deep learning LI-RADS grading system based on contrast enhanced multiphase MRI for differentiation between LR-3 and LR-4/LR-5 liver tumors. *Ann Transl Med*. 2020;8(11):701. doi:10.21037/atm.2019.12.151
21. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. *Radiology*. 2018;286(3):887–896. doi:10.1148/radiol.2017170706
22. Hamm CA, Wang CJ, Savic LJ, et al. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol*. 2019;29(7):3338–3347. doi:10.1007/s00330-019-06205-9