ORIGINAL RESEARCH

# Machine Learning-Based CEMRI Radiomics Integrating LI-RADS Features Achieves Optimal Evaluation of Hepatocellular Carcinoma Differentiation

Hai-Feng Liu, Yang Lu, Qing Wang, Yu-Jie Lu, Wei Xing

Department of Radiology, Third Affiliated Hospital of Soochow University, Changzhou, Jiangsu, 213000, People's Republic of China

Correspondence: Wei Xing, Third Affiliated Hospital of Soochow University, 185 Juqian Street, Changzhou City, Jiangsu Province, People's Republic of China, Email suzhxingwei@suda.edu.cn

**Purpose:** To develop and compare various machine learning (ML) classifiers that employ radiomics extracted from contrast-enhanced magnetic resonance imaging (CEMRI) for diagnosing pathological differentiation of hepatocellular carcinoma (HCC), and validate the performance of the best model.

**Methods:** A total of 251 patients with HCCs (n = 262) were assigned to a training (n = 200) cohort and a validation (n = 62) cohort. A collection of 5502 radiomics signatures were extracted from the CEMRI images for each HCC nodule. To reduce redundancy and dimensionality, Spearman rank correlation, minimum redundancy maximum relevance (mRMR), and the least absolute shrinkage and selection operator (LASSO) approach were employed. Eight ML classifiers were trained to obtain the best radiomics model. The performance of each model was evaluated based on the area under the receiver operating characteristic curve (AUC). The radiomics model was integrated with liver imaging reporting and data system (LI-RADS) features to design a combined model.

**Results:** The eXtreme Gradient Boosting (XGBoost)-based radiomics model outperformed other ML classifiers in evaluating pHCC, achieving an AUC of 1.00 and accuracy of 1.00 in the training cohort. The LI-RADS model demonstrated an AUC value of 0.77 and 0.82 in the training and validation cohorts. The combined model exhibited best performance in both the training and validation cohorts, with AUCs of 1.00 and 0.86 for evaluating HCC differentiation, respectively.

**Conclusion:** CEMRI radiomics integrating LI-RADS features demonstrated excellent performance in evaluating HCC differentiation, suggesting an optimal clinical decision tool for individualized diagnosis of HCC differentiation.

**Keywords:** machine learning, radiomics, LI-RADS, differentiation, hepatocellular carcinoma, HCC

## Introduction

Hepatocellular carcinoma (HCC) accounts for approximately 75–85% of all primary liver malignancies, ranking it as the sixth most common cancer worldwide and a major cause of cancer-related mortality.[1,2] Poor tumor differentiation is a well-established prognostic factor associated with higher rates of postoperative recurrence (up to 50–70%) and worse survival outcomes in patients undergoing curative therapy for HCC.[3,4] Liver biopsy is currently served as the gold standard for preoperative diagnosis of HCC differentiation, but it is accompanied by several complications including invasiveness and potential sampling bias.[5] Consequently, there is an urgent requirement for accurate and noninvasive methods to predict HCC differentiation.

Radiomics enables a comprehensive evaluation of intratumoral heterogeneity on a millimeter scale by extracting numerous quantitative signatures, thereby demonstrating significant capability in assessing the biological characteristics of HCC.[6] Contrast-enhanced magnetic resonance imaging (CEMRI) has proven to be crucial in noninvasive HCC characterization, offering excellent soft-tissue resolution and better representation of the actual heterogeneity of HCC

compared to other imaging modalities.[7,8] Therefore, it is possible to speculate that a CEMRI-based radiomics model might enhance the accuracy of HCC differentiation. Furthermore, machine learning (ML) algorithms provide an artificial intelligence approach to automatically establish more effective decision-making models using radiomics features. Various ML classifiers have recently been applied to assist clinicians for treatment planning and prognosis assessment in cancer patients.[9] However, the use of ML classifiers for CEMRI-based radiomics in evaluating HCC differentiation has yet to be explored and compared.

The Liver Imaging Reporting and Data System (LI-RADS) offers a comprehensive framework for evaluating key features associated with HCC, which has promising value in diagnosing HCC, predicting microvascular invasion, and assessing treatment response.[10] Therefore, one of the aims of our study was to identify MRI LI-RADS features that could potentially serve as indicators of HCC differentiation. However, as far as we know, there is a scarcity of developed models that combine ML-based CEMRI radiomics with LI-RADS features for evaluating HCC differentiation.

The purpose of this study was to establish and compare the performance of CEMRI-based radiomics model using different ML classifiers for predicting HCC differentiation and to validate the most effective model. Furthermore, an investigation was conducted to explore a combined model that incorporated both radiomics and LI-RADS features.

## Materials and Methods

### Patients Selection

This study received approval from the Institutional Ethics Committee of third affiliated hospital of Soochow University and informed consent was not required from the included participants (2022-CL027-01). A retrospective analysis was conducted on a total of 273 consecutive patients who underwent preoperative MRI examinations between January 2017 and May 2023 and were subsequently diagnosed with HCC after surgery. Patients who met any of the following criteria were excluded: 1) HCC-related therapies, such as radiofrequency ablation or transarterial chemoembolization, prior to MRI scanning (n = 6); 2) Patients with incomplete MRI examinations, or poor MRI image quality due to breathing motion artifacts, were deemed incompatible with radiomics analysis (n = 7); 3) Differentiation reports were not included (n = 9). In total, 251 patients were ultimately included and randomly assigned to either the training or validation cohort in chronological order, while maintaining a ratio of 8:2.

### Pathological Differentiation Analysis

The liver specimens were stained with hematoxylin and eosin (HE) staining to determine HCC differentiation by an experienced pathologist who was unaware of the preoperative findings. According to the 2019 WHO classification,[11] the differentiation was classified into well differentiated, moderately differentiated, and poorly differentiated HCC (pHCC). For HCC nodules that exhibited multiple types of differentiation, the final diagnosis was based on the most predominant differentiation observed in the tumor tissue. Specifically, we classified both well differentiated and moderately differentiated HCC as non-poorly differentiated HCC (npHCC).

### MRI Protocol and Features Interpretation

The 3.0-T MRI scanner, Magnetom Verio (Siemens Healthineers, Germany), was utilized for imaging. The following sequences were performed: 1) in-phase and out-phase T1-weighted imaging (T1WI); 2) T2-weighted imaging (T2WI); 3) diffusion-weighted imaging (DWI) in echo planar imaging with a b value of 0 and 800 s/mm$^2$; 4) DCE-MRI: a total of 0.2 mL/kg of Gd-DTPA (Magnevist, Bayer, Germany) was injected at a rate of 1 mL/s for bolus tracking. 3D volumetric interpolated breath-hold examination (VIBE) techniques were employed to acquire images in the arterial phase (AP, 25~35 s), portal venous phase (PVP, 60~70 s), and delayed phase (DP, 180 s). The entire DCE-MRI scan lasted for approximately 3~4 minutes. Detailed information on the sequences and parameters can be found in Table S1.

According to the LI-RADS criteria (Version 2018),[10] MRI features were interpreted by two radiologists with seven and twelve years of experience in liver imaging, respectively. Both radiologists were blinded to the pathological results, and any inconsistencies were resolved by consensus. The following features were collected: (a) major features: nonrim arterial-phase hyperenhancement (NAPHE), nonperipheral washout, enhancing capsule; (b) nodule-in-nodule, mosaic

architecture, blood products or higher fat content in HCC, nonenhancing capsule, which are particular auxiliary features of HCC; (c) auxiliary features favoring malignancy but not HCC in specific, mild-to-moderate T2 hyperintensity, delayed phase hypointensity, iron sparing in the tumor, corona enhancement; (d) baseline features: HCC number, size, margin, shape.

## Imaging Preprocessing and Radiomics Signatures Extraction

Before region of interest (ROI) delineation, N4 bias correction, resampling at a voxel spacing of $1*1*1 \text{ mm}^3$, and grayscale normalization are sequentially performed to eliminate the inhomogeneity of magnetic field intensity, compensate for voxel spatial differences, and maintain grayscale consistency, respectively. Then, all preprocessed CEMRI (AP, PVP, DP) images were imported into ITK-SNAP software (version 3.8.0). Radiologist 1 manually delineated the tumor boundary on each consecutive transverse slice of the images, and a 3D-ROI was automatically generated. During the 3D-ROI segmentation process, it was important to ensure that the entire tumor volume was encompassed while excluding any perfusion disorders or noticeable satellite lesions. Additionally, 30 HCC lesions were randomly selected for both Radiologist 1 and Radiologist 2 to delineate the ROI again.

The Pyradiomics package (Version 2.1.2) was utilized to extract a set of radiomics signatures, including (a) first-order (n = 360), shape (n = 14), (b) gray-level cooccurrence matrix (GLCM) (n = 440), (c) gray-level dependence matrix (GLDM) (n = 280), (d) gray level run length matrix (GLRLM) features (n = 320), (e) gray-level size zone matrix (GLSZM) (n = 320), (f) neighboring gray tone difference matrix (NGTDM) (n = 100). Consequently, a total of 5502 signatures were obtained from each HCC nodule, as 1834 signatures were extracted from each CEMRI sequence. To reduce interference among the radiomics dimensions, all extracted signatures were normalized using the Z-score approach.

## Radiomics Model Development and Machine Learning Classifiers

A three-step procedure was sequentially implemented to identify optical signatures. First, the Spearman rank correlation was used to exclude signatures with a correlation coefficient higher than 0.9, thus reducing redundancy. Second, a tenfold cross-validation was employed with the least absolute shrinkage and selection operator (LASSO) approach for identifying the most accurate signatures with nonzero coefficients, determined by an optimal penalty parameter (λ) based on the minimum criteria. Third, the maximum relevance minimum redundancy (mRMR) approach was utilized to reduce dimensionality by maximizing the correlation between retained signatures.

By calculating and combining selected optical signatures with different algorithms, the logistic regression (LR), support vector machine (SVM), K-nearest neighbor (KNN), random forest (RF), ExtraTrees, eXtreme Gradient Boosting (XGBoost), light gradient boosting machine (LightGBM) and multilayer perceptron (MLP) classifiers were applied to develop eight risk stratification models for evaluating pHCC. To identify the best-performing ML model, the diagnostic ability of each model in evaluating pHCC was compared. After selecting the best model, the optical signatures were weighted and combined based on their respective coefficients to diagnose pHCC using the chosen classifier, contributing to the establishment of a robust radiomics model. The model was trained using the training cohort statistics and tested on the validation cohort.

## Combined Model Development and Verification

The significant MRI LI-RADS features, decided by intra-group statistical comparison, along with the radiomics signatures, were combined to develop a nomogram model. To assess the net clinical benefits and goodness-of-fit of the nomogram model in predicting pHCC, a decision curve analysis (DCA) and calibration curve with the Hosmer–Lemeshow test were performed. In the DCA curve, the red solid line represents the performance of the nomogram, the efficacy of the nomogram model is considered better when the red line is closer to the diagonal dashed line. Additionally, the nomogram model demonstrated improved net clinical benefits for evaluating pHCC when the nomogram model curve was higher than either the "treat all" or "treat none" across a full range of reasonable threshold probabilities. The detailed flowchart, illustrating the process from DEC-MRI scanning to nomogram model development, is shown in Figure 1.
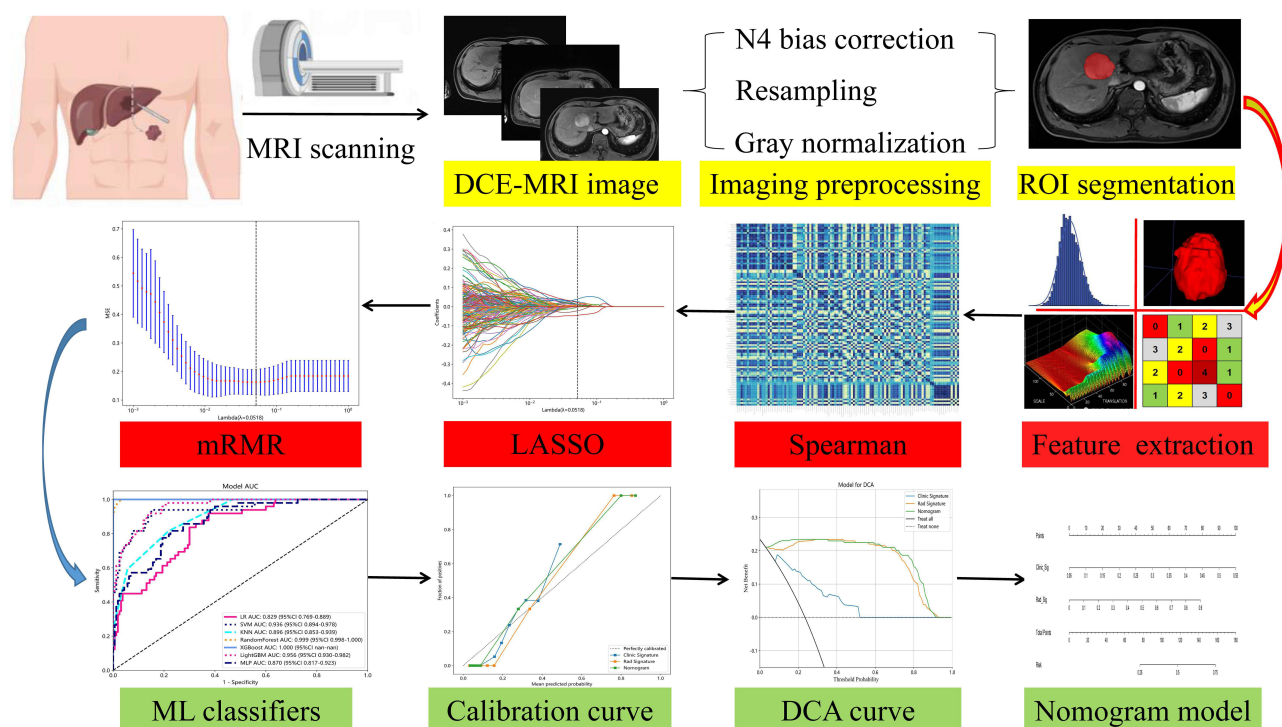
**Figure 1** Detailed flowchart from DCE-MRI scanning to nomogram model development.

## Statistical Analysis

Statistical analyses were performed with R software (version 4.0). Categorical data are listed as the percentage or median (interquartile range) after normality test with the Shapiro–Wilk method, and continuous variables are presented as the mean ± standard deviation (SD). Intraclass correlation coefficient (ICC) was used to analyze inter- and intra-reader variability. Mann–Whitney U or chi-square ($\chi^2$) tests were used to detect differences between pHCC and npHCC. The diagnostic performance for predicting pHCC was evaluated using a receiver operating characteristic (ROC) curve, sensitivity, specificity, accuracy, and area under the ROC curve (AUC). In addition, the Delong method was adopted to compare the differences in AUC values produced by the various models. In each of the tests, statistical significance was achieved when the two-tailed $P$ value <0.05.

# Results

## Patient Characteristics and Pathological Differentiation

There were 251 patients (male/female: 201/50, aged 62.6±10.4 years) with 262 HCCs (eleven patients presented two HCCs), confirmed by pathology results who underwent preoperative MRI examination and were included and randomly classified into a training cohort (201 patients and 210 HCCs) and a validation cohort (50 patients and 52 HCCs). Among the included 262 HCCs, 50 (23.8%) and 160 (76.2%) were diagnosed as pHCC and npHCC in the training cohort, they were 12 (23.1%) and 40 (76.9%) in the validation cohort, respectively. The training and validation cohorts were comparable in terms of age, sex, etiology, alpha-fetoprotein (AFP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (TB), prothrombin time (PT), Child-Pugh grade, performance status (PS), Barcelona clinical liver cancer (BCLC) stage, HCC number and differentiation. The detailed baseline characteristics are shown in Table 1.

## Radiomics Signatures Selection

The inter- and intrareader variability for radiomics signatures was 0.88 (95% CI: 0.81–0.93) and 0.91 (95% CI: 0.84–0.95), respectively. After removing redundant radiomic parameters with Spearman correlation coefficients greater than

Table I Baseline Characteristics of Included Patients Diagnosed with HCC

|  | Total (n=251) | TrainingCohort (n=201) | Validation Cohort (n=50) | t/χ2/Z | P value |
|---|---|---|---|---|---|
| **Age** (Years) | 62.6±10.4 | 62.3±10.3 | 63.7±10.8 | 0.80 | 0.42 |
| **Sex** |  |  |  |  |  |
| Male | 201(80.1%) | 162(83.8%) | 39(77.1%) | 0.17 | 0.68 |
| Female | 50(19.9%) | 39(16.2%) | 11(22.9%) |  |  |
| **Etiology** |  |  |  |  |  |
| HBV/HCV | 188(74.9%) | 150(74.6%) | 38(76.0%) | 0.04 | 0.84 |
| Other-None | 63(25.1%) | 51(25.4%) | 12(24.0%) |  |  |
| **AFP** |  |  |  |  |  |
| ≥20 ng/mL | 98(39.0%) | 81(40.3%) | 19(38.0%) | 0.08 | 0.77 |
| < 20 ng/mL | 153(61.0%) | 120(59.7%) | 31(62.0%) |  |  |
| **ALT** (U/L) | 26.9(18.4–39.7) | 26.9(18.9–39.9) | 26.5(17.1–38.0) | −0.74 | 0.46 |
| **AST** (U/L) | 30.4(25.2–39.2) | 32.0(26.0–44.0) | 29.0(22.5–37.0) | −1.63 | 0.10 |
| **TB** (μmol/L) | 13.4(10.9–19.1) | 13.5(11.2–19.3) | 12.9(10.4–17.9) | −0.53 | 0.60 |
| **PT** (s) | 12.2(11.6–13.0) | 12.2(11.6–12.9) | 12.1(11.3–13.0) | 1.72 | 0.09 |
| **Child-Pugh grade** |  |  |  |  |  |
| A | 225(89.6%) | 183(91.0%) | 42(84.0%) | 2.14 | 0.14 |
| B | 26(10.4%) | 18(9.0%) | 8(16.0%) |  |  |
| **PS** |  |  |  |  |  |
| 0 | 248(98.8%) | 199(99.1%) | 49(98.0%) | 0.34 | 0.56 |
| 1 | 3(1.2%) | 2(0.9%) | 1(2.0%) |  |  |
| **HCC Number** |  |  |  |  |  |
| Solitary | 240(95.6%) | 192(95.5%) | 48(96.0%) | 0.20 | 0.65 |
| Two | 11(4.4%) | 9(4.5%) | 2(4.0%) |  |  |
| **BCLC stage** |  |  |  |  |  |
| 0~A | 221(88.0%) | 175(87.1%) | 46(92.0%) | 1.44 | 0.70 |
| B~C | 30(12.0%) | 26(12.9%) | 4(8.0%) |  |  |
| **HCC differentiation** |  |  |  |  |  |
| Well / Moderately | 31(25.9%)/158(18.3%) | 23(11.4%)/128(63.7%) | 8(16.0%)/30(60.0%) | 0.77 | 0.68 |
| Poorly | 62(55.8%) | 50(24.9%) | 12(24.0%) |  |  |

**Abbreviations**: HCC, hepatocellular carcinoma; HBV, hepatitis B virus; HCV, hepatitis C virus; AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TB, total bilirubin; PT, prothrombin time; PS, performance status; BCLC, Barcelona clinical liver cancer; NA, Not available.

0.90 (Figure 2A), a total of 305, 313, and 318 signatures were retained from AP, PVP, and DP, respectively. Upon applying mRMR (Figure 2B) and the LASSO approach (Figure 2C) for further dimension reduction and redundancy removal, a total of eighteen highly robust signatures (Figure 2D) were selected and retained to train the ML classifier models (Table S2).

## Comparison of Various ML Classifiers Models

The AUCs of different ML classifiers in diagnosing pHCC ranged from 0.85 to 1.00 in the training cohort and from 0.73 to 0.88 in the validation cohort. Among the eight ML classifiers, the XGBoost model achieved the highest AUC (1.00) and accuracy (1.00), followed by Extratrees (AUC = 1.00, accuracy = 1.00), RF (AUC = 1.00, accuracy = 0.995), LightGBM (AUC = 0.96, accuracy = 0.86), SVM (AUC = 0.88, accuracy = 0.88), KNN (AUC = 0.89, accuracy = 0.80), MLP (AUC = 0.88, accuracy = 0.76), and LR (AUC = 0.85, accuracy = 0.78). The detailed information of various ML classifiers for diagnosing pHCC is provided in Table 2 and illustrated in Figures 3.

## MRI LI-RADS Features Model

As presented in Table 3, larger tumor size, nonsmooth margin, mosaic architecture, intratumoral blood products, and corona enhancement were identified as significant risk factors associated with pHCC. The LI-RADS model, which combines these five significant MRI features, demonstrated an accuracy of 0.72, sensitivity of 0.69, specificity of 0.73,
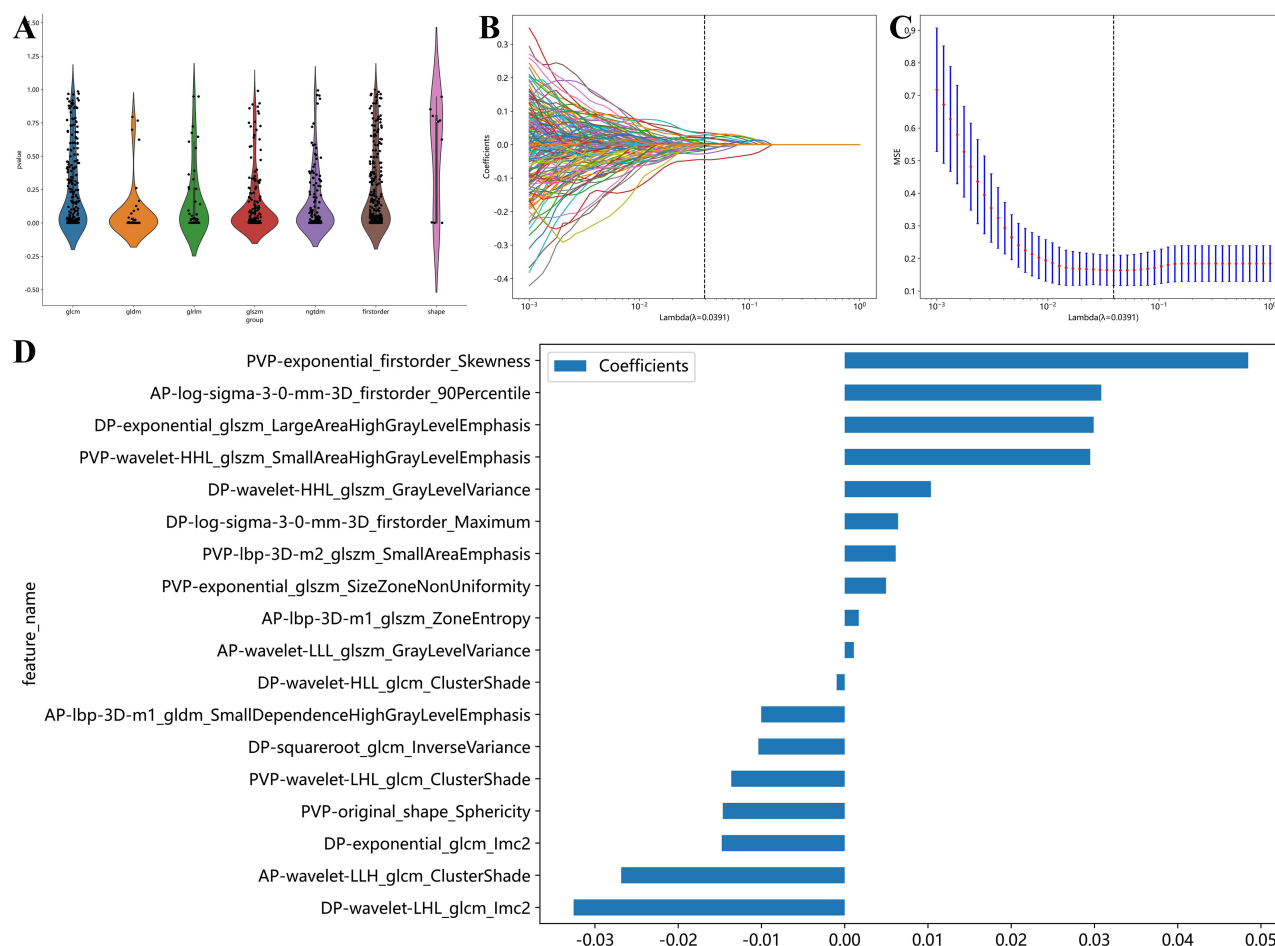
**Figure 2** The distribution plot of retained radiomics signatures after Spearman rank correlation (**A**). Using the LASSO method with tenfold cross-validation based on the minimum binomial deviation, the most valuable signatures with nonzero coefficients were identified (**B**). At lambda ($\lambda$) = 0.0391, eighteen signatures identified by mRMR method were demonstrated (**C**). The weighted importance of the selected eighteen signatures (**D**).

and an AUC value of 0.77 (95% CI: 0.70–0.83) in the training cohort. In the validation cohort, the corresponding values were 0.68, 0.92, 0.60, and 0.82 (95% CI: 0.70–0.94).

## Combined Model Establishment and Verification

The nomogram model that incorporated XGBoost-based Rad-signatures and significant MRI LI-RADS features demonstrated an AUCs value of 1.00 (95% CI: 0.99–1.00) and 0.86 (95% CI: 0.76–0.96), for evaluating pHCC in the training and validation cohorts, respectively, shown in (Figure 4A and B). The calibration curves indicated that three models (Radiomics, LI-RADS, nomogram) presented good agreement in both training and validation cohorts (Figure 4C and D). Furthermore, the DCA results suggested that the nomogram model curve was higher than either "treat all" (pHCC, solid black line) or "treat none" (npHCC, virtual black line) across a full range of reasonable threshold probabilities, suggesting that the nomogram model has improved clinical benefits in individualized clinical decision-making for HCC differentiation (Figure 4E and F).

## Comparison of Diagnostic Ability Among Different Models

The radiomics model demonstrated a significantly higher AUC compared to the LI-RADS model's: 1.00 vs 0.77 ($P<0.05$) in the training cohort, 0.79 vs 0.68 ($P<0.05$) in the validation cohort, indicating the improved discriminative ability of radiomics model in diagnosing pHCC. Furthermore, the nomogram model (Figure 5) demonstrated superior performance,

**Table 2** Various Machine Learning Classifiers for Diagnosing pHCC

| Cohorts | Machine Learning Classifiers | Accuracy | AUC | 95% CI | Sensitivity | Specificity |
|---------|------------------------------|----------|-----|--------|-------------|-------------|
| Training | LR | 0.78 | 0.85 | 0.80–0.91 | 0.90 | 0.74 |
| Validation | LR | 0.85 | 0.86 | 0.75–0.97 | 0.77 | 0.88 |
| Training | SVM | 0.88 | 0.88 | 0.81–0.95 | 0.84 | 0.90 |
| Validation | SVM | 0.72 | 0.81 | 0.69–0.94 | 0.92 | 0.67 |
| Training | KNN | 0.80 | 0.88 | 0.84–0.93 | 0.84 | 0.80 |
| Validation | KNN | 0.64 | 0.75 | 0.62–0.88 | 0.85 | 0.60 |
| Training | RF | 0.995 | 1.00 | 1.00–1.00 | 1.00 | 0.99 |
| Validation | RF | 0.68 | 0.73 | 0.58–0.88 | 0.77 | 0.70 |
| Training | ExtraTrees | 1.00 | 1.00 | 1.00–1.00 | 1.00 | 1.00 |
| Validation | ExtraTrees | 0.72 | 0.77 | 0.64–0.91 | 0.69 | 0.74 |
| Training | XGBoost | 1.00 | 1.00 | 1.00–1.00 | 1.00 | 1.00 |
| Validation | XGBoost | 0.79 | 0.84 | 0.73–0.95 | 0.85 | 0.80 |
| Training | LightGBM | 0.86 | 0.96 | 0.93–0.98 | 0.96 | 0.83 |
| Validation | LightGBM | 0.62 | 0.77 | 0.62–0.92 | 0.92 | 0.53 |
| Training | MLP | 0.78 | 0.88 | 0.82–0.93 | 0.90 | 0.74 |
| Validation | MLP | 0.77 | 0.88 | 0.79–0.97 | 1.00 | 0.70 |

with AUCs of 1.00 and 0.86, respectively, indicating the stronger ability of the nomogram model in the prediction of HCC differentiation.

## Discussion

To identify the most effective and noninvasive approach for predicting HCC differentiation, we established and validated a CEMRI-based radiomics model with various ML classifiers. Our study concluded that the CEMRI-radiomics model, specifically when utilizing the XGBoost classifier, exhibited excellent diagnostic ability with an AUC of 1.00 in diagnosing pHCC. More meaningfully, the nomogram model, which incorporated significant LI-RADS features and XGBoost-based signatures, demonstrated the best performance in both the training and validation cohorts, with high goodness-of-fit and improved net benefits in diagnosing pHCC.

The majority of HCC cases are typically diagnosed at an advanced stage with a poor prognosis, consequently, this diminishes the viability of surgical resection as the primary treatment option for HCC. To improve the prognosis and effectiveness monitoring of HCC, clinical studies have explored programmed death-1 (PD-1) inhibitors with antiangiogenic therapy through TACE,[12] as well as the combination of radiotherapy with sorafenib.[13] These studies have shown promising therapeutic outcomes in cases of unresectable HCC, providing valuable guidance for selecting the most effective treatment approach for HCC in the future. Moreover, an increasing number of studies have been conducted to select more accurate biomarkers for predicting treatment response and survival prognosis in HCC patients by developing a nomogram model. For instance, both alkaline phosphatase (ALP)[14] and albumin-bilirubin (ALBI)[12] nomogram model demonstrated good performance for predicting survival rate for HCC. Similarly, Li et al[15] introduced an aspartate aminotransferase lymphocyte rate (ASR score) nomogram model, which presented valuable performance to predict the prognosis of HCC treated with TACE or TACE combined with systemic therapy. In this study, the nomogram model that combining CEMRI-based radiomics with LI-RADS features provides an optimal clinical decision approach for individualized clinical diagnosis of HCC differentiation, indicating the potential value of nomogram model.

By applying ML algorithms and statistical analysis to the extracted radiomics features, it becomes develop to create predictive models that capture the underlying characteristics and heterogeneity of tumors at a microscopic level. These radiomics models have the potential to differentiate hepatocellular carcinoma (HCC) from other liver lesions, evaluate tumor stage and progression, predict treatment response, and even provide prognostic information.[6,16] Recently, several radiomics studies have been performed using hepatobiliary phase (HBP) to predict HCC differentiation. However, these
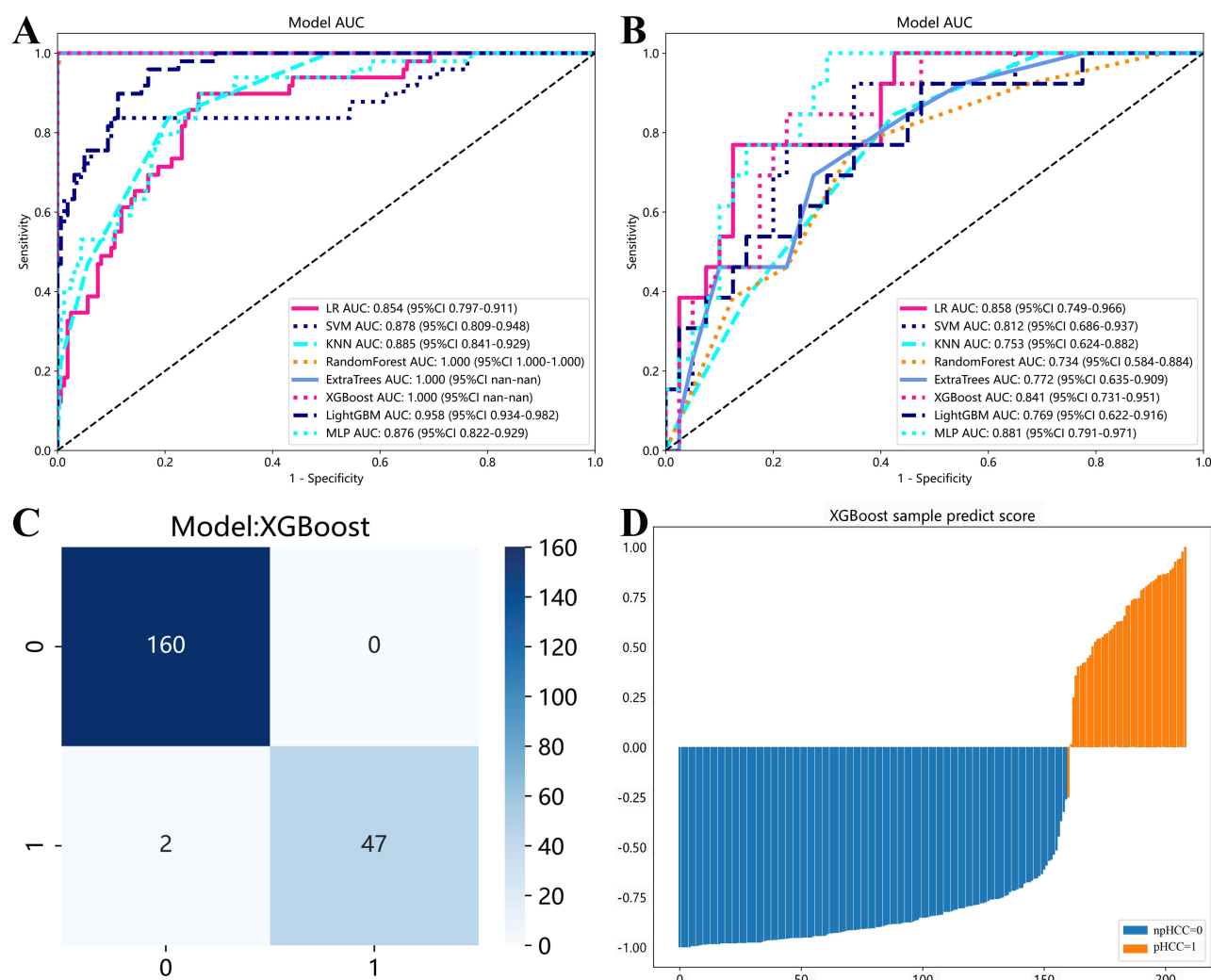
**Figure 3** The ROC curve of eight machine learning classifiers in diagnosing HCC differentiation in the training (**A**) and validation (**B**) cohorts. The confusion matrix (**C**) and sample prediction histogram (**D**) of XGBoost classifier in evaluating HCC differentiation.

studies demonstrated relatively lower AUC values,[17–19] ranging from 0.61 to 0.83, partly because only HBP-based radiomics was not sufficiently accurate in characterizing HCC heterogeneity. Moreover, both Brancato et al[20] and Yang et al[21] developed CEMRI-based radiomics models for predicting HCC, but the inclusion of a relatively limited number of patients (n = 38~188) and the extraction of a restricted number of radiomics signatures (38~108 per sequence) were likely to decrease the reliability of the study findings. Furthermore, no relevant studies have yet been conducted to compare the diagnostic differences using different ML classifiers in predicting HCC differentiation.

In our study, we observed that the CEMRI-based radiomics model showed higher AUCs and accuracies compared to previous studies, indicating it was able to optimally predict HCC differentiation. We speculate that our research findings may be attributed to the following two factors: 1) A large number (n = 5502) of quantitative radiomics signatures were extracted from the CEMRI images for each HCC nodule. Then, Spearman rank correlation, LASSO, mRMR approaches were employed to select the most robust signatures for developing the radiomics model. This comprehensive selection process possibly contributed to the higher AUC. 2) Comparison of eight ML classifiers, including regression, classification, and clustering algorithms. These classifiers were used to establish the most effective radiomics model, addressing both linear and nonlinear classification problems. In contrast, most previous studies solely employed linear models, leading to lower accuracy. The inclusion of diverse ML classifiers in our study enhanced the accuracy for diagnosing HCC differentiation. Therefore, our study suggests that a larger number of extracted radiomics signatures, combined with

**Table 3** MRI Features of Each HCC Nodule

| MRI Features | Training Cohort | | | Validation Cohort | | |
|---|---|---|---|---|---|---|
| | pHCC (n = 50) | npHCC (n = 160) | P value | pHCC (n = 12) | npHCC (n = 40) | P value |
| Tumor size (cm) | 7.14±3.83 | 4.01±2.25 | <0.001* | 6.12±4.23 | 3.73±2.35 | <0.001* |
| Nonsmooth margin | 34(68.0%)/16(32.0%) | 50(31.3%)/110(68.7%) | <0.001* | 7(58.3%)/5(41.7%) | 9(22.5%)/31(77.5%) | 0.045* |
| Irregular shape | 13(26.0%)/37(74.0%) | 57(35.6%)/103(64.4%) | 0.17 | 4(33.3%)/8(66.7%) | 11(27.5%)/29(72.5%) | 0.03* |
| NAPHE | 49(98.0%)/1(2.0%) | 156(97.5%)/4(2.5%) | 1.00 | 12(100.0%)/0(0.0%) | 38(95.0%)/2(5.0%) | 1.00 |
| Nonperipheral washout | 49(98.0%)/1(2.0%) | 150(93.8%)/10(6.2%) | 0.42 | 11(91.7%)/1(8.3%) | 38(95.0%)/2(5.0%) | 1.00 |
| Enhancing capsule | 46(92.0%)/4(8.0%) | 148(92.5%)/12(7.5%) | 1.00 | 3(25.0%)/9(75.0%) | 2(5.0%)/38(95.0%) | 0.13 |
| Nonenhancing capsule | 0(0.0%)/50(100.0%) | 6(3.8%)/154(96.2%) | 0.37 | 1(8.3%)/11(91.7%) | 0(0.0%)/40(100.0%) | 0.52 |
| Nodule-in-nodule | 5(10.0%)/45(90.0%) | 11(6.9%)/149(93.1%) | 0.67 | 0(0.0%)/12(100.0%) | 3(7.5%)/37(92.5%) | 0.79 |
| Mosaic architecture | 31(62.0%)/19(38.0%) | 49(30.6%)/111(69.4%) | <0.001* | 8(66.7%)/4(33.3%) | 10(25.0%)/30(75.0%) | 0.02* |
| Intratumoral blood products | 27(54.0%)/23(46.0%) | 40(25.0%)/120(75.0%) | <0.001* | 5(41.7.1%)/7(58.3%) | 8(20.0%)/32(80.0%) | 0.25 |
| Intratumoral fat content | 20(40.0%)/30(60.0%) | 44(27.5%)/116(72.5%) | 0.13 | 4(33.3%)/8(66.7%) | 6(15.0%)/34(85.0%) | 0.32 |
| Mild-moderate T2 hyperintensity | 50(100.0%)/0(0.0%) | 160(100.0%)/0(0.0%) | 1.0 | 12(100.0%)/0(0.0%) | 40(100.0%)/0(0.0%) | 1.0 |
| Corona enhancement | 31(62.0%)/19(38.0%) | 47(29.4%)/113(70.6%) | <0.001* | 8(66.7%)/4(33.3%) | 11(27.5%)/29(72.5%) | 0.03* |
| Iron sparing | 0(0.0%)/50(100.0%) | 1(0.6%)/159(99.4%) | 1.00 | 0(0.0%)/12(100.0%) | 0(0.0%)/40(100.0%) | 1.00 |
| Delayed phase hypointensity | 49(98.0%)/1(2.0%) | 156(97.5%)/4(2.5%) | 1.00 | 11(91.7%)/1(8.3%) | 39(97.5%)/1(2.5%) | 0.95 |

**Note**: *Represents P < 0.05.
**Abbreviation**: NAPHE, non-rim arterial phase hyperenhancement.
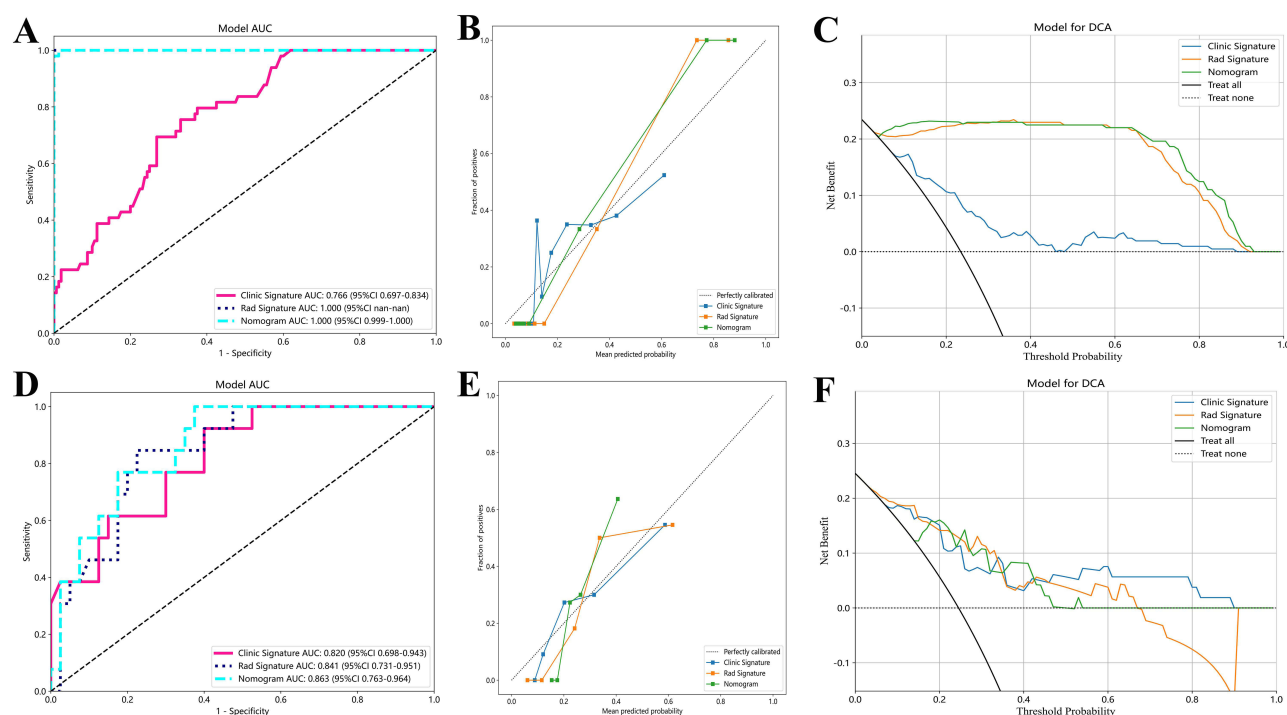
**Figure 4** The ROC curve, calibration curve, and DCA curve of three models in evaluating HCC differentiation in the training (**A**–**C**) and validation (**D**–**F**) cohorts, respectively. The radiomics model presented higher diagnostic accuracy, better goodness-of-fit and improved net clinical benefits in the training cohort.
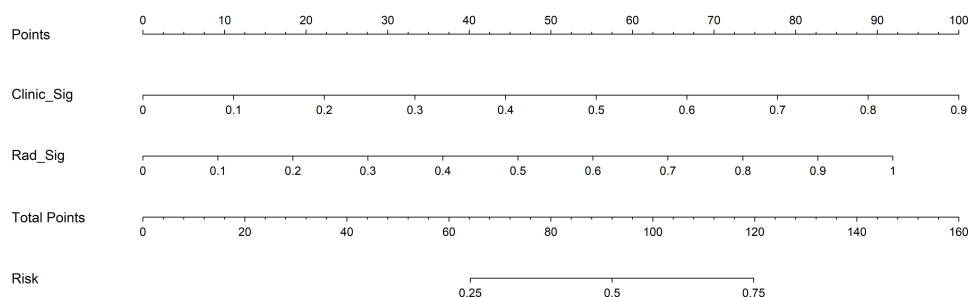


**Figure 5** The nomogram model for diagnosing HCC differentiation.

a variety of ML classifiers, can improve the accuracy of diagnosing HCC differentiation, which is useful for advancing the current understanding of radiomics model with ML classifiers.

Shinkawa et al[3] discovered that tumor size is a significant predictor related to the prognosis of recurrence and survival prognosis following HCC resection. This aligns with our study's conclusion that a larger tumor size was shown to have a significantly higher incidence of pHCC, revealing that the biological behavior may worsen into increased aggressiveness, infiltrative growth, and destruction of the peritumoral capsule with tumor size.[22] The presence of nonsmooth margins in tumors was characterized by extranodular growth or the multinodular type on imaging. Additionally, tumors with nonsmooth margins often exhibited destruction of the tumor capsule and invasion of adjacent liver parenchyma, thus reflecting aggressive biological behavior and providing support for our study's conclusion that nonsmooth margin is strongly associated with pHCC, which is consistent with the findings of Huang et al.[23]

The mosaic architecture feature in tumors reflects various characteristics such as tumor viability, fatty infiltration, hemorrhage, necrosis, or copper deposition.[24] It is typically composed of random compartments or internal nodules within a tumor, which differ in terms of enhancement, attenuation, intensity, shape, size, and are separated by fibrous tissues. In our study, we observed the mosaic architecture feature more frequently in pHCC, mainly because it is

associated with rapid disease progression, diverse biological heterogeneities, and a higher likelihood of vascular invasion. Furthermore, we identified corona enhancement, which refers to late AP or early PVP enhancement within the drainage vessels surrounding HCC. This enhancement is attributed to the venous drainage and indicates aggressive vascular invasion and intrahepatic metastases of HCC. Consistent with these findings, Liu et al[25] also identified corona enhancement as a risk factor for microvascular invasion in HCC.

Furthermore, the LI-RADS model provides valuable application for assessing pHCC and has proven to be fundamental in assisting radiologists in interpreting MRI naked-eyes features associated with pHCC. However, the radiomics model yielded significant higher AUCs of 1.00 for predicting pHCC, compared to LI-RADS model. This can be attributed to three following key factors: 1) HCC differentiation was evaluated based on the heterogeneous pathologies,[4,11] including mitotic count, tumor differentiation, and necrosis; thus, radiomics has the advantage of accurately predicting pHCC as it provides a more comprehensive assessment of tumor heterogeneity. 2) Radiomics signatures, which are not visible to the naked eye in MRI features, serve as quantitative markers for biomedical applications and present a crucial role in characterizing the microscopic behavior of HCC. 3) This study carefully selected eighteen optimal radiomics signatures and combined them to construct the most effective model using eight ML classifiers with different algorithms, this approach was instrumental in improving the diagnostic performance.

The integration of clinical and radiomics factors has led to the development of computational assisted nomogram models, which have emerged as highly effective and easy-to-understand alternatives for evidence-based management of cancer patients. In the study conducted by Xu et al[26], a nomogram was developed that combined serum markers and CT-based radiomics outperformed both the radiomics model alone and the clinical model, making it the superior option for predicting microvascular invasion and stratifying prognosis in HCC patients. Kong et al[8] found that a nomogram model integrating clinical indicators and the radiomics presented increased AUCs of 0.86 to 0.88, in predicting tumor response to HCC following TACE. In this study, our developed nomogram model outperformed both the single radiomics and LI-RADS models in evaluating pHCC, with an accuracy of 1.00 and an AUC of 1.00, suggesting its high precision and discriminatory ability. Furthermore, the decision curve analysis (DCA) curve showed improved net clinical benefits, these findings highlight the potential of our nomogram model in enhancing clinical decision-making and improving individualized diagnosis of HCC differentiation.

There were several shortcomings should be addressed in our study. 1) In this retrospective study, we excluded several suspected HCC patients without pathology, who may have been removed due to the potential threat of TACE or RFA. This patient selection criteria could introduce selection bias and undermine the reliability of our study's conclusions. Therefore, future prospective studies with larger cohorts from multiple centers are needed to validate the clinical application of our findings. 2) Considering that radiomics signatures were merely extracted from CEMRI images, it is possible that some significant signatures from DWI or T2WI might have been overlooked, through CEMRI images were deemed more representative of the actual heterogeneity of HCC.[7] 3) Moreover, this study only focused on exploring the relationship between HCC differentiation and the radiomics model. Therefore, future investigations examining the correlation between the radiomics model and HCC recurrence and prognosis will undoubtedly enhance the clinical applicability of our findings.

## Conclusion

In summary, our study revealed that CEMRI-based radiomics, specifically with the use of the XGBoost classifier, exhibited outstanding performance in evaluating HCC differentiation. More importantly, the integration of the radiomics model with the LI-RADS model in our nomogram model contributed to an optimal clinical decision tool for the individualized diagnosis of HCC differentiation.

## Abbreviations

ML, Machine learning; CEMRI, Contrast-enhanced magnetic resonance imaging; HCC, Hepatocellular carcinoma; ROC, Receiver operating characteristic curve; AUC, Area under the ROC curve; LI-RADS, liver-imaging reporting and data system; XGBoost, eXtreme Gradient Boosting; pHCC, poorly differentiated HCC; npHCC, non-poorly differentiated HCC; LASSO, least absolute shrinkage and selection operator; mRMR, maximum relevance minimum redundancy.

## Data Sharing Statement

The study data may be provided by contacting the corresponding author.

## Ethics Approval and Informed Consent

This retrospective study, conducted at the Third Affiliated Hospital of Soochow University, was approved by the Ethics Committee (approval number: 2022-CL027-01), in compliance with the Helsinki Declaration. Informed consent from patients was waived by the committee, considering the retrospective nature of the study and the utilization of anonymized data, while ensuring patient privacy and confidentiality. All patient data were handled in strict accordance with relevant privacy regulations, guidelines, and the principles outlined in the Declaration of Helsinki, ensuring confidentiality and anonymity throughout the study.

## Funding

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–249. doi:10.3322/caac.21660
2. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet Lond Engl.* 2018;391(10127):1301–1314. doi:10.1016/S0140-6736(18)30010-2
3. Shinkawa H, Tanaka S, Kabata D, et al. The prognostic impact of tumor differentiation on recurrence and survival after resection of hepatocellular carcinoma is dependent on tumor size. *Liver Cancer.* 2021;10(5):461–472. doi:10.1159/000517992
4. Zhou L, Rui JA, Zhou WX, et al. Edmondson-Steiner grade: a crucial predictor of recurrence and survival in hepatocellular carcinoma without microvascular invasio. *Pathol Res Pract.* 2017;213(7):824–830. doi:10.1016/j.prp.2017.03.002
5. Kleiner DE. Hepatocellular carcinoma: liver biopsy in the balance. *Hepatology.* 2018;68(1):13–15. doi:10.1002/hep.29831
6. Gong XQ, Tao YY, Wu Y, et al. Progress of MRI radiomics in hepatocellular carcinoma. *Front Oncol.* 2021;11:698373. doi:10.3389/fonc.2021.698373
7. Vietti Violi N, Lewis S, Liao J, et al. Gadoxetate-enhanced abbreviated MRI is highly accurate for hepatocellular carcinoma screening. *Eur Radiol.* 2020;30(11):6003–6013. doi:10.1007/s00330-020-07014-1
8. Kong C, Zhao Z, Chen W, et al. Prediction of tumor response via a pretreatment MRI radiomics-based nomogram in HCC treated with TACE. *Eur Radiol.* 2021;31(10):7500–7511. doi:10.1007/s00330-021-07910-0
9. Swanson K, Wu E, Zhang A, et al. From patterns to patients: advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell.* 2023;186(8):1772–1791. doi:10.1016/j.cell.2023.01.035
10. Moura Cunha G, Chernyak V, Fowler KJ, et al. Up-to-date role of CT/MRI LI-RADS in hepatocellular carcinoma. *J Hepatocell Carcinoma.* 2021;8:513–527. doi:10.2147/JHC.S268288
11. Loy LM, Low HM, Choi JY, et al. Variant hepatocellular carcinoma subtypes according to the 2019 WHO classification: an imaging-focused review. *Am J Roentgenol.* 2022;219(2):212–223. doi:10.2214/AJR.21.26982
12. Li H, Wu Z, Chen J, et al. External radiotherapy combined with sorafenib has better efficiency in unresectable hepatocellular carcinoma: a systematic review and meta-analysis. *Clin Exp Med.* 2023;23(5):1537–1549. doi:10.1007/s10238-022-00972-4
13. Su K, Shen Q, Tong J, et al. Construction and validation of a nomogram for HBV-related hepatocellular carcinoma: a large, multicenter study. *Ann Hepatol.* 2023;28(4):101109. doi:10.1016/j.aohep.2023.101109
14. Su K, Huang W, Xu K, et al. Evaluation of lactate dehydrogenase and alkaline phosphatase as predictive biomarkers in the prognosis of hepatocellular carcinoma and development of a new nomogram. *J Hepatocell Carcinoma.* 2023;10:69–79. doi:10.2147/JHC.S398632
15. Li H, Su K, Jiang Y, et al. Construction and validation of TACE therapeutic efficacy by ALR score and nomogram: a large, multicenter study. *J Hepatocell Carcinoma.* 2023;10:1009–1017. doi:10.2147/JHC.S414926
16. Castaldo A, De Lucia DR, Pontillo G, et al. State of the art in artificial intelligence and radiomics in hepatocellular carcinoma. *Diagnostics.* 2021;11(7):1194. doi:10.3390/diagnostics11071194
17. Han YE, Cho Y, Kim MJ, et al. Hepatocellular carcinoma pathologic grade prediction using radiomics and machine learning models of gadoxetic acid-enhanced MRI: a two-center study. *Abdom Radiol.* 2022;48(1):244–256. doi:10.1007/s00261-022-03679-y
18. Ameli S, Venkatesh BA, Shaghaghi M, et al. Role of MRI-derived radiomics features in determining degree of tumor differentiation of hepatocellular carcinoma. *Diagnostics.* 2022;12(10):2386. doi:10.3390/diagnostics12102386
19. Mao Y, Wang J, Zhu Y, et al. Gd-EOB-DTPA-enhanced MRI radiomic features for predicting histological grade of hepatocellular carcinoma. *Hepatobiliary Surg Nutr.* 2022;11(1):13–24. doi:10.21037/hbsn-19-870
20. Brancato V, Garbino N, Salvatore M, et al. MRI-based radiomic features help identify lesions and predict histopathological grade of hepatocellular carcinoma. *Diagnostics.* 2022;12(5):1085. doi:10.3390/diagnostics12051085
21. Yang X, Yuan C, Zhang Y, et al. Magnetic resonance radiomics signatures for predicting poorly differentiated hepatocellular carcinoma: a SQUIRE-compliant study. *Medicine.* 2021;100:19..

22. Usta S, Kayaalp C. Tumor diameter for hepatocellular carcinoma: why should size matter? *J Gastrointest Cancer.* 2020;51(4):1114–1117. doi:10.1007/s12029-020-00483-z

23. Huang K, Dong Z, Cai H, et al. Imaging biomarkers for well and moderate hepatocellular carcinoma: preoperative magnetic resonance image and histopathological correlation. *BMC Cancer.* 2019;19(1):364. doi:10.1186/s12885-019-5574-8

24. Kovac JD, Ivanovic A, Milovanovic T, et al. An overview of hepatocellular carcinoma with atypical enhancement pattern: spectrum of magnetic resonance imaging findings with pathologic correlation. *Radiol Oncol.* 2021;55(2):130–143. doi:10.2478/raon-2021-0004

25. Liu HF, Zhang YZZ, Wang Q, et al. A nomogram model integrating LI-RADS features and radiomics based on contrast-enhanced magnetic resonance imaging for predicting microvascular invasion in hepatocellular carcinoma falling the Milan criteria. *Transl Oncol.* 2023;27:101597. doi:10.1016/j.tranon.2022.101597

26. Xu L, Li L, Wang P, et al. Novel prognostic nomograms for hepatocellular carcinoma patients with microvascular invasion: experience from a single center. *Gut Liver.* 2019;13(6):669–682. doi:10.5009/gnl18489