# RESEARCH



# Gadoxetic acid-enhanced MRI for identifying cholangiocyte phenotype hepatocellular carcinoma by interpretable machine learning: individual application of SHAP

Wei Liu<sup>1+</sup>, Zhiping Cai<sup>1+</sup>, Yifan Chen<sup>1+</sup>, Xingqun Guan<sup>1</sup>, Jieying Feng<sup>2</sup>, Haixiong Chen<sup>1</sup>, Baoliang Guo<sup>1</sup>, Fusheng OuYang<sup>1</sup>, Chun Luo<sup>3</sup>, Rong Zhang<sup>1</sup>, Xinjie Chen<sup>1</sup>, Xiaohong Li<sup>1</sup>, Cuiru Zhou<sup>1</sup>, Shaomin Yang<sup>4\*†</sup>, Ziwei Liu<sup>1\*†</sup> and Qiugen Hu<sup>1\*†</sup>

# Abstract

**Purpose** Cholangiocyte phenotype hepatocellular carcinoma (HCC) is highly invasive. This study aims to develop and validate an optimal machine learning model to predict cholangiocyte phenotype HCC based on T1 mapping gadoxetic acid-enhanced MRI and to implement individual applications via the Shapley Additive explanation (SHAP).

**Methods** We included 180 patients with histologically confirmed HCC from two institutions. Clinical and MRI imaging features were screened for predicting cholangiocyte phenotype hepatocellular carcinoma using Least Absolute Shrinkage and Selection Operator (LASSO) and the logistic regression analysis. Five machine learning models were constructed based on these features. A Kaplan–Meier survival analysis aims to compare prognostic differences between cholangiocyte phenotype-positive HCC groups and classical (cholangiocyte phenotype-negative) HCC groups, and was conducted to explore the prognostic information of the optimal model.

**Results** The most significant clinicoradiological features, including the platelet-to-lymphocyte ratio (PLR), tumor capsule, target sign on hepatobiliary phase (HBP), and T1 relaxation time of 20 min (T1rt-20 min), were selected to construct the prediction model. Finally, we selected the eXtreme Gradient Boosting (XGBoost) model as the optimal predictive model, which achieved AUCs of 0.835, 0.830, 0.816 and 0.776 in training, internal validation, external

 $^{\rm t}\mbox{Wei}$  Liu, Zhiping Cai and Yifan Chen contributed equally to this work as first author.

<sup>+</sup>Shaomin Yang, Ziwei Liu and Qiugen Hu contributed equally to this work.

\*Correspondence: Shaomin Yang ysmsin@aliyun.com Ziwei Liu Izw13590605816@163.com Qiugen Hu hu6009@163.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

validation, and prospective validation cohorts, respectively, for visual analysis via SHAP, in which T1rt-20 min made a significant contribution. Survival analysis showed a statistically significant difference in relapse-free survival (RFS) between cholangiocyte phenotype-positive HCC groups and classical HCC groups from institution I (hazard ratio [HR] 1.994; 95% CI, 1.059–3.758; P=0.027), and the construction XGBoost model can be used to stratify RFS according to prognosis (HR, 1.986; 95% CI, 1.061–3.717; P=0.029).

**Conclusion** The machine learning model utilizing T1 mapping gadoxetic acid-enhanced MRI demonstrates significant potential in identifying cholangiocyte phenotype HCC. Furthermore, personalized prediction is enhanced through the application of SHAP, providing valuable insights to support clinical decision-making processes.

Keywords Cholangiocyte phenotype, Hepatocellular carcinoma, Magnetic resonance imaging, Machine learning

# Introduction

Hepatocellular carcinoma (HCC) is a common malignancy worldwide. Complex pathological phenotypes and tumor heterogeneity are the main causes of poor prognosis in patients with HCC [1, 2]. HCC expressing cytokeratin (CK) 19 or CK7 are considered cholangiocyte phenotypes with highly aggressive behavior. Compared to classical HCC, cholangiocyte phenotype HCC exhibits higher lymph node metastasis rate, and higher risk of vascular invasion, which are closely related to poor prognosis [3-5]. Therefore, the preoperative identifying cholangiocyte phenotype HCC provides a basis for the selection of surgical strategies and formulation of adjuvant treatments, which is important for the evaluation of recurrence and prognosis. However, invasive biopsy may result in complications. Moreover, owing to tumor heterogeneity, the degree of marker expression in different biopsy regions may be underestimated or overestimated [6]. Therefore, it is necessary to develop a non-invasive method to identifying cholangiocyte phenotype HCC.

T1 mapping, a non-invasive and quantitative method for analyzing T1 values of tissues, is performed by fitting a series of images collected at different time points during the T1 relaxation time recovery process [7]. The T1 relaxation time measured on T1 mapping can reflect the intrinsic characteristics of the organization and is not affected by the scanning sequence parameters [8]. It can be combined with gadoxetic acid-enhanced MRI efficiently to accurately and objectively reflect changes in liver cells absorbing gadoxetic acid, thereby reflecting the biological characteristics of HCC [9]. Previous studies have evaluated the pathological differentiation [10], microvascular invasion [11], and resection recurrence of HCC [12]. However, there is currently no research has used gadoxetic acid-enhanced MRI combined with T1 mapping to predict cholangiocyte phenotype HCC.

Machine learning has great potential in medical research and prognostic model construction. Due to the varying logic and complexity of different machine learning algorithms, their model results are often difficult to interpret, leading to the "black-box" problem [13]. Previous studies have focused on the accuracy of model prediction, neglecting their interpretability and limiting their clinical application; therefore, interpretable machine learning algorithms have become the current focus of research [14]. SHapley Additive exPlanation (SHAP) originated from the cooperative game theory. It can explain the "black-box" model at the global and local levels and interpret the predicted value of the model as the sum of the contribution values of each input feature, that is, the Shapley value [15]. Machine learning combined with SHAP can provide an explicit explanation of individualized prediction and provide physicians with an intuitive understanding of the influence of key features in the model [16, 17]. To the best of our knowledge, only a few studies have focused on the interpretable machine learning prediction of cholangiocyte phenotype HCC using SHAP.

Therefore, this study aimed to develop and validate an interpretable machine-learning model based on clinicoradiological features to identifying cholangiocyte phenotype HCC. SHAP was used to intuitively explain the predictive results by comparing the predictive performance of the five machine learning models to ultimately determine the optimal model, guiding the clinical development of personalized diagnosis and treatment plans.

# Materials and methods

# Participants

This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Medical Ethics Committee of Shunde Hospital, Southern Medical University. The requirement to obtain informed consent was waived because of the observational design of the study. Preoperative T1 mapping gadoxetic acidenhanced MRI and clinical data were retrospectively collected from two institutions between January 2019 and May 2022. Data from Institution I were designated as the training and internal validation cohort, while data from Institution II served as the external validation cohort. Additionally, patients from Institution I were prospectively enrolled from June 2022 to December 2022 to form the prospective validation cohort. The training and internal validation cohorts were utilized as retrospective datasets, whereas the external validation cohort also comprised retrospective data. The prospective validation cohort, however, was derived from prospectively collected data. The data were reviewed on June 1, 2023.

The inclusion criteria were as follows: (a) patients with pathologically confirmed HCC and CK19 status; (b) underwent curative hepatic resection; (c) those with preoperative T1 mapping gadoxetic acid-enhanced MRI imaging. The exclusion criteria were as follows: (a) patients who did not receive curative resection; (b) those with a lack of complete clinical data or whose MRI and pathological images were unavailable; (c) those who received previous treatment; (d) those with imaging data of poor quality with obvious artifacts; and (e) those who underwent MRI examination more than one month before surgery. The patient recruitment process and study design are shown in Fig. 1.

### **Clinical data collection**

Clinical and laboratory data of the patients were recorded, including sex, age, hepatitis, levels of alpha-fetoprotein (AFP,  $\mu$ g/L), alanine aminotransferase (ALT, U/L), aspartate aminotransferase (AST, U/L), and gamma-glutamyltransferase (GGT, U/L), neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte

ratio (PLR), Child-Pugh classification, and the modus operandi.

### **MRI** examination

All patients from institutions I and II were underwent scanning using the Magnetom Skyra, Lumina, or Verio 3.0T MRI system (Siemens Healthcare Sector, Erlangen, Germany) equipped with a dedicated abdominal coil. All patients fasted for more than 6 h and underwent breathing training before the scan. The patients were instructed to adopt a head-advanced supine position, with examinations ranging from the upper edge to the lower edge of the liver. The standard imaging protocol consisted of T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and T1 mapping. Gadoxetic acid was used for enhanced MRI (Primovist; Bayer Schering Pharma AG, Berlin, Germany, 0.1 mmol/kg), the flow rate was set at 1.0 mL/s, and 30 mL physiological saline was then used for rinsing. Multiphase enhancements, including the arterial, portal, equilibrium, and hepatobiliary phases (HBP), were obtained at 20-30 s, 60-90 s, 150-180 s, and 20 min, respectively. T1 mapping included pre-enhancement and enhancement after 20 min. The specific scanning parameters are listed in Supplement Table 1.



Fig. 1 Patient recruitment and study design

### **MRI** feature analysis

The MRI qualitative features were independently evaluated by two abdominal radiologists (with 5 and 10 years of experience, respectively) from institution I who were blinded to the clinical and pathological information. In case of conflicting opinions, a third senior abdominal radiologist (with 15 years of experience) would participate in the discussion for a consensus. In patients with multiple tumors, the largest tumor was analyzed. The following MRI image features were evaluated (Supplement Method 1 and Supplement Fig. 1): (1) tumor margin; (2) tumor capsule; (3) cystic or necrosis portion; (4) fat deposition; (5) signal intensity on T2WI; (6) hemorrhage; (7) target sign on DWI; (8) target sign on HBP; (9) arterial rim enhancement; (10) peritumoral enhancement; (11) peritumoral hypointensity; and (12) satellite nodules.

Quantitative MRI features were measured by an abdominal radiologist (with 5 years of experience) from institution I using the RadiAnt DICOM Viewer 2022.1.1 software (https://www.radiantviewer.com). Tumor size was defined as the maximum tumor diameter on the axial or coronal HBP. The tumor region of interest (ROI) was then drawn to select the largest slice of the tumor on the apparent diffusion coefficient (ADC) and T1 mapping to avoid blood vessels. The ROI was set as the maximum cross-sectional area of the tumor. Another abdominal radiologist (with 10 years of experience) from institution I edited and confirmed the ROIs. The average of the results measured by the two radiologists was used as the final value, and the interclass correlation coefficient (ICC) was evaluated. The ADC value, T1 relaxation time of pre-enhancement (T1rt-pre), and T1 relaxation time of 20 min after enhancement (T1rt-20 min) were recorded, and the reduction rate of T1 relaxation time (rrT1rt) was calculated using the following formula: rrT1rt = (T1rtpre - T1rt-20 min) / T1rt-pre.

### Histopathological examination

All histopathological examinations were conducted by two pathologists (with over 5 and 10 years of experience in liver pathology) who were blinded to the clinical and imaging information. If all of the following criteria are met, the pathological diagnosis is HCC with cholangiocyte phenotype. (1) microscopic morphological features of HCC; (2) positive expression of hepatocyte paraffin antigen 1 (HepPar-1), glypican-3 (GPC-3), or glutamine synthetase (GS) in tumor cells; and (3) positive expression of CK19 in tumor cells ( $\geq$  15%) [3]. All patients were divided into two groups: cholangiocyte phenotype-positive HCC groups and classical (cholangiocyte phenotypenegative) HCC groups. The detailed measurements of CK19 are described in the Supplement Method 2.

### Model development and validation

The clinical and MRI features were screened using the least absolute shrinkage and selection operator (LASSO), and the most significant clinicoradiological features were obtained through logistic regression analysis with stepwise selection. A five-fold stratified cross-validation was performed. Four groups, accounting for 90% of the total sample, formed the training cohort for model construction, whereas the remaining patients formed the internal validation cohort to evaluate model performance. Stratified sampling ensured that the distribution of the patients in the two cohorts was similar to the total sample to reduce systematic errors caused by the division of datasets.

Five machine learning models were constructed based on the most significant clinicoradiological features: random forest (RF), K-nearest neighbor (KNN), support vector machine (SVM), eXtreme Gradient Boosting (XGBoost), and logistic regression (LR). The entire process was repeated 100 times via bootstrapping to ensure model stability; one round of cross-validation is shown in Supplement Fig. 2. One external validation cohort and one prospective cohort were used to validate the prediction performance.

### **Explanation of the SHAP algorithm**

Through a comprehensive comparison, an optimal machine-learning model was obtained, and the model results were visualized using SHAP. SHAP is a game-theoretic approach for interpreting machine learning model predictions by quantifying the contribution of each feature to the final output. The Shapley algorithm calculates the Shapley value of each variable in the training cohort, explains the relationship between the input variables and the output results of the model, and improves model interpretability. The Shapley value of the input variable reflects the contributing weight of the feature in the model, and the red and blue colors represent positive and negative effects, respectively. A schematic of the machine learning analysis algorithm for the entire process is shown in Fig. 2. The detailed demonstration process and the formulas are provided in the Supplement explanation.

### Follow-up

From institution I, Two patients who were lost to followup, and the remaining 115 patients were followed up by ultrasonography, CT, or MRI every 3–6 months after surgery until recurrence or final review of the data. The Relapse-free survival (RFS) was defined as intrahepatic recurrence or distant metastasis, including residual liver lesions, and organ, lymph node, and peritoneal metastasis. The RFS rates were recorded. Then, XGBoost models were constructed based on the data of the training and



Fig. 2 Algorithm schematic of machine learning analysis for the entire process. The most significant features were screened by LASSO and logistic regression analysis. Five different machine learning models were constructed by these features. To obtain the best prediction model, the prediction performance of the models was compared by the mean ROC, and the SHAP was used to analyze the diagnostic process of the best model

internal validation cohorts. Patients were divided into high-risk and low-risk groups based on the best cut-off of the XGBoost predicted values.

### Statistical analysis

The quantitative data are expressed as means ± standard deviations or medians (minimum, maximum range). Qualitative data are expressed as numbers and percentages. The ICC was used to evaluate the consistency of measurements between the two radiologists (ICC>0.75: good, 0.65–0.75: general, and ICC < 0.65: poor). The prediction efficiency of each model was evaluated using the area under the curve (AUC), accuracy, sensitivity, and specificity values obtained from the mean receiver operating characteristic curve. The DeLong test was used to compare the differences in AUCs among the different models. The confidence intervals (CIs) in this study were set to 95%, and P < 0.05 indicated statistically significant differences. Survival curves were drawn using the Kaplan Meier method and compared using the log rank test. All statistical analyses were performed using the SPSS (version 25.0) or R (version 3.6.1; http://www.rproject. org) software.

# Results

## **Clinical data and MRI features**

The clinical data and CT image features of the participants are shown in Tables 1 and 2. In the training and internal validation cohorts, 117 patients with HCC (106 males, mean age  $59.0 \pm 10.9$  years) were included: 40 cholangiocyte phenotype-positive HCC patients and 77

classical HCC patients. In the external validation cohort, 33 patients with HCC (31 males, mean age  $64.8 \pm 10.3$ years) were included: 12 cholangiocyte phenotypepositive HCC patients and 21 classical HCC patients patients. Additionally, in the prospective cohort, 30 patients with HCC (26 males, mean age  $59.0 \pm 10.9$  years) were included: 9 cholangiocyte phenotype-positive HCC patients and 21 classical HCC patients patients. The ICC values for tumor size, ADC, T1rt-pre, and T1rt-20 min measured by the two radiologists were shown in Supplement Table 2.

### Model development and evaluation

The most significant clinicoradiological features were obtained using LASSO and stepwise-selection logistic regression analyses. Finally, the PLR, tumor capsule, target sign on HBP, and T1rt-20 min were selected to construct the prediction model. The cut-off value was 797 ms of T1rt-20 min.

The prediction efficiencies of the RF, KNN, SVM, XGBoost, and LR models are listed in Table 3. Comparison of prediction performance (AUCs) of machine learning models, XGBoost performed slightly better than others in the cohorts, as shown in Supplement Table 3, which achieved AUCs of 0.835, 0.830, 0.816 and 0.776 in training, internal validation, external validation, and prospective validation cohorts, respectively. Finally, we chose the XGBoost model as the optimal model.

| Characteristics      | Training and internal validation<br>Cohort ( <i>n</i> = 117) |  | External validation cohort<br>(n=33)     |  | Prospective validation cohort<br>(n = 30) |   |
|----------------------|--|--|--|--|---|---|
|                      | cholangiocyte<br>phenotype (-)<br>(N=77)                     | cholangiocyte<br>phenotype (+)<br>(N=40) | cholangiocyte<br>phenotype (-)<br>(N=21) | cholangiocyte<br>phenotype (+)<br>(N=12) | cholangiocyte<br>phenotype (-)<br>(N=21)  | cholangio-<br>cyte pheno-<br>type (+)<br>( <i>N</i> =9) |
| Sex                  |  |  |  |  |   |   |
| Female               | 7 (9.1%)   | 4 (10.0%)                                | 1 (4.8%)                                 | 1 (8.3%)                                 | 3 (14.3%)                                 | 1 (11.1%)   |
| Male                 | 70 (90.9%)   | 36 (90.0%)                               | 20 (95.2%)                               | 11 (91.7%)                               | 18 (85.7%)                                | 8 (88.9%)   |
| Age                  |  |  |  |  |   |   |
| $Mean \pm SD$        | $57.3 \pm 10.2$  | $55.4 \pm 11.1$                          | $50.7 \pm 8.08$                          | $50.8 \pm 12.6$                          | $61.2 \pm 14.6$                           | $59.3 \pm 12.2$   |
| Hepatitis            |  |  |  |  |   |   |
| Absent               | 12 (15.6%)   | 6 (15.0%)                                | 3 (14.3%)                                | 2 (16.7%)                                | 4 (19.0%)                                 | 3 (33.3%)   |
| Present              | 65 (84.4%)   | 34 (85.0%)                               | 18 (85.7%)                               | 10 (83.3%)                               | 17 (81.0%)                                | 6 (66.7%)   |
| AFP (ug/L)           |  |  |  |  |   |   |
| ≤20                  | 45 (58.4%)   | 12 (30.0%)                               | 12 (57.1%)                               | 4 (33.3%)                                | 13 (61.9%)                                | 3 (33.3%)   |
| >20                  | 32 (41.6%)   | 28 (70.0%)                               | 9 (42.9%)                                | 8 (66.7%)                                | 8 (38.1%)                                 | 6 (66.7%)   |
| ALT (U/L)            |  |  |  |  |   |   |
| Median<br>(Min, Max) | 31.0 (10.0, 241)   | 37.0 (14.0, 241)                         | 34.0 (10.0, 259)                         | 34.0 (19.0, 241)                         | 26.0 (8.00, 134)                          | 24.0 (8.00,<br>237)                                     |
| AST (U/L)            |  |  |  |  |   |   |
| Median<br>(Min, Max) | 36.0 (14.0, 169)   | 41.0 (18.0, 287)                         | 36.0 (14.0, 202)                         | 43.0 (21.0, 287)                         | 33.0 (21.0, 73.0)                         | 44.0 (21.0,<br>181)                                     |
| GGT (U/L)            |  |  |  |  |   |   |
| Median<br>(Min, Max) | 49.0 (12.0, 1060)  | 85.0 (13.0, 735)                         | 71.0 (20.0, 611)                         | 92.5 (21.0, 735)                         | 65.0 (16.0, 366)                          | 50.0 (31.0,<br>449)                                     |
| NLR                  |  |  |  |  |   |   |
| Median<br>(Min, Max) | 1.87 (0.480, 7.30)   | 2.62 (0.74, 14.3)                        | 2.00 (1.12, 5.42)                        | 3.65 (0.740, 14.3)                       | 2.77 (1.03, 49.7)                         | 2.30 (1.49,<br>9.57)                                    |
| PLR                  |  |  |  |  |   |   |
| Median<br>(Min, Max) | 103 (44.7, 251)  | 120 (31.7, 1650)                         | 105 (60.7, 236)                          | 130 (47.8, 1650)                         | 96.2 (64.8, 222)                          | 111 (49.1,<br>228)                                      |
| Child-Pugh           |  |  |  |  |   |   |
| А                    | 66 (85.7%)   | 33 (82.5%)                               | 17 (81.0%)                               | 9 (75.0%)                                | 20 (95.2%)                                | 6 (66.7%)   |
| В                    | 11 (14.3%)   | 7 (17.5%)                                | 4 (19.0%)                                | 3 (25.0%)                                | 1 (4.8%)                                  | 3 (33.3%)   |
| Modus operandi       |  |  |  |  |   |   |
| Non-anatomic         | 33   | 14                                       | 9  | 3  | 9   | 3   |
| Anatomic             | 44   | 26                                       | 12                                       | 9  | 12  | 6   |

# Table 1 Clinical data in different cohort

Notes: Continuous variables with normal distribution are presented as mean ± standard deviation (SD) and those with abnormal distribution as median (minimum, maximum range). Categorical variables are presented as N (%). AFP, alpha-fetoprotein, ALT, alanine aminotransferase, AST, aspartate aminotransferase, GGT, gamma-glutamyl transferase, NLR, neutrophil-to-lymphocyte ratio, PLR, platelet to lymphocyte ratio

# **Clinical application by SHAP**

We calculated the overall and individual Shapley values of the XGBoost model, which can be helpful for its interpretation and clinical application. For overall prediction, the SHAP bar graph (Fig. 3A) shows the degree of influence of the four most significant features on the final predicted probability, and the absolute average Shapley values are 0.0431, 0.0445, 0.0523, and 0.0704, respectively, among which T1rt-20 min has the greatest impact. The SHAP scatterplot (Fig. 3B) shows the positive and negative effects of each feature on the prediction probability using different colors. In predicting the probability of cholangiocyte phenotype-positive HCC, PLR, target sign on HBP, and T1rt-20 min had a positive effect, whereas the tumor capsule had a negative effect. The results of the SHAP summary effort plot (Fig. 3C) show the positive and negative effects of each feature on predictive probability in all cases. The SHAP decision plot (Fig. 3D) shows the transition path of a clinical sample's prediction from the baseline to the final outcome, allowing clinicians to identify diagnostic outliers or critical decision boundaries. For individual predictions, Figs. 4A-F show six examples of accurately predictions. The SHAP effort plot shows both the positive and negative effects of each feature on predictive outcomes in a single case. The base value represents the basic prediction probability of the XGBoost model and f (x) represents the final prediction probability of the optimal model.

| Characteristics            | Training cohort and internal valida-<br>tion Cohort ( <i>n</i> = 117) |  | External validation cohort<br>(n = 33)   |  | Prospective validation cohort (n = 30)   |  |
|----------------------------|---|--|--|--|--|--|
|                            | cholangiocyte<br>phenotype (-)<br>(N=77)                              | cholangiocyte<br>phenotype (+)<br>(N=40) | cholangiocyte<br>phenotype (-)<br>(N=21) | cholangiocyte<br>phenotype (+)<br>(N=12) | cholangiocyte<br>phenotype (-)<br>(N=21) | cholangio-<br>cyte pheno-<br>type (+)<br>(N=9) |
| Tumor margin               |   |  |  |  |  |  |
| Absent                     | 35 (45.5%)  | 8 (20.0%)                                | 7 (33.3%)                                | 2 (16.7%)                                | 10 (47.6%)                               | 1 (11.1%)                                      |
| Present                    | 42 (54.5%)  | 32 (80.0%)                               | 14 (66.7%)                               | 10 (83.3%)                               | 11 (52.4%)                               | 8 (88.9%)                                      |
| Tumor capsule              |   |  |  |  |  |  |
| Absent                     | 31 (40.3%)  | 5 (12.5%)                                | 5 (23.8%)                                | 1 (8.3%)                                 | 11 (52.4%)                               | 3 (33,3%)                                      |
| Present                    | 46 (59.7%)  | 35 (87.5%)                               | 16 (76.2%)                               | 11 (91.7%)                               | 10 (47.6%)                               | 6 (66,7%)                                      |
| Cystic or necrosis portion |   |  |  |  |  |  |
| Absent                     | 52 (67 5%)  | 18 (45 0%)                               | 8 (38 1%)                                | 4 (33 3%)                                | 12 (57 1%)                               | 6 (66 7%)                                      |
| Present                    | 25 (32 5%)  | 22 (55 0%)                               | 13 (61 9%)                               | 8 (66 7%)                                | 9 (42 9%)                                | 3 (33 3%)                                      |
| Fat deposition             | 25 (32.570)   | 22 (33.676)                              | 13 (01.570)                              | 0 (00.770)                               | 5 (12.576)                               | 5 (55.570)                                     |
| Absent                     | 66 (85 7%)  | 34 (85 0%)                               | 18 (85 7%)                               | 10 (83 3%)                               | 18 (85 7%)                               | 8 (88 9%)                                      |
| Present                    | 11 (1/ 3%)  | 6 (15 0%)                                | 3 (1/ 3%)                                | 2 (16 7%)                                | 3 (1/ 3%)                                | 1 (11 1%)                                      |
| Signal intensity on T2W/   | 11 (14.570)   | 0(13.070)                                | 5 (14.570)                               | 2 (10.770)                               | 5 (14.570)                               | 1 (11.170)                                     |
| Abcont                     | 30 (50 6%)  | 15 (27 50%)                              | 4 (10,0%)                                | 1 (33 30%)                               | 10 (47 6%)                               | 5 (55 606)                                     |
| Procont                    | 39 (30.0%)  | 75 (57.5%)                               | 4 (19.0%)                                | 4 (55.5%)<br>8 (66 7%)                   | 10 (47.0%)                               | J (JJ.070)                                     |
| Hemerrhage                 | 38 (49.470)   | 23 (02.3%)                               | 17 (01.0%)                               | 8 (00.7 %)                               | 11 (32.470)                              | 4 (44.470)                                     |
| Abcont                     | 56 (72 70%)   | 20 (75 004)                              | 12 (57 104)                              | 0 (75 00%)                               | 15 (71 404)                              | 7 (77 00%)                                     |
| Absent                     | 20 (72.7%)<br>21 (27.2%)  | 50 (75.0%)<br>10 (25.0%)                 | 12 (37.1%)                               | 9 (75.0%)<br>2 (25.0%)                   | 15 (71.4%)                               | 7 (77.0%)                                      |
| Fresent                    | 21 (27.5%)  | 10 (23.0%)                               | 9 (42.9%)                                | 5 (25.0%)                                | 0 (28.0%)                                | 2 (22.2%)                                      |
| larget sign on Dvvi        | F7 (74 00()   | 24 (60.00()                              |  | 0 (75 00()                               | 17 (01 00()                              | 0 (1000()                                      |
| Absent                     | 57 (74.0%)  | 24 (60.0%)                               | 15 (71.4%)                               | 9 (75.0%)                                | 17 (81.0%)                               | 9 (100%)                                       |
| Present                    | 20 (26.0%)  | 16 (40.0%)                               | 6 (28.6%)                                | 3 (25.0%)                                | 4 (19.0%)                                | 0 (0%)   |
| larget sign on HBP         | (00, (00)   | 07 (67 50)                               | 24 (4 2 2 2 ()                           | 0 (75 00)                                | 10 (05 70))                              | 2 (22 24)                                      |
| Absent                     | 69 (89.6%)  | 27 (67.5%)                               | 21 (100%)                                | 9 (75.0%)                                | 18 (85./%)                               | 3 (33.3%)                                      |
| Present                    | 8 (10.4%)   | 13 (32.5%)                               | 0 (0%)                                   | 3 (25.0%)                                | 3 (14.3%)                                | 6 (66./%)                                      |
| Arterial rim enhancement   |   |  |  |  |  |  |
| Absent                     | 43 (55.8%)  | 14 (35.0%)                               | 6 (28.6%)                                | 2 (16.7%)                                | 10 (47.6%)                               | 4 (44.4%)                                      |
| Present                    | 34 (44.2%)  | 26 (65.0%)                               | 15 (71.4%)                               | 10 (83.3%)                               | 11 (52.4%)                               | 5 (55.6%)                                      |
| Peritumoral enhancement    |   |  |  |  |  |  |
| Absent                     | 49 (63.6%)  | 17 (42.5%)                               | 9 (42.9%)                                | 3 (25.0%)                                | 12 (57.1%)                               | 3 (33.3%)                                      |
| Present                    | 28 (36.4%)  | 23 (57.5%)                               | 12 (57.1%)                               | 9 (75.0%)                                | 9 (42.9%)                                | 6 (66.7%)                                      |
| Peritumoral hypointensity  |   |  |  |  |  |  |
| Absent                     | 52 (67.5%)  | 17 (42.5%)                               | 11 (52.4%)                               | 4 (33.3%)                                | 16 (76.2%)                               | 5 (55.6%)                                      |
| Present                    | 25 (32.5%)  | 23 (57.5%)                               | 10 (47.6%)                               | 8 (66.7%)                                | 5 (23.8%)                                | 4 (44.4%)                                      |
| Satellite nodules          |   |  |  |  |  |  |
| Absent                     | 59 (76.6%)  | 21 (52.5%)                               | 15 (71.4%)                               | 6 (50.0%)                                | 20 (95.2%)                               | 6 (66.7%)                                      |
| Present                    | 18 (23.4%)  | 19 (47.5%)                               | 6 (28.6%)                                | 6 (50.0%)                                | 1 (4.8%)                                 | 3 (33.3%)                                      |
| Tumor size                 |   |  |  |  |  |  |
| Median (Min, Max)          | 36.0 (11.9, 117)  | 39.9 (14.1, 132)                         | 62.3 (17.7, 158)                         | 57.2 (17.2, 125)                         | 40.0 (11.2, 120)                         | 34.3 (26.2,<br>89.9)                           |
| ADC                        |   |  |  |  |  |  |
| Median (Min, Max)          | 1030 (462, 2140)  | 956 (628, 1910)                          | 904 (656, 1550)                          | 957 (798, 1530)                          | 845 (674, 1410)                          | 892 (752,<br>1200)                             |
| T1rtpre                    |   |  |  |  |  | -  |
| Mean±SD<br>T1rt20min       | 1310±250  | 1440±275                                 | 1500±288                                 | $1550 \pm 345$                           | 1280±264                                 | 1430±229                                       |
| Median (Min, Max)          | 790 (420, 1260)   | 847 (495, 1410)                          | 877 (630, 1440)                          | 969 (759, 1410)                          | 747 (340, 991)                           | 812 (550,<br>1410)                             |

# Table 2 MRI image features in different cohort

| Characteristics | Training cohort and internal valida-<br>tion Cohort (n = 117) |  | External validation cohort<br>(n=33)     |  | Prospective validation cohort<br>(n=30)  |  |
|-----------------|---|--|--|--|--|--|
|                 | cholangiocyte<br>phenotype (-)<br>(N=77)                      | cholangiocyte<br>phenotype (+)<br>(N=40) | cholangiocyte<br>phenotype (-)<br>(N=21) | cholangiocyte<br>phenotype (+)<br>(N=12) | cholangiocyte<br>phenotype (-)<br>(N=21) | cholangio-<br>cyte pheno-<br>type (+)<br>(N=9) |
| rrT1rt          |   |  |  |  |  |  |
| Mean±SD         | $0.399 \pm 0.128$   | 0.389±0.120                              | $0.369 \pm 0.126$                        | 0.334±0.113                              | 0.437±0.144                              | $0.396 \pm 0.162$                              |

### Table 2 (continued)

Notes: Continuous variables with normal distribution are presented as mean ± standard deviation (SD) and those with abnormal distribution as median (m maximum range). DWI, diffusion-weighted imaging, HBP, hepatobiliary phase, T1rt, T1 relaxation time, rrT1rt, the reduction rate of T1 relaxation time

| Table 3 | Prediction | performance | of machine | learning | models |
|---------|------------|-------------|------------|----------|--------|
|---------|------------|-------------|------------|----------|--------|

|                            | AUC (95% CI)       | Accuracy (95% CI)  | Sensitivity (95% CI) | Specificity (95% CI) |
|----------------------------|--------------------|--------------------|----------------------|----------------------|
| RF                         |                    |                    |                      |                      |
| Training cohort            | 0.808(0.806-0.810) | 0.774(0.771-0.777) | 0.700(0.693-0.707)   | 0.812(0.805-0.819)   |
| Internal validation cohort | 0.709(0.699–0.719) | 0.739(0.730-0.748) | 0.698(0.681-0.715)   | 0.760(0.742-0.778)   |
| External validation cohort | 0.829(0.827-0.831) | 0.804(0.799–0.808) | 0.712(0.700-0.723)   | 0.856(0.844-0.868)   |
| Prospective validation     | 0.813(0.810-0.815) | 0.831(0.828-0.834) | 0.762(0.757-0.767)   | 0.860(0.855–0.865)   |
| KNN                        |                    |                    |                      |                      |
| Training cohort            | 0.834(0.830-0.837) | 0.798(0.794–0.801) | 0.719(0.711-0.727)   | 0.839(0.832-0.845)   |
| Internal validation cohort | 0.709(0.699–0.720) | 0.772(0.763-0.780) | 0.654(0.638-0.670)   | 0.839(0.824–0.853)   |
| External validation cohort | 0.859(0.856-0.862) | 0.846(0.843-0.848) | 0.744(0.734-0.753)   | 0.904(0.897-0.911)   |
| Prospective validation     | 0.754(0.750-0.758) | 0.820(0.817-0.822) | 0.653(0.644-0.662)   | 0.891(0.885–0.898)   |
| SVM                        |                    |                    |                      |                      |
| Training cohort            | 0.720(0.694-0.746) | 0.814(0.807-0.821) | 0.657(0.628-0.686)   | 0.895(0.889-0.902)   |
| Internal validation cohort | 0.609(0.595-0.622) | 0.751(0.744–0.758) | 0.522(0.503-0.542)   | 0.870(0.857–0.883)   |
| External validation cohort | 0.728(0.702-0.753) | 0.797(0.788–0.805) | 0.681(0.651-0.711)   | 0.863(0.851-0.875)   |
| Prospective validation     | 0.690(0.667-0.713) | 0.752(0.739–0.766) | 0.713(0.690-0.735)   | 0.769(0.746-0.793)   |
| XGBoost                    |                    |                    |                      |                      |
| Training cohort            | 0.835(0.826-0.843) | 0.788(0.779–0.798) | 0.774(0.762-0.786)   | 0.796(0.783-0.809)   |
| Internal validation cohort | 0.830(0.817-0.842) | 0.816(0.808-0.824) | 0.799(0.780-0.818)   | 0.827(0.815-0.839)   |
| External validation cohort | 0.816(0.808-0.824) | 0.778(0.770–0.786) | 0.775(0.760-0.790)   | 0.780(0.763–0.796)   |
| Prospective validation     | 0.776(0.771-0.782) | 0.797(0.790–0.803) | 0.693(0.681-0.705)   | 0.841(0.830-0.853)   |
| LR                         |                    |                    |                      |                      |
| Training cohort            | 0.790(0.788-0.792) | 0.776(0.772-0.780) | 0.689(0.678–0.699)   | 0.821(0.811-0.832)   |
| Internal validation cohort | 0.764(0.755-0.773) | 0.788(0.780-0.795) | 0.715(0.700-0.730)   | 0.826(0.812-0.839)   |
| External validation cohort | 0.795(0.794–0.795) | 0.845(0.844-0.846) | 0.597(0.594–0.599)   | 0.986(0.983–0.989)   |
| Prospective validation     | 0.802(0.801-0.803) | 0.844(0.841–0.846) | 0.684(0.679–0.689)   | 0.912(0.907-0.917)   |

Notes: AUC, the area under curve, RF, random forest, KNN, K-nearest neighbor, SVM, support vector machine, XGBoost, eXtreme Gradient Boosting, LR, logistic regression. Cl, confidence intervals

# Prognostic analysis of the optimal model

115 patients were followed up from institution I. Two patients who were lost to follow-up. The RFS rate was 36.5% (42/115). Statistically significant differences in RFS rates between cholangiocyte phenotype-positive HCC and classical HCC patients were observed (hazard ratio [HR], 1.994; 95% confidence interval [CI], 1.059–3.758; P=0.027) (Fig. 5A).

In the optimal models, 117 patients with HCC were included. To evaluate the prognostic stratification value of the model, patients were divided into predicted cholangiocyte phenotype negative (low-risk < 0.5) and cholangiocyte phenotype positive (high-risk > 0.5) groups based on the cut-off value by maximizing their Youden

index. Statistically significant differences in RFS rates between XGBoost (high-risk) and XGBoost (low-risk) patients were observed (HR, 1.986; 95% CI, 1.061–3.717; P = 0.029) (Fig. 5B).

## Discussion

In this study, the four most significant predictive features i.e., PLR, tumor capsule, target sign on HBP, and T1rt-20 min, were screened using LASSO and stepwise-selection logistic regression, and five machine learning models were compared to ultimately determine the model with optimal prediction performance. SHAP was used to visually interpret the models from both the overall and individual perspectives, and the predictive performances of



Fig. 3 Visualization of the model through the Shapley Additive Explanations (SHAP) algorithm. The SHAP bar graph (A) shows the degree of influence of the four most significant features on the final predicted probability. The SHAP scatter plot (B) shows the positive or negative effects of each feature on the prediction probability through different colors. The SHAP summary effort plot (C) shows the positive and negative effects of each feature of all cases in the model on the predictive probability. The SHAP decision plot (D) shows the transition path of a clinical sample's prediction from the baseline to the final outcome, allowing clinicians to identify diagnostic outliers or critical decision boundaries

the models were successfully validated using three validation cohorts. In addition, our results suggest that the construction XGBoost model can be used to stratify RFS according to prognosis.

The PLR and NLR are commonly used as inflammation-related indicators. Xu et al. [18] showed that NLR and PLR are closely related to the poor prognosis of tumors. In our study, PLR was an important predictor of cholangiocyte phenotype positive expression. Wei et al. [19] found that platelet count was highly correlated with positive CK19 expression, and Chen et al. [20] pointed out that PLR is an independent risk factor for extrahepatic metastasis after radical resection of an HCC, similar to the findings of this study. In addition, Lee et al. [21] found a strong correlation between NLR and CK19 expression. However, NLR was not included in the establishment of the final model in our study, which might be due to the different sample sizes. We consider that PLR is an important component of inflammation, which is involved in the formation of the tumor tissue microenvironment, thus changing the local regulation of tissue homeostasis, cell proliferation, and genetic stability, making tumors more invasive, and leading to worse prognosis, which was consistent with the findings of Park Y [22].

Encapsulation occurs due to the interaction between the tumor and hepatic parenchyma, which is considered a physical barrier that restricts tumor cells within the tumor boundary, thus inhibiting their proliferation (23–24). If the encapsulation appears to be ruptured on imaging, it indicates that the tumor has invaded the surrounding tissue, has higher invasiveness, results in a poorer prognosis, and is prone to recurrence [25]. In our study, incomplete encapsulation or non-capsular HCC were closely associated with cholangiocyte phenotypepositive HCC. Chen et al. [26] showed that the tumor capsule in HCC with a progenitor phenotype was mostly incomplete or absent, which is similar to our results. This may be related to the more aggressive growth type and



Fig. 4 Group A-C shows three examples of correctly predicted cholangiocyte phenotype negative. Group D-F shows three examples of correctly predicted cholangiocyte phenotype positive. A: T1rt-20 min (-), Tumor capsule (-), Target sign on HBP (-), PLR (-). B: T1rt-20 min (-), Tumor capsule (+), Target sign on HBP (-), PLR (+). C: T1rt-20 min (-), Tumor capsule (+), Target sign on HBP (+), PLR (-). D: T1rt-20 min (+), Tumor capsule (+), Target sign on HBP (+), PLR (+). E: T1rt-20 min (+), Tumor capsule (+), Target sign on HBP (-), PLR (-). D: T1rt-20 min (+), Tumor capsule (+), Target sign on HBP (-), PLR (-). E: T1rt-20 min (+), Tumor capsule (+), Target sign on HBP (-), PLR (-).

higher histological grade of cholangiocyte phenotype-positive HCC.

The LR-M features in the 2018 version of Liver Imaging Report and Data System (LI-RADS) definition include the target sign [27], which includes arterial rim enhancement, target sign on DWI, and target sign on HBP. In our study, the target sign was relatively common in cholangiocyte phenotype-positive HCC, which was similar to the findings of previous studies [26, 28]. Additionally, the target sign on HBP was an important predictor. This might be related to bile duct phenotypic differentiation in HCC with positive expression of CK19 or CK7, which may lead to the formation of a fibroproliferative interstitium within the tumor, causing the retention of gadoxetic acid in the HBP. Unfortunately, in this study, the arterial rim enhancement and target sign on DWI were not included in the model. This may be because of the presence of internal cysts or tumor necrosis. Larger HCCs are prone to intratumoral cystic and necrotic degeneration, which is not unique to cholangiocyte phenotype positive HCC. It is also frequently observed in other tumors, such as metastatic tumors and intrahepatic



Fig. 5 A: Kaplan Meier curves comparing RFS between patients with pathologically confirmed cholangiocyte phenotype-positive HCC (cholangiocyte phenotype +) and classical HCC groups (cholangiocyte phenotype -). B: Kaplan Meier curves comparing RFS between patients with XGBoost calculated high-risk and low-risk groups

cholangiocarcinomas. Nevertheless, these findings require further validation.

T1 mapping can be combined with gadoxetic acidenhanced MRI to provide more accurate and objective quantitative images with functional information. Previous studies have found that T1 mapping combined with gadoxetic acid-enhanced MRI could effectively predict Ki-67 expression in HCC [29], and previous reports also showed that T1 mapping combined with gadoxetic acid-enhanced MRI could effectively predict the biological characteristics of HCC [10-12]. In our study, T1rt-20 min was another important predictor of cholangiocyte phenotype-positive HCC, and it contributed most to the model. This finding was consistent with the findings of the studies by Zhao et al. [30] and Choi et al. [31]. To our knowledge, there is currently no research on quantitative MRI features for cholangiocyte phenotype positive HCC. We found that T1rt-20 min greater than 797ms is indicative of cholangiocyte phenotype positive expression. We speculate that cholangiocyte phenotype-positive HCC possessed high invasiveness, leading to active tumor cell proliferation and a higher density. Thus, normal hepatocytes are replaced by tumor cells, and the absorption of gadoxetic acid was reduced, resulting in a higher T1rt-20 min value.

Some studies have suggested that AFP [32], tumor margin [33], peritumoral enhancement, and satellite nodules [19] are risk factors for CK19 expression. However, these features were not included in model construction in this study, their significance cannot be overlooked and needs to be verified further. In addition, according to our results, cholangiocyte phenotype positive HCC has a higher recurrence rate, and model-based risk stratification can reflect prognostic differences.

Our results indicate that the XGBoost model demonstrates superior predictive performance and higher stability, which is consistent with previous studies [34, 35]. To facilitate clinical application, we employed the SHAP method to elucidate the model's prediction process. SHAP can quantify the importance of each feature in the model and visualize it, aiding physicians in effectively utilizing the model to identify high-risk patients with cholangiocytic phenotype HCC. Additionally, SHAP helps uncover hidden important features that may be overlooked in traditional diagnostic approaches. In this study, we were surprised to find that the absolute mean SHAP value of T1rt-20 min exceeded that of the other three key features, with an optimal cutoff value of 797 milliseconds. This highlights the significant additional value of T1 mapping imaging in conventional diagnostic methods. Therefore, SHAP can serve as a powerful tool to assist clinicians in interpreting model predictions, enhancing confidence in the clinical application of models, and optimizing physicians' diagnostic and therapeutic decisions. Previous studies have also demonstrated that SHAP has broad clinical applicability. Yixin et al. [36] developed a multicenter radiomics-clinical model to evaluate responses to whole-brain radiotherapy, employing SHAP to interpret the model. Their results demonstrated robust predictive performance, with SHAP providing clinician-friendly visual explanations of the model's decision process. Similarly, Yiqi et al. [37] utilized SHAP to interpret a multiparametric MRI-based radiomics model for predicting complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. These SHAP-driven interpretability approaches enable physicians to tailor more precise treatment strategies based on individual patient characteristics. In predicting heart failure risk, SHAP analysis revealed that ejection fraction, NT-proBNP levels, and renal function markers contributed most significantly to risk stratification. For patients with severely reduced ejection fraction and markedly elevated NT-proBNP levels, treatment strategies aimed at improving myocardial contractility and reducing cardiac load should be prioritized [38].

### Limitations

Our study had some limitations. First, the sample size of this study was small. Future studies with a larger sample size are required to validate the effectiveness of our prediction model. Second, this study only included patients with HCC who underwent surgical resection and excluded patients who were inoperable for clinical evaluation, which may have led to selection bias. Third, we have assessed CK19 expression while CK7, MUC-1, CA19-9, etc. expression are not assessed, which may have led to selection bias. Fourthly, the accuracy of assessment of lesion MRI features was dependent on the experience of radiologists. Therefore, radiomics requires careful consideration in future and further analyses on interpretable machine-learning methods involved in medicine should be conducted.

### Conclusion

In summary, we have developed and validated an optimal machine learning model based on T1 mapping gadoxetic acid-enhanced MRI for identifying cholangiocyte phenotype HCC. T1 mapping offers substantial incremental value. Personalized predictions were achieved through SHAP analysis, offering valuable insights to support clinical decision-making and potential applications in patient management.

### Abbreviations

| Abbieviatio | 113   |
|-------------|---|
| ADC         | Apparent Diffusion Coefficient                  |
| AFP         | Alpha-Fetoprotein                               |
| ALT         | Alanine Aminotransferase                        |
| AST         | Aspartate Aminotransferase                      |
| CI          | Confidence Intervals                            |
| CK          | Cytokeratin                                     |
| DWI         | Diffusion-Weighted Imaging                      |
| GGT         | Gamma-Glutamyltransferase                       |
| GPC-3       | Glypican-3                                      |
| GS          | Glutamine Synthetase                            |
| HBP         | Hepatobiliary Phase                             |
| HCC         | Hepatocellular Carcinoma                        |
| HepPar-1    | Hepatocyte Paraffin Antigen 1                   |
| ICC         | Interclass Correlation Coefficient              |
| KNN         | K-Nearest Neighbor                              |
| LASSO       | Least Absolute Shrinkage And Selection Operator |
| LI-RADS     | Liver Imaging Report And Data System            |
| LR          | Logistic Regression                             |
| NLR         | Neutrophil-To-Lymphocyte Ratio                  |
| PLR         | Platelet-To-Lymphocyte Ratio                    |
| RF          | Random Forest                                   |
| RFS         | Relapse-Free Survival                           |
| ROC         | Region Of Interest                              |
| SHAP        | Shapley Additive Explanation                    |
| SVM         | Support Vector Machine                          |
| T1WI        | T1-Weighted Imaging                             |

T2WI T2-Weighted Imaging XGBoost eXtreme Gradient Boosting

### **Supplementary Information**

The online version contains supplementary material available at https://doi.or g/10.1186/s12885-025-14147-3.

Supplementary Material 1

### Author contributions

Wei Liu: Participate in writing and revising papers. Zhiping Cai: Data collection, collation and statistical analysis. Yifan Chen: Collecting, sorting and analyzing the original results. Jieying Feng and Xingqun Guan: Put forward the concept/ basic framework of Collect cases. Haixiong Chen: Investigation and collation of documents. Baoliang Guo: Index detection. Fusheng OuYang: Data collection. Chun Luo: Data collection. Rong Zhang: Data collection. Xinjie Chen: Collated documents. Xiaohong Li: Sorting and analyzing the original results. Cuiru Zhou: Investigation and collation of documents. Shaomin Yang: Feasibility analysis of research scheme and revision of paper. Ziwei Liu and Qiugen Hu: Design ideas, design research schemes and feasibility analysis of research schemes, revising and reviewing papers. All authors reviewed the manuscript.

### Funding

This research was supported by The science and technology planning project of Foshan (2420001003711, 2320001006750); Scientific Research Start Plan of Shunde Hospital, Southern Medical University (SRSP2023008); Foshan Medical Imaging Artificial Intelligence Engineering Technology Application Research Center; Administration of Traditional Chinese Medicine of Guangdong Province of China (20241312); Guangdong Province Key Research Projects in Key Areas for General Colleges and Universities (2024ZDZX2029); Guangdong Medical Science and Technology Research Fund (A2024022, A2023204).

### Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request. [Qiugen Hu]

### Declarations

### Ethics approval and consent to participate

The need for written informed consent was waived by the Medical Ethics Committee of Shunde Hospital, Southern Medical University due to the observational design of the study.

#### **Consent for publication**

Informed consent for publication was obtained from all participants.

#### **Competing interests**

The authors declare no competing interests.

#### Author details

<sup>1</sup>Department of Radiology, The Eighth Affiliated Hospital of Southern Medical University (The First People's Hospital of Shunde Foshan), No. 1 Jiazi Road, Lunjiao, Shunde District, Foshan, Guangdong Province 528308, China

<sup>2</sup>Department of Radiology, The Sixth Affiliated Hospital, South China University of Technology, Foshan, Guangdong Province 528247, China <sup>3</sup>Department of Radiology, The First Peoples Hospital of Foshan, Foshan, Guangdong Province 528000, China

<sup>4</sup>Xingtan Hospital Affiliated of Southern Medical University Shunde Hospital, No. 222 Xinglong Road, Shunde, China

Received: 22 December 2023 / Accepted: 14 April 2025 Published online: 28 April 2025

- 1. Koshy A. Evolving global etiology of hepatocellular carcinoma (HCC): insights and trends for 2024. J Clin Exp Hepatol. 2025;15(1):102406.
- 2. Wang Y, Deng B. Hepatocellular carcinoma: molecular mechanism, targeted therapy, and biomarkers. Cancer Metastasis Rev. 2023;42(3):629–52.
- Chen Y, Chen J, Yang C, et al. Preoperative prediction of cholangiocyte phenotype hepatocellular carcinoma on contrast-enhanced MRI and the prognostic implication after hepatectomy. Insights Imaging. 2023;14(1):190. Published 2023 Nov 14.
- Zhuo JY, Lu D, Tan WY, Zheng SS, Shen YQ, Xu X. CK19-positive hepatocellular carcinoma is a characteristic subtype. J Cancer. 2020;11(17):5069–77. Published 2020 Jun 28.
- Huang X, Long L, Wei J, et al. Radiomics for diagnosis of dual-phenotype hepatocellular carcinoma using Gd-EOB-DTPA-enhanced MRI and patient prognosis. J Cancer Res Clin Oncol. 2019;145(12):2995–3003.
- Lehrich BM, Zhang J, Monga SP, Dhanasekaran R. Battle of the biopsies: role of tissue and liquid biopsy in hepatocellular carcinoma. J Hepatol. 2024;80(3):515–30.
- Lee S, Kim P, Im DJ, et al. The image quality and diagnostic accuracy of T1-mapping-based synthetic late gadolinium enhancement imaging: comparison with conventional late gadolinium enhancement imaging in real-life clinical situation. J Cardiovasc Magn Reson. 2022;24(1):28. Published 2022 Apr 14.
- Haimerl M, Verloh N, Zeman F, et al. Gd-EOB-DTPA-enhanced MRI for evaluation of liver function: comparison between signal-intensity-based indices and T1 relaxometry. Sci Rep. 2017;7:43347. Published 2017 Mar 7.
- 9. Ichikawa S, Goshima S. Gadoxetic Acid-Enhanced liver MRI: everything you need to know. Invest Radiol. 2024;59(1):53–68.
- Peng Z, Jiang M, Cai H, et al. Gd-EOB-DTPA-enhanced magnetic resonance imaging combined with T1 mapping predicts the degree of differentiation in hepatocellular carcinoma. BMC Cancer. 2016;16:625. Published 2016 Aug 12.
- Rao C, Wang X, Li M, Zhou G, Gu H. Value of T1 mapping on Gadoxetic acid-enhanced MRI for microvascular invasion of hepatocellular carcinoma: a retrospective study. BMC Med Imaging. 2020;20(1):43. Published 2020 Apr 28.
- Wang WT, Zhu S, Ding Y, et al. T1 mapping on Gadoxetic acid-enhanced MR imaging predicts recurrence of hepatocellular carcinoma after hepatectomy. Eur J Radiol. 2018;103:25–31.
- Zhang G, Shi Y, Yin P, et al. A machine learning model based on ultrasound image features to assess the risk of Sentinel lymph node metastasis in breast cancer patients: applications of scikit-learn and SHAP. Front Oncol. 2022;12:944569. Published 2022 Jul 25.
- 14. Nyrup R, Robinson D. Explanatory pragmatism: a context-sensitive framework for explainable medical AI. Ethics Inf Technol. 2022;24(1):13.
- Mastropietro A, Feldmann C, Bajorath J. Calculation of exact Shapley values for explaining support vector machine models using the radial basis function kernel. Sci Rep. 2023;13(1):19561. Published 2023 Nov 10.
- Wang K, Tian J, Zheng C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med. 2021;137:104813.
- Ma J, Bo Z, Zhao Z, et al. Machine learning to predict the response to lenvatinib combined with transarterial chemoembolization for unresectable hepatocellular carcinoma. Cancers (Basel). 2023;15(3):625. Published 2023 Jan 19.
- 18. Xu Y, Yuan X, Zhang X, et al. Prognostic value of inflammatory and nutritional markers for hepatocellular carcinoma. Med (Baltim). 2021;100(25):e26506.
- Shuyao W, Mingyang B, Feifei M, Xiaoqin H. CK19 predicts recurrence and prognosis of HBV positive HCC. J Gastrointest Surg. 2022;26(2):341–51.
- Chen Y, Zeng J, Guo P, Zeng J, Liu J. Prognostic significance of Platelet-to-Lymphocyte ratio (PLR) in extrahepatic metastasis of hepatocellular carcinoma after curative resection. Cancer Manag Res. 2021;13:1395–405. Published 2021 Feb 12.
- 21. Lee CW, Lin SE, Yu MC, et al. Does neutrophil to lymphocyte ratio have a role in identifying cytokeratin 19-Expressing hepatocellular carcinoma?? J Pers Med. 2021;11(11):1078. Published 2021 Oct 24.

- 22. Park Y, Chang AR. Neutrophil to lymphocyte ratio and platelet to lymphocyte ratio in hepatocellular carcinoma treated with stereotactic body radiotherapy. Korean J Gastroenterol. 2022;79(6):252–9.
- Lim JH, Choi D, Park CK, Lee WJ, Lim HK. Encapsulated hepatocellular carcinoma: CT-pathologic correlations. Eur Radiol. 2006;16(10):2326–33.
- Cho ES, Choi JY. MRI features of hepatocellular carcinoma related to biologic behavior. Korean J Radiol. 2015;16(3):449–64.
- Iguchi T, Aishima S, Sanefuji K, et al. Both fibrous capsule formation and extracapsular penetration are powerful predictors of poor survival in human hepatocellular carcinoma: a histological assessment of 365 patients in Japan. Ann Surg Oncol. 2009;16(9):2539–46.
- 26. Chen J, Wu Z, Xia C, et al. Noninvasive prediction of HCC with progenitor phenotype based on Gadoxetic acid-enhanced MRI. Eur Radiol. 2020;30(2):1232–42.
- Elsayes KM, Kielar AZ, Elmohr MM, et al. White paper of the society of abdominal radiology hepatocellular carcinoma diagnosis disease-focused panel on LI-RADS v2018 for CT and MRI. Abdom Radiol (NY). 2018;43(10):2625–42.
- Chen Y, Chen J, Zhang Y, et al. Preoperative prediction of cytokeratin 19 expression for hepatocellular carcinoma with deep learning radiomics based on Gadoxetic Acid-Enhanced magnetic resonance imaging. J Hepatocell Carcinoma. 2021;8:795–808. Published 2021 Jul 22.
- Liu Z, Yang S, Chen X, et al. Nomogram development and validation to predict Ki-67 expression of hepatocellular carcinoma derived from Gd-EOB-DTPA-enhanced MRI combined with T1 mapping. Front Oncol. 2022;12:954445. Published 2022 Oct 14.
- Zhao Y, Tan X, Chen J, et al. Preoperative prediction of cytokeratin-19 expression for hepatocellular carcinoma using T1 mapping on Gadoxetic acid-enhanced MRI combined with diffusion-weighted imaging and clinical indicators. Front Oncol. 2023;12:1068231. Published 2023 Jan 19.
- Choi SY, Kim SH, Park CK, et al. Imaging features of Gadoxetic Acid-enhanced and Diffusion-weighted MR imaging for identifying cytokeratin 19-positive hepatocellular carcinoma: A retrospective observational study. Radiology. 2018;286(3):897–908.
- 32. Cui DJ, Wu Y, Wen DH. CD34, PCNA and CK19 expressions in AFP- hepatocellular carcinoma. Eur Rev Med Pharmacol Sci. 2018;22(16):5200–5.
- Wang W, Gu D, Wei J, et al. A radiomics-based biomarker for cytokeratin 19 status of hepatocellular carcinoma with Gadoxetic acid-enhanced MRI. Eur Radiol. 2020;30(5):3004–14.
- Liu W, Zhang L, Xin Z, et al. A promising preoperative prediction model for microvascular invasion in hepatocellular carcinoma based on an extreme gradient boosting algorithm. Front Oncol. 2022;12:852736. Published 2022 Mar 4.
- Mao B, Zhang L, Ning P, et al. Preoperative prediction for pathological grade of hepatocellular carcinoma via machine learning-based radiomics. Eur Radiol. 2020;30(12):6924–32.
- Wang Y, Lang J, Zuo JZ, et al. The radiomic-clinical model using the SHAP method for assessing the treatment response of whole-brain radiotherapy: a multicentric study. Eur Radiol. 2022;32(12):8737–47.
- 37. Wang Y, Zhang L, Jiang Y, et al. Multiparametric magnetic resonance imaging (MRI)-based radiomics model explained by the Shapley additive explanations (SHAP) method for predicting complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a multicenter retrospective study. Quant Imaging Med Surg. 2024;14(7):4617–34.
- Lu S, Chen R, Wei W, Belovsky M, Lu X. Understanding Heart Failure Patients EHR Clinical Features via SHAP Interpretation of Tree-Based Machine Learning Model Predictions. AMIA Annu Symp Proc. 2022;2021:813–822.

# Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.