Journal of Hepatocellular Carcinoma

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Preoperative Prediction of Cytokeratin 19 Expression for Hepatocellular Carcinoma with Deep Learning Radiomics Based on Gadoxetic Acid-Enhanced Magnetic Resonance Imaging

Yuying Chen,[1,*] Jia Chen,[2,*] Yu Zhang,[3,*] Zhi Lin,[1] Meng Wang,[1] Lifei Huang,[2] Mengqi Huang,[1] Mimi Tang,[1] Xiaoqi Zhou,[1] Zhenpeng Peng,[1] Bingsheng Huang,[2] Shi-Ting Feng[1]

[1]Department of Radiology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, People's Republic of China; [2]Medical AI Lab, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, Guangdong, People's Republic of China; [3]Department of Pathology, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, People's Republic of China

*These authors contributed equally to this work

Correspondence: Bingsheng Huang
Medical AI Lab, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, Guangdong, People's Republic of China
Tel +86 755-86172208
Fax +86 755-86171820
Email huangb@szu.edu.cn

Shi-Ting Feng
Department of Radiology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, People's Republic of China
Tel +86 20-87755766 Ext 8471
Fax +86 20-87615805
Email fengsht@mail.sysu.edu.cn

Purpose: Cytokeratin 19 (CK19) expression is a proven independent prognostic predictor of hepatocellular carcinoma (HCC). This study aimed to develop and validate the performance of a deep learning radiomics (DLR) model for CK19 identification in HCC based on preoperative gadoxetic acid-enhanced magnetic resonance imaging (MRI).

Patients and Methods: A total of 141 surgically confirmed HCCs with preoperative gadoxetic acid-enhanced MRI from two institutions were included. Prediction models were established based on hepatobiliary phase (HBP) images using a training set (n=102) and validated using time-independent (n=19) and external (n=20) test sets. A receiver operating characteristic curve was used to evaluate the performance for CK19 prediction. Recurrence-free survival (RFS) was also analyzed by incorporating the CK19 expression and other factors.

Results: For predicting CK19 expression, the area under the curve (AUC) of the DLR model was 0.820 (95% confidence interval [CI]: 0.732–0.907, $P<0.001$) with sensitivity, specificity, accuracy of 0.800, 0.766, and 0.775, respectively, and reached 0.781 in the external test set. Combined with alpha fetoprotein, the AUC increased to 0.833 (95% CI: 0.753–0.912, $P<0.001$) and the sensitivity was 0.960. Intratumoral hemorrhage and peritumoral hypointensity on HBP were independent risk factors for HCC recurrence by multivariate analysis. Based on predicted CK19 expression and the independent risk factors, a nomogram was developed to predict RFS and achieved C-index of 0.707.

Conclusion: This study successfully established and verified an optimal DLR model for preoperative prediction of CK19-positive HCCs based on gadoxetic acid-enhanced MRI. The prediction of CK19 expression in HCC using a non-invasive method can help inform preoperative planning.

Keywords: hepatocellular carcinoma, gadoxetic acid, magnetic resonance imaging, cytokeratin 19, deep learning radiomics

Plain Language Summary

Cytokeratin 19 (CK19) expression is an important prognostic predictor of hepatocellular carcinoma (HCC). We developed a non-invasive artificial intelligence method for identifying CK19 in HCC based on preoperative magnetic resonance imaging. This method achieved good prediction of CK19 expression in HCC patients. Our work gives an interpretation of the correlation between imaging results and gene phenotypes, and provides a basis for revealing the prediction of HCC recurrence and prognosis before the treatment.

Received: 14 April 2021
Accepted: 6 July 2021
Published: 22 July 2021

Introduction

Hepatocellular carcinoma (HCC) is the second leading cause of cancer-related mortality worldwide, and has shown an increasing incidence.[1] The recurrence rate of HCC following radical resection is up to 69%,[2] and early recurrence within 2 years following surgery is associated with intrinsic gene expression in the tumor.[3]

Cytokeratin 19 (CK19), which is normally expressed in hepatic progenitor cells and cholangiocytes but not hepatocytes, has an important value in prognosis, diagnosis, and treatment of tumors.[4,5] Studies have found that approximately 20% of HCCs express markers of progenitor cells or cholangiocytes, and this type of HCC is known as dual phenotype HCC (DPHCC).[5,6] Compared with CK19-negative HCCs, CK19-positive HCCs exhibit highly aggressive behavior, which is associated with poorer prognosis and a higher rate of recurrence.[6,7] Expression of CK19 is an important independent prognostic predictor of HCC.[8] Preoperative prediction of CK19 expression is of great significance for prognostic evaluation and formulation of treatment strategies for patients with HCCs.

Currently, CK19 expression is clinically confirmed by preoperative biopsy or postoperative pathology.[4] Owing to intratumoral heterogeneity, local sampling is unable to reflect the overall characteristics.[9] On the other hand, imaging can reveal the overall perspective and peritumor status of lesions, especially gadoxetic acid-enhanced MRI.[10] Several studies have shown that imaging is closely related to the biological behavior and gene expression of HCC.[11,12] Some studies have evaluated the imaging findings related to CK19 expression in HCCs;[13,14] however, significant limitations in predicting CK19 using imaging findings remain. First, the analysis of image features is subjective and poor in terms of repeatability. Second, the sensitivity of MRI findings for prediction of CK19 is relatively low.

Radiomics can extract a large number of high-dimensional quantitative features from multimodal medical images, and then reveal the correlation between these features and the diagnosis, pathology, and prognosis of the tumor.[15] Recently, studies have used radiomics based on MRI to predict CK19 in HCCs.[16,17] Radiomics has progressed quite significantly, but the following problems remain. First, accurate image segmentation relies on manual delineation, which is time-consuming and easily affected by the operator. Second, the different designs of the image features can lead to different analytic results. Furthermore, pre-designed image features such as intensity, shape, texture, and wavelets can be computed, while all the features of lesions are included.[18] Deep learning (DL) does not require definition of specific radiological features to interpret the image, and it may even be possible to discover other important features that are not yet known.[19–21] Several studies achieved good results with the use of the combination of DL and radiomics to solve scientific problems related to medical imaging.[18,22–24] Li et al[18] proposed the concept of DL-based radiomics (DLR) to overcome the shortcomings of radiomics, and achieved successful prediction of IDH1 expression in low-grade glioma through MRI-based DLR. Huang et al[24] proposed the framework of deep semantic segmentation feature-based radiomics utilizing a segmentation network to extract effective features, and achieved better classification performance than traditional radiomics.

In this study, we aimed to develop a DLR model to predict CK19 in HCC based on gadoxetic acid-enhanced MRI. Furthermore, we aimed to investigate the prognostic effect of CK19 on recurrence-free survival (RFS) in HCC. We hypothesized that DLR could extract effective features associated with CK19 expression based on gadoxetic acid-enhanced MRI and give help in preoperative prediction of CK19 expression in HCC.

Patients and Methods

Patients

The study was conducted in accordance with Declaration of Helsinki and approved by our Institutional Ethics Review Board, including waiver of informed consent (approval number: [2021]124). The reasons for waiver of informed consent are as follows: First, due to the retrospective nature of this study, it is unrealistic or impossible to obtain informed consent of all patients. Second, the study does not exceed the minimum risk after review by the ethics committee. Third, all patient data accessed complied with relevant data protection and privacy regulations; Fourth, the rights and interests of all patients have not been invaded.

The flow chart of data collection and study design is shown in Figure 1. We retrospectively collected patients from the First Affiliated Hospital of Sun Yat-sen University (institution I) and Sun Yat-sen University Cancer Center (institution II). Patients who underwent gadoxetic acid-enhanced MRI within 1 month before radical surgery for solid single HCC confirmed by pathology along with routine immunochemical staining for CK19 were included. The exclusion criteria were as follows: 1) administration of other preoperative antitumor therapies and 2) suboptimal MR image quality for interpretation. MR images of 102

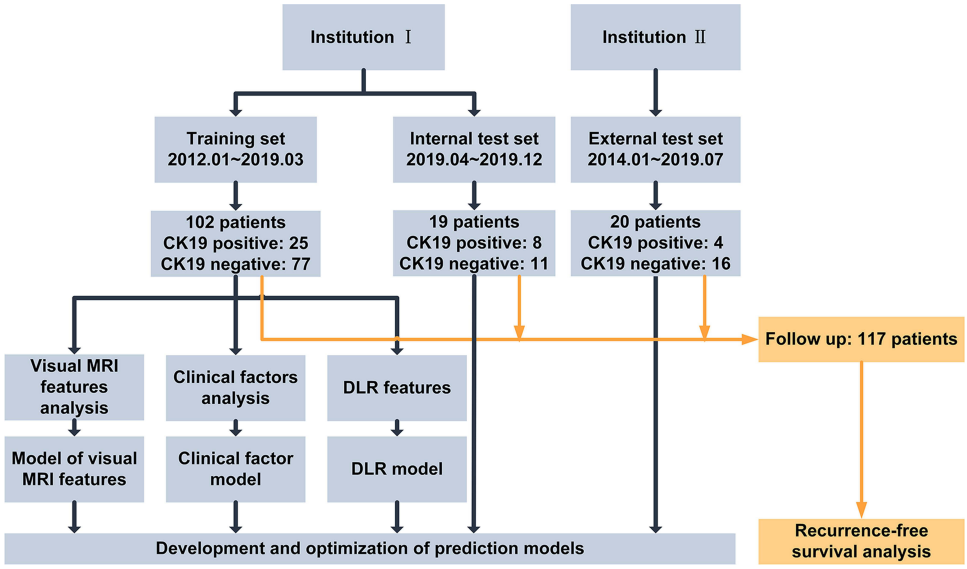


Figure 1 Study flowchart.
Abbreviations: Institution I, The First Affiliated Hospital of Sun Yat-sen University; Institution II, Sun Yat-sen University Cancer Center; CK19, Cytokeratin-19; DLR, deep learning radiomics.

patients from institution I were used as the training set to establish prediction models to predict CK19 expression in HCC based on the hepatobiliary phase (HBP) of gadoxetic acid-enhanced MRI. The predictive performance of models was evaluated by test sets (19 cases from institution I and 20 cases from institution II).

Magnetic Resonance Imaging Protocol

MRI examination in all patients from institution I and institution II were performed using a Magnetom Trio A Tim 3.0T system (Siemens Healthcare Sector, Erlangen, Germany) and GE 3.0T (750W, Pioneer; GE Healthcare, Milwaukee, WI) MR scanning system, respectively. The scanning range covered from the top to the lower edge of the liver with an 8-channel phased-array coil as the receiver coil. Gadoxetic acid-enhanced MRI was obtained including the unenhanced phase, enhanced arterial phase (20–40s), portal phase (50–70s), equilibrium phase (100–120s), transitional period (3–5min), and 20-min HBP images. Gadoxetate disodium (gadoxetic acid) (Primegen; Bayer Schering Pharma, Berlin, Germany) was injected into the cubital vein at a flow rate of 1 mL/s and a dose of 0.025 mmol/kg, followed by 20 mL of normal saline for flushing. One hundred and eighteen patients underwent T_1 mapping imaging.

A more detailed description of the MRI methods and specific sequences and parameters of MRI scans are shown in Supplementary Table 1.

The ground truth of tumor lesions on the HBP of MRI on HCC patients was contoured by an experienced radiologist without knowledge of the clinical and pathological findings. The radiologist labeled three layers which included the first, largest, and last layer of tumor appearance on cross-section MRI based on Insight Toolkit (ITK)-snap software. The marked quadrangles all contained tumor areas (Figure 2).

Clinicopathological Analyses

The pathological sampling of HCC specimens was based on the baseline 7-point sampling scheme in the standardized pathological diagnosis guidelines for primary liver cancer and included sampling at the junction of carcinoma and adjacent liver tissues. The diagnostic criteria for HCC were based on morphological criteria defined by the World Health Organization. The expression of CK19 was semi-quantitatively evaluated by immunochemical staining. The hepatocytes and bile ducts of normal liver tissues were used as negative and positive controls, respectively. Tumors were classified as negative (<5% of tumor cells) or positive (≥5% of tumor cells) for CK19 by an experienced pathologist who was blinded to clinical and imaging information.

The relevant clinical indicators and reference range of the laboratory in this study are shown in the Supplementary Table 2.

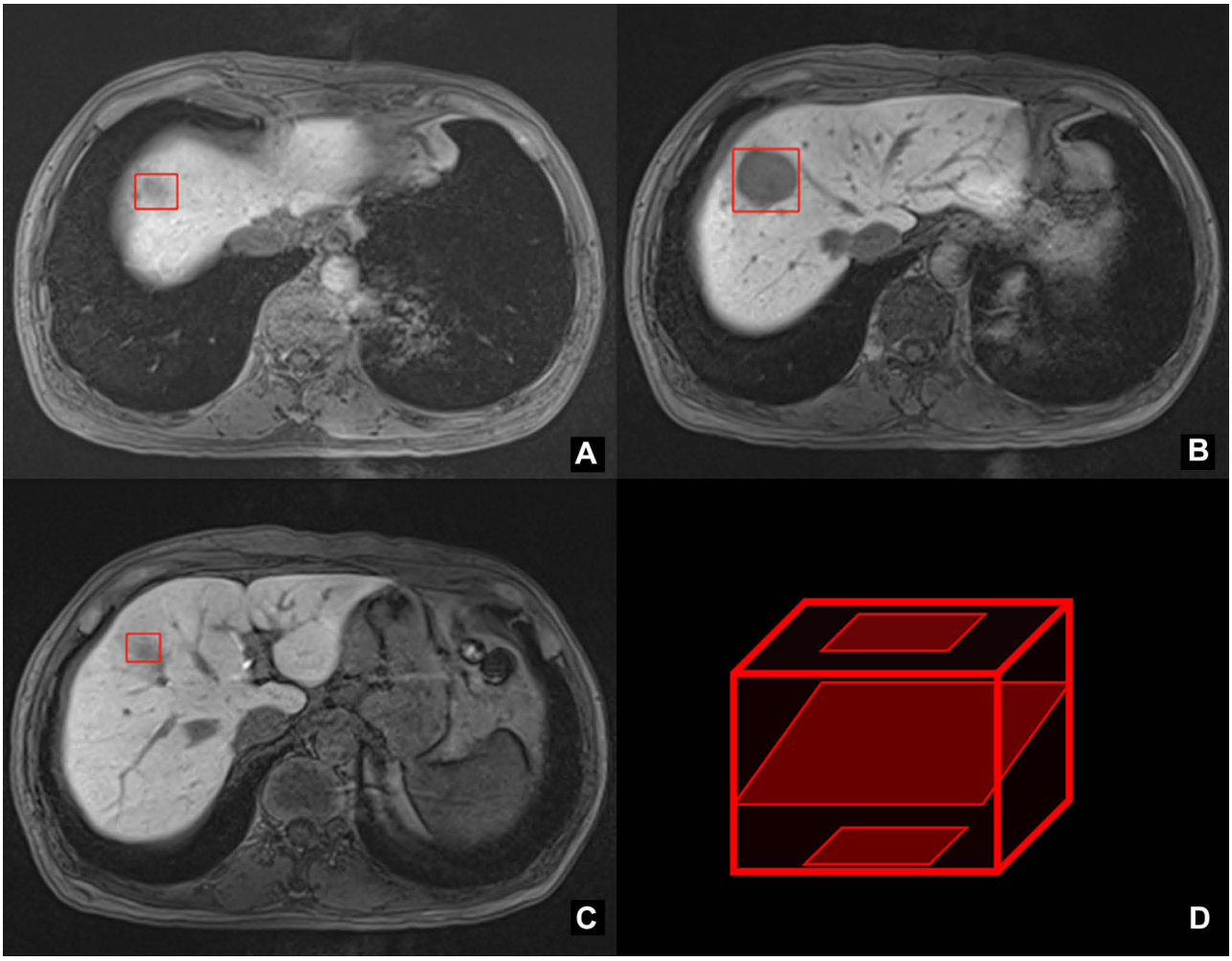


Figure 2 Lesion labeling on hepatobiliary phase image for deep learning radiomics analysis. The first (**A**), largest (**B**) and last layer (**C**) of tumor was roughly contoured by a radiologist on cross-section hepatobiliary phase images to build a cube area of interest including the tumor lesion (**D**).

Univariable analysis was used to assess the association between clinical factors and CK19 expression of HCC. Variables with a *P* value <0.05 in the multivariable logistic regression analysis were identified as potential clinical risk factors and the model based on clinical factors was constructed at the same time.

Postoperative Follow-Up

Imaging examinations (ultrasound/CT/MRI) were conducted every 3–6 months for evaluation of tumor recurrence thereafter, and the RFS of each patient was recorded. Tumor recurrence of HCC included intrahepatic recurrence and distant metastasis, which were defined as the appearance of new lesions in the residual liver and new metastatic lesions in solid organs, lymph nodes, and peritoneum, respectively.

Traditional Features of MRI

Traditional MRI features were evaluated by two senior radiologists without knowledge of the clinical, surgical, and pathological findings. The regions of interest manually delineated on ADC maps, HBP, and T_1mapping images was as large as possible to avoid areas such as the cystic portion, necrosis, hemorrhage, fat, and artifacts. The traditional quantitative features acquired in these ROIs included 1) lesion size, 2) the average apparent diffusion coefficient (ADC, *b*=800 s/mm^2) value, 3) the average HBP signal intensity (SI [HBP]), and 4) T_1 value (T_{1N}) and T_1 reduction rate (T_{1D}). The traditional qualitative MRI features included (Supplementary Figure 1) 1) tumor margin, 2) dynamic enhanced pattern, 3) abnormal peritumoral perfusion on the arterial phase, 4) tumor vessels on the arterial phase, 5) rim enhancement on the arterial phase, 6) SI on the arterial phase compared to normal

liver parenchyma, 7) peritumoral hypointensity on HBP images, 8) target sign on HBP/diffusion-weighted imaging (DWI; b=800 s/mm^2), 9) intratumoral cystic portion or necrosis, 10) intratumoral hemorrhage, 11) fat content, 12) tumor capsule, 13) vessel invasion, and 14) intratumoral septum.

A more detailed description of traditional MRI features is shown in the Supplementary Method.

The association between traditional MR features and CK19 expression of HCC was assessed by univariable analysis. Potential risk factors were identified in the multivariable analysis with a P value <0.05 and the model based on traditional MR features was constructed at the same time.

Training and Validation of CK19 Predicting Models with DLR

In this section, we describe the details of the proposed framework. To facilitate understanding, the pipeline of our framework is presented in Figure 3. The proposed method included a deep semantic feature extraction module and a classifier training module. First, the convolutional neural network segmentation network was trained to obtain the deep feature extraction module. Second, semantic features were extracted from the HCC dataset and then used to construct feature sets for the training of classifiers. Finally,

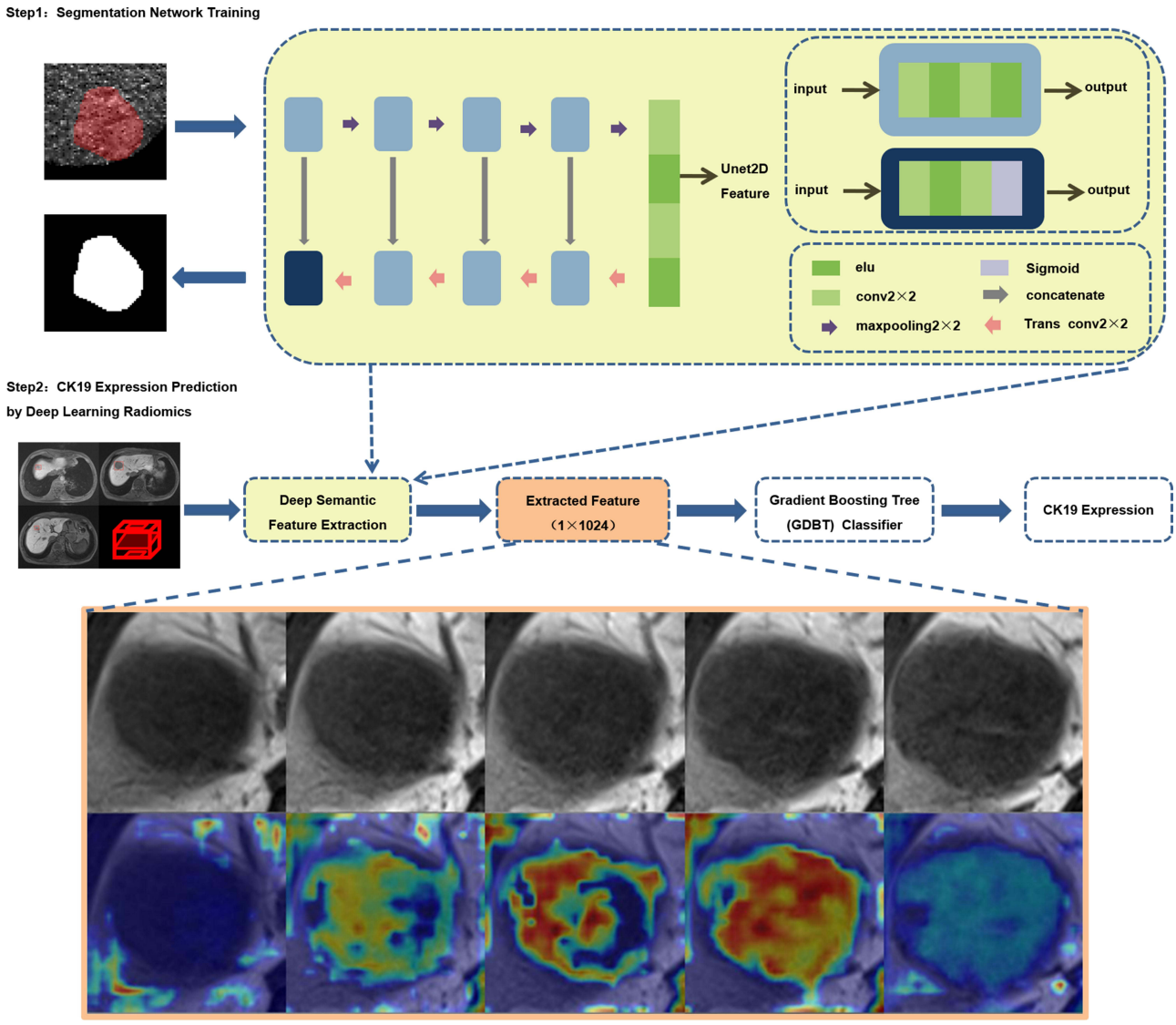


Figure 3 Deep learning radiomics analysis workflow. The deep learning radiomics analysis consists of two steps. First, the deep feature extraction module is obtained by training the CNN segmentation network. From MRI images of hepatocellular carcinoma, semantic features are extracted by the network and then used to construct feature sets for the training of classifiers. Second, the feature set is fed into a machine learning classifier to establish the prediction model for CK19 expression.
Abbreviations: CK19, cytokeratin-19; CNN, convolutional neural network; DLR, deep learning radiomics.

the feature set was fed into a machine learning classifier for training and testing to build a classification model.

Deep Learning Radiomics Feature Extraction

In this study, computed tomography (CT) images from 80 patients with HCC were specifically used for training the U-Net segmentation model 64, 8, and 8 patients were regarded as the training, validation, and test sets, respectively. Please note that the CT dataset we collected were only used for the training and testing of the segmentation network and not for the extraction of deep semantic features. More details of image preprocessing and the training strategy for segmentation network can be found in the Supplementary Method.

After the segmentation model was established, the HBP images of hepatocellular carcinoma after image preprocessing were fed into the segmentation model for semantic feature extraction.[24] More details of HBP MRI image feature extraction can be found in the Supplementary Method.

Training and Testing of Machine Learning Models

Based on the extracted deep semantic features, a prediction model of CK19 expression was constructed by using the Gradient Boosting Tree (GDBT) classifier. The classifiers were all implemented using the Python (version 3.7; https://www.python.org/) machine-learning library known as scikit-learn (version 0.23.2; https://scikit-learn.org/stable). CK19 expression (positive or negative) was used as a classification label, and deep semantic features were used as the input features. The training applied a grid search method to obtain the optimal parameter combination to establish the classification model. Training and testing of the classification model included two steps: 1) 10-fold internal cross-validation (CV). A classification model was established and tested by randomly dividing internal dataset (102 patients from institution I) into training (9 folds) and test set (1 fold), and the step was repeated for 10 times with a different fold as the test set each time. 2) independent validation, in which we evaluated the performance of the 10 models acquired in the internal CV with internal test data (19 patients from institution I) and external test data (20 patients from institution II). We obtained the average prediction probability with the 10 models as the final prediction probability value and evaluated the predicting performance.

Comparison of the Performance of Prediction Models

To better illustrate the potential clinical value of DLR model, a comparison of the DLR model and the model based on clinical factors and traditional MR features will be conducted. To determine whether the combination of DLR algorithm, clinical factors and traditional MRI features can improve predictive performance, a comparison between DLR model and the integrated model will be performed.

RFS Analysis by Incorporating CK19 Expression and Other Risk Factors

To test whether the expression of CK19, clinical factors and MRI features were complementary for the prediction of recurrence in patients with HCC, we created a combined model integrating all the potential factors together to a multivariable Cox regression model. Clinical factors and traditional MRI features with significant association with recurrence were entered into the multivariable combined model along with actual CK19.

Statistical Analysis

SPSS software (Version 24.0; IBM Corp., Armonk, NY) and R software (version 3.4.1) were used for analysis. The kappa test was performed to assess the consistency between two radiologists in the traditional MRI features analysis (kappa value determination: $\kappa>0.75$, excellent; $0.40\leq\kappa\leq0.75$, good; $\kappa<0.40$ poor). A Student *t*-test (mean ± standard deviation) or Wilcoxon rank-sum test (median, P_{25}~P_{75}) was performed for continuous variables. The categorical variables were compared by χ^2 or Fisher's exact tests. Multivariable logistic regression analyses were performed to identify the independent predictors of CK19-positive HCCs. The receiver operator characteristic (ROC) curve was used to evaluate the performance of predicting the expression of CK19. The comparison among ROC curves was analyzed using the MedCalc Statistical Software version 15.8 (MedCalc Software bvba, Ostend, Belgium, https://www.medcalc.org/manual/sampling_ROC1.php). The Kaplan–Meier method was used to evaluate the cumulative RFS. Log rank test was used to evaluate the differences between groups. Univariate and multivariate Cox proportional hazards models were used to determine the risk factors for RFS. The performance of nomogram for predicting the prognosis was evaluated by calculating C-index. All differences were considered statistically significant with a P value of <0.05. For the internal CV of the CK19 prediction study, we investigated whether the sample size in our study was sufficient to detect an statistical difference based on the following conditions by using MedCalc Statistical Software: statistical power, 80%; a two-tailed significance level, 0.05; the true

AUC values of the classification model for CK19 prediction; the null hypothesis of AUC=0.5; ratio of classes, the real ratios in our study, for example, 25 CK19-positive/77 CK19-negative in the internal dataset of CK19.

Results

Clinicopathological Features of the Training Set and the Test Sets

A total of 121 patients from institution I were included in this study, including 33 CK19-positive HCCs and 88 CK19-negative HCCs. Twenty patients from institution II were included, among which four were positive and 16 were negative regarding CK19 expression. The distribution of clinicopathological features in the training group and the verification group is shown in Table 1.

The results of univariate analysis of clinical factors in the training set showed that AFP was a significant variable associated with CK19 ($P<0.05$, Supplementary Table 3), which was then incorporated into the logistic regression model. The result indicated that AFP >400 ug/L was an independent factor for positivity of CK19 in HCC ($P=0.003$, OR=4.544, 95% confidence interval [CI]: 1.696–12.175).

Traditional MRI Features of HCCs Related to CK19 Expression in the Training Set

The results showed there were no significant differences in quantitative variables between CK19-positive HCCs and CK19-negative HCCs. Consistency analysis of qualitative features indicated that kappa values were all above 0.600 (0.637–1.000, $P<0.001$, Supplementary Table 4), which proved that the two radiologists were consistent in the analysis of traditional qualitative features. CK19-positive HCCs more frequently showed arterial rim enhancement ($P=0.018$) and target sign on HBP images ($P=0.005$) or DWI ($P=0.001$) compared to CK19-negative HCCs (Supplementary Table 4, Figure 4). Target sign on DWI ($P=0.001$, OR=4.875, 95% CI: 1.838–12.927) was an independent significant variable associated with CK19-positive HCCs.

Performance of CK19 Predicting Models

The predictive efficacy of each model in the training set and the test sets is shown in Table 2. ROC analysis among models was followed by the DeLong test to compare the predictive performance (Table 3). The ROC curves of each model in training set are shown in Figure 5. The results indicated DLR model significantly achieved better performance than the clinical factor model (AUC 0.820 vs 0.656, $P=0.008$) and the model of traditional MRI features (AUC 0.820 vs 0.669, $P=0.023$). The combination of the model of traditional MRI features with AFP level showed significantly better predictive efficiency ($P=0.008$). After the DLR model was combined with AFP level, the AUC of the model was increased to 0.833, with a sensitivity up to 96%.

A sample size of 29 patients (7 CK19-positive/22 CK19-negative) is required for the classification model establishment of the CK19 prediction study. Therefore, the internal sample size (102 patients in institution I) in this study was considered able to detect an AUC (0.792 for

Table 1 Baseline Clinical Characteristics and Pathological Parameters of the Training and Test Dataset

Variables	Total n=141 (%)	Training Set n=102 (%)	Test 1 n=19 (%)	*P* value	Test 2 n=20 (%)	*P* value
Age (years)	54±13	53±13	54±14	0.965	55±11	0.519
Sex (male)	120 (85.1)	94 (92.2)	12 (63.2)	<0.001	14 (70.0)	0.004*
History of hepatitis B	127 (90.1)	94 (92.2)	17 (89.5)	0.697	16 (80.0)	0.095
Child–Pugh class A	137 (97.2)	99 (97.1)	18 (94.7)	0.603	20 (100)	1.000
ALT >40 U/L	45 (31.9)	34 (33.3)	7 (36.8)	0.767	4 (20.0)	0.239
AST >37 U/L	47 (33.3)	34 (33.3)	7 (36.8)	0.767	6 (30.0)	0.772
TBIL >20 μmol/L	24(17.0)	21 (17.6)	3 (15.8)	0.630	0 (0)	0.026*
ALB <35 g/L	14 (9.9)	12 (11.8)	2 (10.5)	0.877	0 (0)	0.106
PT >14 s	4 (2.8)	1 (1.0)	1 (5.3)	0.179	2 (10.0)	0.017*
AFP >400 μg/L	37 (26.2)	25 (24.5)	7 (36.8)	0.263	5 (25.0)	0.963
CA19–9 >35 U/mL	14 (9.9)	6 (5.9)	1 (5.3)	0.915	7 (35.0)	<0.001*
CK19-positive HCCs	37 (26.2)	25 (24.5)	8 (42.1)	0.114	4 (20.0)	0.665

Notes: *$P<0.05$. Values are represented as mean ± standard deviation or number (percentage). *P*- values represent the result of comparison of the training set with the two test sets, respectively.

Abbreviations: Test 1, internal test set; Test 2, external test set; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; ALB, albumin; PT, prothrombin time; AFP, alpha fetoprotein; CA19-9, carbohydrate antigen 19-9; CK19, cytokeratin 19.

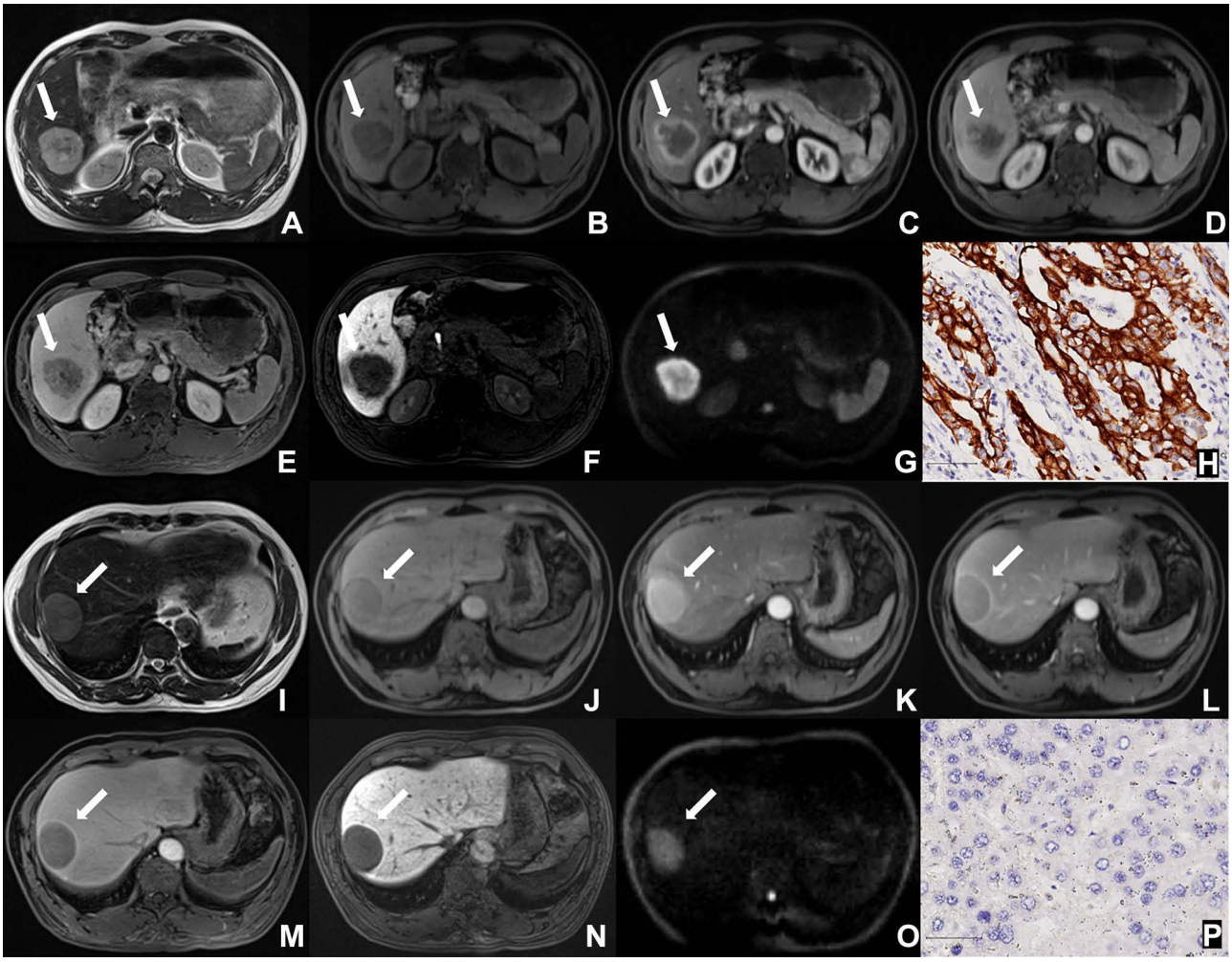


Figure 4 Gadoxetic acid-enhanced magnetic resonance images and pathological immunohistochemistry of CK19-positive and CK19-negative HCCs. CK19-positive HCC in a 41-year-old male patient (**A**–**H**, white arrow). There was rim enhancement on the arterial phase (**C**) with targetoid sign on the hepatobiliary phase (**F**) and diffusion-weighted imaging (**G**; *b*=800 s/mm^2). On histopathological analyses, this tumor was positive for CK19 (**H**; scale bar: 50μm). (**I**–**P**) Images show a CK19-negative hepatocellular carcinoma (white arrow) in a 56-year-old male patient. There is a round hypointensity lesion on T_1 weighted image (**I**) with non-rim hyperenhancement on the arterial phase (**K**). The tumor shows non-peripheral wash out with enhancing capsule on the portal phase (**L**). It shows homogeneous marked diffusion restriction on diffusion-weighted imaging (**O**; *b*=800 s/mm^2). Immunohistochemistry suggests negative expression of CK19 (**P**; scale bar: 50μm).
Abbreviations: CK19, cytokeratin-19; HCC, hepatocellular carcinoma.

the model based on DLR, AFP and traditional MRI features) different from 0.500 with 80% power.

RFS Analysis of HCC and Nomogram Construction

One hundred and seventeen of the 141 cases included had follow-up data, and 39 cases (33.3%) had recurrence after curative recession of HCC. The Kaplan Meier curves illustrated that the RFS rates were lower in patients with CK19-positive HCCs compared with CK19-negative HCCs, but not significantly (Figure 6A, P=0.301). The univariate and multivariate Cox regression analyses were used to identify the risk factors for recurrence in HCC. Intratumoral hemorrhage (HR: 2.261, 95% CI: 1.205–4.627; P = 0.012) and peritumoral hypointensity on HBP (HR: 2.427, 95% CI: 1.188–4.957; P = 0.015) were found to be independent risk factors for recurrence (Table 4). Based on CK19 expression of HCC and the independent risk factors including intratumoral hemorrhage and peritumoral hypointensity on HBP identified by the multivariate analysis, a nomogram was developed to predict 6-month, 1-year, and 2-year RFS probabilities for HCC patients (Figure 6B, C-index=0.707).

Discussion

In this study, a DLR algorithm based on gadoxetic acid-enhanced MRI was used for the first time to successfully

Table 2 Performance of Prediction Models in Internal and External Data

Models		ACC	SEN	SPC	AUC	95% CI of AUC	*P* value
AFP	Internal CV	0.745	0.480	0.831	0.656	0.524–0.787	0.020*
	Independent test1	0.632	0.500	0.727	0.614	0.349–0.878	0.409
	Independent test2	0.750	0.500	0.813	0.656	0.330–0.983	0.345
Traditional MRI features	Internal CV	0.745	0.520	0.818	0.669	0.539–0.799	0.011*
	Independent test1	0.632	0.375	0.818	0.597	0.328–0.865	0.483
	Independent test2	0.700	0.500	0.750	0.625	0.299–0.951	0.450
DLR	Internal CV	0.775	0.800	0.766	0.820	0.732–0.907	<0.001*
	Independent test1	0.631	0.750	0.545	0.591	0.314–0.868	0.509
	Independent test2	0.850	0.750	0.875	0.781	0.514–1.000	0.089
Traditional MRI features+AFP	Internal CV	0.706	0.800	0.675	0.761	0.650–0.871	<0.001*
	Independent test1	0.632	0.750	0.545	0.676	0.425–0.927	0.201
	Independent test2	0.650	0.750	0.625	0.719	0.422–1.000	0.186
DLR+ AFP	Internal CV	0.706	0.960	0.623	0.833	0.753–0.912	<0.001*
	Independent test1	0.684	0.500	0.818	0.614	0.342–0.885	0.409
	Independent test2	0.800	0.750	0.815	0.750	0.487–1.000	0.131
DLR+ Traditional MRI features	Internal CV	0.676	0.960	0.584	0.815	0.732–0.899	<0.001*
	Independent test1	0.789	0.625	0.909	0.682	0.411–0.952	0.186
	Independent test2	0.800	0.500	0.875	0.688	0.394–0.981	0.257
DLR+ AFP+ Traditional MRI features	Internal CV	0.686	0.840	0.636	0.792	0.690–0.893	<0.001*
	Independent test1	0.737	0.500	0.909	0.648	0.378–0.917	0.283
	Independent test2	0.650	0.750	0.625	0.672	0.383–0.961	0.299

Notes: *$P<0.05$. The *P* values indicate the significance level of the model to predict CK19 expression in HCC.
Abbreviations: Independent test 1, time-independent internal test; Independent test 2, independent external test; CV, cross-validation; DLR, deep learning radiomics; AFP, alpha fetoprotein; ACC, accuracy; SEN, sensitivity; SPC, specificity; AUC, area under the curve; CI, confidence interval.

Table 3 Comparison of the Performance of Prediction Models Based on ROC Curves by DeLong's Test (P-values Presented)

Models		1	2	3	2+1	3+1	3+2	3+1+2
	AUC	0.656	0.669	0.820	0.761	0.833	0.815	0.792
1	0.656		0.874	0.008*	0.086	0.003*	0.012*	0.024*
2	0.669			0.023*	0.008*	0.014*	0.025*	0.100
3	0.820				0.311	0.528	0.877	0.375
2+1	0.761					0.232	0.371	0.064
3+1	0.833						0.446	0.136
3+2	0.815							0.439
3+1+2	0.792							

Notes: *$P<0.05$. Values in table indicate the significance level of the AUCs comparison between different two models. 1, model based on clinical factors (alpha fetoprotein); 2, model based on traditional features (target sign on diffusion-weighted imaging); 3, deep learning radiomics model.
Abbreviations: AUC, area under the curve; ROC, receiver operating characteristic.

develop a model for CK19 expression prediction in HCC. The results showed that the DLR model provided optimal performance in predicting CK19 in HCC, with good robustness. The DLR algorithm combined with clinical parameters is a promising non-invasive method to help preoperative prediction of CK19.

We found that targetoid features in MRI, such as arterial rim enhancement and target sign on DWI/HBP, were closely related to CK19 in HCC, which was similar to the findings of Hu et al.[14] Previous studies have indicated that arterial rim enhancement and target sign on DWI were important independent predictors for the diagnosis of

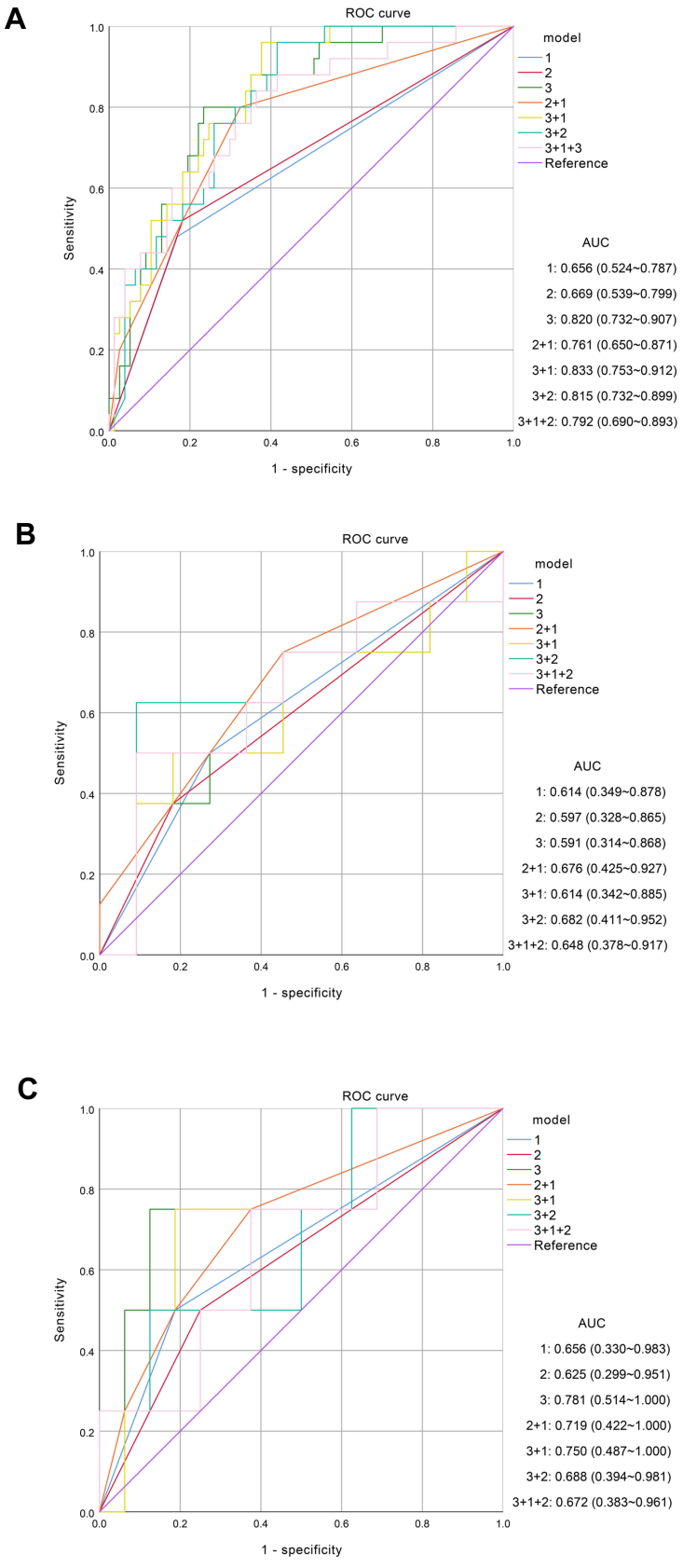
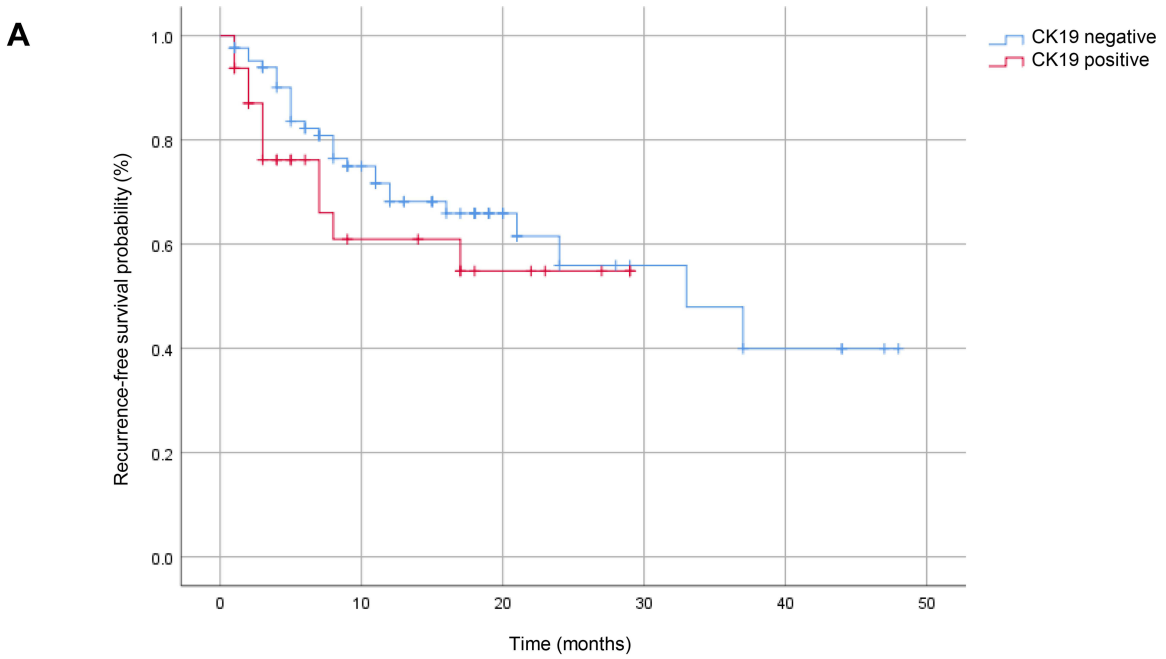


Figure 5 Comparison of receiver operating characteristics (ROC) curves for predicting CK19 status of HCC. ROC curves of each prediction model in the internal CV (**A**), internal test set (**B**), and external test set (**C**). 1: model based on clinical factors (alpha fetoprotein); 2: model based on traditional features (target sign on DWI); 3: deep learning radiomics model.
Abbreviations: HCC, hepatocellular carcinoma; CK19, cytokeratin 19; CV, cross-validation.



Group		Total (n=117)	Recurrence-free survival rate (%)			*P value*
			6 months	12 months	24 months	
CK19	Positive	32	76.2	60.9	54.8	0.301
	Negative	85	82.2	68.2	55.9	

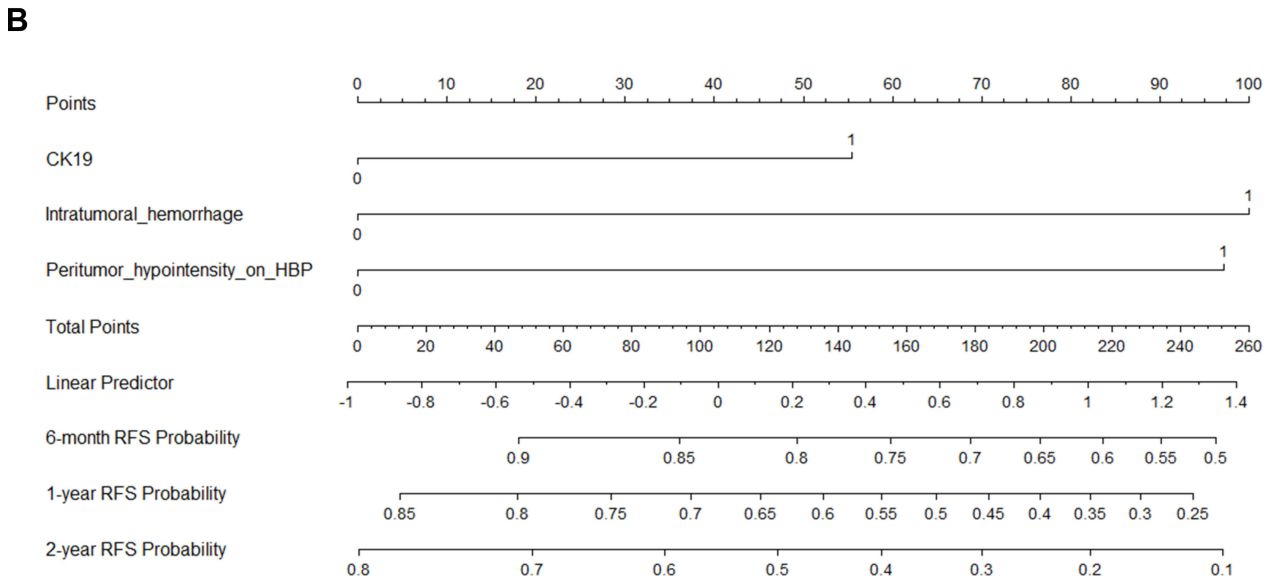


Figure 6 Kaplan–Meier survival analysis (**A**) and nomogram for predicting RFS (**B**) in patients with HCC. (**A**) Kaplan–Meier curve for the RFS of CK19-negative and CK19-positive patients with HCC. (**B**) The nomogram for predicting the 6-month, 1-year, and 2-year RFS. Each risk factor was allocated a predicting score, and the sum of three scores was located on the total points axis, suggesting the prediction of 6-month, 1-year, and 2-year RFS probabilities.
Abbreviations: HCC, hepatocellular carcinoma; RFS, recurrence-free survival; CK19, cytokeratin 19.

intrahepatic cholangiocarcinoma (ICC),[25,26] which are related to the pathological morphology of peripheral hyperproliferation and central stromal fibrosis.[27] The formation of stromal fibrosis seemed to be more common in CK19-positive HCCs than in CK19-negative HCCs, which indicated the morphological characteristics of DPHCC could be between those of typical HCC and ICC.[28] Based on traditional MRI features of HCCs, the sensitivity

Table 4 Univariate and Multivariate COX Analyses of Recurrence-Free Survival in HCC

Variable	Total	Univariate		*P* value	Multivariate		*P* value
	(n=117)	HR	95% CI		HR	95% CI	
AST(U/L)				0.020*			0.053
>37	39	2.116	1.123–3.986				
≤37	78	1.000					
Intratumoral hemorrhage				0.002*			0.012*
Present	42	2.760	1.432–5.318		2.361	1.205–4.627	
Absent	75	1.000			1.000		
Peritumor hypointensity on HBP				0.004*			0.015*
Present	65	2.798	1.389–5.638		2.427	1.188–4.957	
Absent	52	1.000			1.000		

Notes: Variables with a *P* value of < 0.05 on the univariate analysis were included for the multivariate analysis via the forward stepwise model. *$P<0.05$. The *P* values indicate the significance level of the difference in variables between the two groups in this model.
Abbreviations: AST, aspartate aminotransferase; HR, hazard ratios; CI, confidence interval.

for prediction of CK19 expression was not high in our study, which was similar to the results of previous studies;[13,14] therefore, there remain barriers to its application in daily practice.

We found that compared with the model based on traditional features evaluated by radiologists, the DLR model had significantly better performance in predicting CK19 in HCC. The DLR model manifested optimal performance in comparison to the model based on traditional features. Furthermore, the application of the DLR model in external test data demonstrated good robustness with different MRI scanners and with different parameter settings. This study utilized CT images to train the DL segmentation network, which could help obtain abstract features for better characterization of HCC lesions. Also, the DLR algorithm fully utilized the image information including intratumoral and peritumoral regions of the hepatic lesions to provide more valuable relevant information for better prediction.[24]

Our study showed that a significant increase in preoperative AFP level (>400 ug/L) was an independent clinical predictor of CK19 expression. AFP level was an important tumor marker for HCC, which has been proven to be correlated with CK19 expression.[29] A comparative analysis in our study suggested that combining preoperative serum AFP level of HCC patients was helpful for improving the predictive efficacy of the model based on MRI features evaluated by radiologists or extracted by DLR algorithm. Some studies have established models based on AFP, MRI features evaluated by radiologists or radiomic features, and have also achieved good results in predicting HCC with progenitor phenotype.[16,17] The combination of clinical information and MRI features can provide complementary information in development of a model and improve predictive performance.[22] However, the performance of a combined model based on DLR algorithm and other independent factors related to CK19, was not significantly improved. Features extracted by the DLR algorithm were more abstract than traditional features, and were less affected by subjective factors. Features extracted based on the DLR algorithm and other factors associated with CK19 may have redundant information, which may lead to a decrease in model performance when training the integrated model. The effective combination of high-throughput features extracted by the DL segmentation network and clinical factors needs to be studied further.

Our study showed that the RFS of the CK19-positive group was lower than that of the CK19-negative group within 2 years, but not significantly (p=0.301), which was consistent with the results of some previous studies.[30,31] The small sample size may have also contributed to the statistical non-significance. The presence of intratumoral hemorrhage, as a poor prognostic factor for patients with HCC, implies a fast-growing and worse tumor microenvironment.[32] As an underlying key factor of recurrence, peritumoral hypointensity on HBP was recognized to be involved in the presence of microvascular invasion (MVI).[33,34] Peritumoral perfusion changes due to microvascular blockage might influence organic anion-transporting polypeptide transporters of hepatocytes around the tumor, resulting in decreased uptake of gadoxetic acid in hepatocytes around tumors with MVI. In this study, finally, these two MRI features were incorporated in the nomogram along with CK19 expression and the model achieved good performance for predicting RFS in patients

with HCC. Overall survival was not investigated in this study, which may have been affected by competing risk factors for death and subsequent treatment outcomes. Large-scale and comprehensive studies on the prognostic value of CK19 expression are warranted.

There were several limitations to our study. First, due to the heterogeneity of the tumor, it may be difficult to reflect the overall tumor using pathological sampling of CK19 expression. Our pathological sampling followed the baseline 7-point sampling scheme in the standardized pathological diagnosis guidelines for primary liver cancer to minimize the impact of selection deviation. Second, since the positivity of CK19 expression in patients with HCC was about 20%, there were limited cases of CK19-positive HCCs in this study. The sample size for prospective and external test sets was not large enough; therefore, the prediction model needs to be further optimized through large-scale and multicenter studies in future studies.

Conclusion

In conclusion, the DLR model showed good performance and robustness in predicting CK19 expression in HCC. Thus, this may be a useful non-invasive method for preoperative prediction of CK19 expression and assist clinicians in preoperative planning. Further study was warranted for more comprehensive interpretation of the results, shedding light on more accurate preoperative prediction of recurrence and prognosis in HCC.

Abbreviations

HCC, hepatocellular carcinoma; CK19, cytokeratin 19; DLR, deep learning radiomics; ICC, Intrahepatic cholangiocarcinoma; DPHCC, dual-phenotype hepatocellular carcinoma; DWI, diffusion weighted imaging; ADC, apparent diffusion coefficient; HBP, hepatobiliary phase; AP, arterial phase; PP, portal phase; CNN, convolutional neural networks; ALT, alanine aminotransferase; AST, aspartate transaminase; TBIL, total bilirubin; ALB, albumin; PT, prothrombin time; AFP, alpha fetoprotein; AUC, area under curve; ROC, receiver operator characteristic; CI, confidence interval; OR, odds ratio; HR, hazard ratio; RFS, recurrence-free survival; MVI, microvascular invasion.

Ethics Approval and Informed Consent

This study was approved by the institutional review board of The First Affiliated Hospital, Sun Yat-sen University as a retrospective study, and the requirement for informed consent was waived. All patient data accessed complied with relevant data protection and privacy regulations. Following the national legislation and institutional requirements, written informed consent for participation was not required for this study.

Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agreed to be accountable for all aspects of the work.

Funding

This work was funded by National Natural Science Foundation of China (81971684, 81771908), Natural Science Foundation of Guangdong Province (No. 2020A1515010571), Shenzhen – Hong Kong Institute of Brain Science – Shenzhen Fundamental Research Institutions (No. 2019SHIBS0003), Shenzhen Science and Technology Project (JCYJ20200109114014533) and SZU Top Ranking Project – Shenzhen University (860/000002100108).

Disclosure

The authors declare no conflict of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Akoad ME, Pomfret EA. Surgical resection and liver transplantation for hepatocellular carcinoma. *Clin Liver Dis*. 2015;19(2):381–399. doi:10.1016/j.cld.2015.01.007
3. Colecchia A, Schiumerini R, Cucchetti A, et al. Prognostic factors for hepatocellular carcinoma recurrence. *World J Gastroenterol*. 2014;20(20):5935–5950. doi:10.3748/wjg.v20.i20.5935
4. Mehrpouya M, Pourhashem Z, Yardehnavi N, Oladnabi M. Evaluation of cytokeratin 19 as a prognostic tumoral and metastatic marker with focus on improved detection methods. *J Cell Physiol*. 2019;234(12):21425–21435. doi:10.1002/jcp.28768
5. Shibuya M, Kondo F, Sano K, Takada T, Asano T. Immunohistochemical study of hepatocyte, cholangiocyte and stem cell markers of hepatocellular carcinoma. *J Hepato Bil Pan Sci*. 2011;18(4):537–543.
6. Lu XY, Xi T, Lau WY, et al. Hepatocellular carcinoma expressing cholangiocyte phenotype is a novel subtype with highly aggressive behavior. *Ann Surg Oncol*. 2011;18(8):2210–2217. doi:10.1245/s10434-011-1585-7

7. Zhuo J, Lu D, Tan W, Zheng S, Shen Y, Xu X. CK19-positive hepatocellular carcinoma is a characteristic subtype. *J Cancer*. 2020;11(17):5069–5077. doi:10.7150/jca.44697
8. Govaere O, Komuta M, Berkers J, et al. Keratin 19: a key role player in the invasion of human hepatocellular carcinomas. *Gut*. 2014;63(4):674–685. doi:10.1136/gutjnl-2012-304351
9. Nault JC, Villanueva A. Intratumor molecular and phenotypic diversity in hepatocellular carcinoma. *Clin Cancer Res*. 2015;21(8):1786–1788. doi:10.1158/1078-0432.CCR-14-2602
10. Phongkitkarun S. Added value of hepatobiliary phase gadoxetic acid-enhanced MRI for diagnosing hepatocellular carcinoma in high-risk patients. *World J Gastroenterol*. 2013;19(45):8357. doi:10.3748/wjg.v19.i45.8357
11. Feng S, Jia Y, Liao B, et al. Preoperative prediction of microvascular invasion in hepatocellular cancer: a radiomics model using Gd-EOB-DTPA-enhanced MRI. *Eur Radiol*. 2019;29(9):4648–4659. doi:10.1007/s00330-018-5935-8
12. Dong Z, Huang K, Liao B, et al. Prediction of sorafenib treatment–related gene expression for hepatocellular carcinoma: preoperative MRI and histopathological correlation. *Eur Radiol*. 2019;29(5):2272–2282. doi:10.1007/s00330-018-5882-4
13. 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. doi:10.1148/radiol.2017162846
14. Hu X, Wang W, Yang L, et al. MR features based on LI-RADS identify cytokeratin 19 status of hepatocellular carcinomas. *Eur J Radiol*. 2019;113:7–14. doi:10.1016/j.ejrad.2019.01.036
15. Parekh VS, Jacobs MA. Deep learning and radiomics in precision medicine. *Expert Rev Precis Med Drug Dev*. 2019;4(2):59–72. doi:10.1080/23808993.2019.1585805
16. 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):1232–1242. doi:10.1007/s00330-019-06414-2
17. 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):3004–3014. doi:10.1007/s00330-019-06585-y
18. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep learning based radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep*. 2017;7(1):1–11.
19. Hamm CA, Wang CJ, Savic LJ, et al. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol*. 2019;29(7):3338–3347. doi:10.1007/s00330-019-06205-9
20. McBee MP, Awan OA, Colucci AT, et al. Deep learning in radiology. *Acad Radiol*. 2018;25(11):1472–1480. doi:10.1016/j.acra.2018.02.018
21. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. *Radiology*. 2018;286(3):887–896. doi:10.1148/radiol.2017170706
22. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun*. 2020;11(1):1–9.
23. Wang K, Lu X, Zhou H, et al. Deep learning radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut*. 2019;68(4):729–741. doi:10.1136/gutjnl-2018-316204
24. Huang B, Tian J, Zhang H, et al.Deep semantic segmentation feature-based radiomics for the classification tasks in medical image analysis. *IEEE J Biomed Health*. 2020;1.
25. Park HJ, Kim YK, Park MJ, Lee WJ. Small intrahepatic mass-forming cholangiocarcinoma: target sign on diffusion-weighted imaging for differentiation from hepatocellular carcinoma. *Abdom Imaging*. 2013;38(4):793–801. doi:10.1007/s00261-012-9943-x
26. Ni T, Shang XS, Wang WT, Hu XX, Zeng MS, Rao SX. Different MR features for differentiation of intrahepatic mass-forming cholangiocarcinoma from hepatocellular carcinoma according to tumor size. *Br J Radiol*. 2018;91(1088):20180017. doi:10.1259/bjr.20180017
27. Kim H, Choi GH, Na DC, et al. Human hepatocellular carcinomas with "Stemness"-related marker expression: keratin 19 expression and a poor prognosis. *Hepatology*. 2011;54(5):1707–1717. doi:10.1002/hep.24559
28. Aishima S, Nishihara Y, Kuroda Y, et al. Histologic characteristics and prognostic significance in small hepatocellular carcinoma with biliary differentiation: subdivision and comparison with ordinary hepatocellular carcinoma. *Am J Surg Pathol*. 2007;31(5):783–791. doi:10.1097/01.pas.0000213421.53750.0a
29. Lu Y, Zhu M, Li W, et al. Alpha fetoprotein plays a critical role in promoting metastasis of hepatocellular carcinoma cells. *J Cell Mol Med*. 2016;20(3):549–558. doi:10.1111/jcmm.12745
30. Lee S, Lee J, Na G, You Y, Kim D. Immunohistochemical markers for hepatocellular carcinoma prognosis after liver resection and liver transplantation. *Clin Transplant*. 2017;31(1):e12852. doi:10.1111/ctr.12852
31. Lee JI. Prognosis of hepatocellular carcinoma expressing cytokeratin 19: comparison with other liver cancers. *World J Gastroenterol*. 2012;18(34):4751. doi:10.3748/wjg.v18.i34.4751
32. Kim JM, Kwon CHD, Joh J, et al. Nomograms in hepatectomy patients with hepatitis b virus-related hepatocellular carcinoma. *J Gastrointest Surg*. 2019;23(8):1559–1567. doi:10.1007/s11605-018-04074-z
33. Huang M, Liao B, Xu P, et al. Prediction of microvascular invasion in hepatocellular carcinoma: preoperative Gd-EOB-DTPA-dynamic enhanced MRI and histopathological correlation. *Contrast Media Mol Imaging*. 2018;2018:1–9.
34. Lee S, Kim SH, Lee JE, Sinn DH, Park CK. Preoperative gadoxetic acid-enhanced MRI for predicting microvascular invasion in patients with single hepatocellular carcinoma. *J Hepatol*. 2017;67(3):526–534. doi:10.1016/j.jhep.2017.04.024