



Differentiation combined hepatocellular and cholangiocarcinoma from intrahepatic cholangiocarcinoma based on radiomics machine learning

Jun Zhang¹, Zixing Huang¹, Likun Cao², Zhen Zhang¹, Yi Wei¹, Xin Zhang³, Bin Song¹

¹Department of Radiology, West China Hospital, Sichuan University, Chengdu 610041, China; ²Department of Radiology, Peking Union Medical College Hospital (Dongdan Campus), Beijing 100730, China; ³Pharmaceutical Diagnostic team, GE Healthcare, Life Sciences, Beijing 100176, China

Contributions: (I) Conception and design: B Song, J Zhang; (II) Administrative support: B Song; (III) Provision of study materials or patients: B Song, J Zhang; (IV) Collection and assembly of data: J Zhang, Z Zhang, L Cao; (V) Data analysis and interpretation: J Zhang, Z Huang, Y Wei, X Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Bin Song. Department of Radiology, West China Hospital, Sichuan University, No. 37, Guoxue Alley, Chengdu 610041, China. Email: songlab_radiology@163.com.

Background: Combined hepatocellular and cholangiocarcinoma (CHC) and intrahepatic cholangiocarcinoma (ICC) are hard to identify in clinical practice preoperatively. This study looked to develop and confirm a radiomics-based model for preoperative differentiation CHC from ICC.

Methods: The model was developed in 86 patients with ICC and 46 CHC, confirmed in 37 ICC and 20 CHC, and data were collected from January 2014 to December 2018. The radiomics scores (Radscores) were built from radiomics features of contrast-enhanced computed tomography in 12 regions of interest (ROI). The Radscore and clinical-radiologic factors were integrated into the combined model using multivariable logistic regression. The best-combined model constructed the radiomics-based nomogram, and the performance was assessed concerning its calibration, discrimination, and clinical usefulness.

Results: The radiomics features extracted from tumor ROI in the arterial phase (AP) with preprocessing were selected to build Radscore and yielded an area under the curve (AUC) of 0.800 and 0.789 in training and validation cohorts, respectively. The radiomics-based model contained Radscore and 4 clinical-radiologic factors showed the best performance (training cohort, AUC =0.942; validation cohort, AUC =0.942) and good calibration (training cohort, AUC =0.935; validation cohort, AUC =0.931).

Conclusions: The proposed radiomics-based model may be used conveniently to the preoperatively differentiate CHC from ICC.

Keywords: Combined hepatocellular and cholangiocarcinoma (CHC); intrahepatic cholangiocarcinoma (ICC); machine learning; radiomics

Submitted Dec 11, 2019. Accepted for publication Jan 14, 2020.

doi: 10.21037/atm.2020.01.126

View this article at: <http://dx.doi.org/10.21037/atm.2020.01.126>

Introduction

Combined hepatocellular and cholangiocarcinoma (CHC) account for 0.4–14.2% of primary hepatic malignancies (1,2), originate from hepatic progenitor cells (3) and show both hepatocytic and cholangiocytic differentiation (4,5). In contrast, intrahepatic cholangiocarcinoma (ICC) is the second most common hepatic malignancy after

hepatocellular carcinoma (6) and histologically derived from the biliary epithelial cell.

Although the 7th and 8th editions of the American Joint Committee on Cancer Staging and the International Union for Cancer Control tumor-node-metastasis staging group CHC and ICC into one classification, the distinctions between CHC and ICC clinical characteristics remain

unclear. Some studies proved clinical features, and prognosis of CHC is markedly different from those of ICC, while others reported CHC and ICC show similar genetic (7) and clinical characteristics (8). Lee *et al.* reported that the patients with CHC had 18 months' median overall survival which was shorter than patients with ICC (9). A recent study proved that the rate of recurrence-free survival at 5 years was significantly different between CHC and ICC, with 0% and 37.7%, respectively (10). Surgical resection is the sole therapy choice that offers a chance of long-term survival for patients with ICC, while patients with CHC experience more frequently recurrence than patients with ICC after curative resection (9). However, drug-targeted therapy may bring desirable results to patients with CHC. Futsukaichi *et al.* reported a patient with CHC successfully treated with sorafenib (11). Liver transplantation is also one of choice for patients with ICC or CHC (12). Nonetheless, ICC patient surgical results following liver transplantation are quite poor due to higher recurrence and reduced survival (13). Liver transplantation may improve surgical outcomes of patients with CHC than ICC patients (14). Therefore, it has significant clinical implications to differentiate CHC from ICC preoperatively accurately.

In clinical practice, using needle biopsy to differentiate CHC from ICC may increase the possibility of tumors metastasis, so its clinical application is limited. Additionally, radiographic imaging plays a vital role in the diagnosis and differential diagnosis of liver tumors. However, CHC can share some similar radiographic characteristics with ICC (15,16), so it is challenging to differentiate diagnosis preoperatively through those methods (17). For example, there have been some studies regarding CHC's radiographic characteristics (18-25), but these studies' results are contradictory. It poses a big challenging in clinical practice.

Radiomics is a method for imaging analysis based on data-mining and statistical analysis of high-throughput imaging features (26). In recent years, it successfully showed favorable abilities in clinical management (27). Xu *et al.* reported that integrating clinical-radiologic and radiomic features can predict microvascular invasion and outcome in patients with hepatocellular carcinoma (26). Ji *et al.* showed that the radiomics-based model could predict lymph node metastasis and survival outcomes in patients with biliary tract cancer (28). However, to our knowledge, there is still a lack of radiomics-based methods or tools for differentiation between CHC and ICC. Our research aimed to establish

and validate a radiomics-based model for ICC preoperative differentiation of CHC.

Methods

Patients

The institutional review board approved this study in our hospital. All procedures were performed following the 1975 Helsinki declaration and its later amendments.

The medical records were reviewed from January 2014 to December 2018 to identify all patients who underwent resection for CHC and ICC in our hospital. The inclusion criteria were as follows: (I) postoperatively pathological confirmation for CHC and ICC; (II) CHC is a classical type according to WHO classification criteria; (III) cholangiocarcinoma-dominant in CHC; (IV) preoperative CT imaging, including arterial phase (AP) and portal vein phase (PVP), were performed in our radiology unit; (V) without anti-tumor therapy before CT examination; (VI), routine preoperative laboratory examinations, including total bilirubin (TB), serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), γ -glutamyl transpeptidase (GGT), serum a-fetoprotein (AFP), carcinoembryonic antigen (CEA), cancer antigen 199 (CA-199), and hepatitis B surface antigen. The exclusion criteria were as follows: (I) biopsy confirmation for CHC and ICC without operation; (II) treated with other diseases before CT examination; (III) imaging with artifacts or incomplete; (IV), deficient of clinical data. Finally, a total of 189 patients (ICC =123; CHC =66) were enrolled during the study period. The flowchart of patients' enrollment was shown in *Figure 1*.

Imaging acquisition

All patients' scans were performed using multi-slice CT (Philips Brilliance64, Philips Medical Systems, Best, The Netherlands; Siemens SOMATOM Definition Flash, Siemens Healthcare, Erlangen, Germany) in our institution. The scanning parameters were as follows: voltage, 120 kV; current, 200–250 mAs; slice thickness, 5 mm. After plain scanning was completed, a nonionic iodine contrast agent (Omnipaque 300, GE Healthcare, Ireland) was injected into the antecubital vein with a dose of 1.2 mL per kilogram of body weight at a rate of 3.0 mL/s, followed by a saline flush of 20 mL. AP and PVP scanning started at 25 and 60 s after

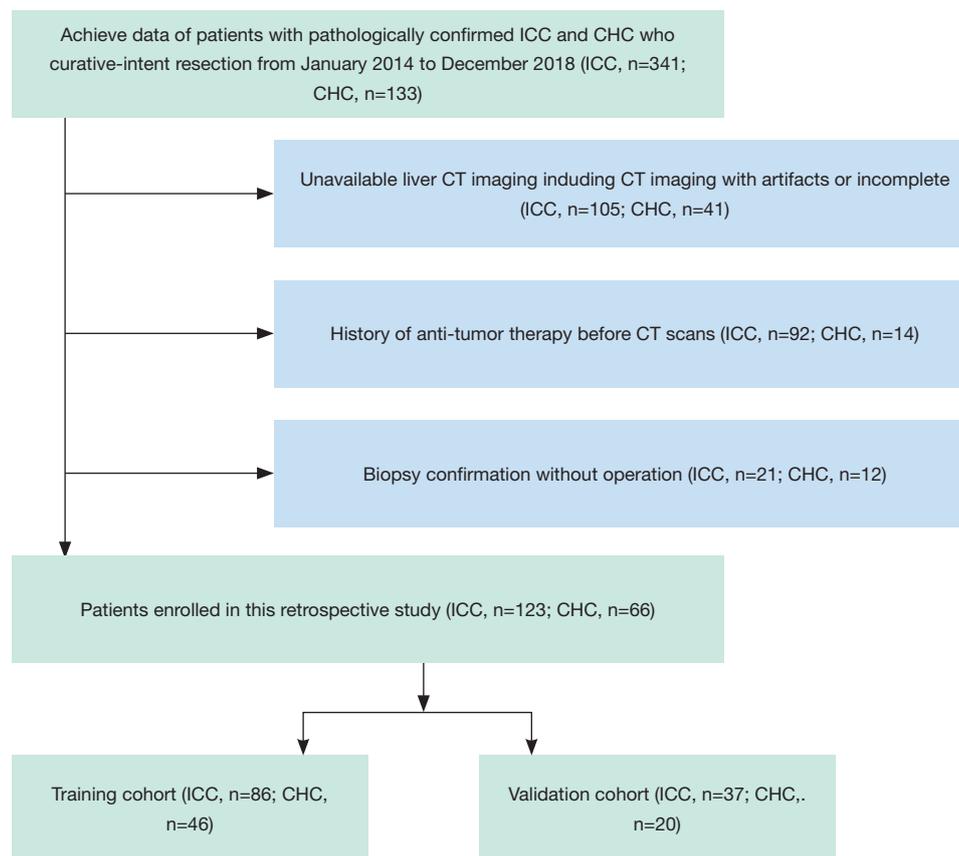


Figure 1 Flowchart of ICC and CHC patients' enrollment. ICC, intrahepatic cholangiocarcinoma; CHC, combined hepatocellular and cholangiocarcinoma.

the contrast medium was injected.

Imaging analysis

All CT images were reviewed in picture archiving and communication system (Syngo-Imaging, version VB36A; Siemens Medical Solutions, PACS) and were retrospectively interpreted by two radiologists (Yi Wei, six years of liver imaging experience, and Zixing Huang, fourteen years of liver imaging experience), both of whom were blinded to pathological outcomes and clinical factors. Before the study, they followed a complete day lecture-based and firsthand instruction session that explained in detail a structural liver imaging score (LI-score) system. During the course, 100 cases randomly selected from the institutional PACS database (not included in this study cohort) were reviewed individually, scored, and then reviewed as a group. The two radiologists independently reviewed all radiologic imaging

individually first and then reviewed them together. Any discrepancies between two readers were discussed until a final consensus was generated. Individual scores were used for the calculation of inter-observer agreement, and consensus scores were used for classifying performance.

The LI-score, most referring to Xu *et al.* (26), were obtained depending on characteristics of tumor imaging: (I) lobes with lesions involvement (0, one lobe; 1, more than one lobes); (II) numbers of lesions (0, single; 1, multiple); (III) maximum diameter of the lesions (0, ≤ 5 ; 1, > 5 cm); (IV) lesions margin (0, well-defined; 1, ill-defined); (V) hypodense halos (0, absent; 1, present); (VI) enhancement patterns (1, wash-in and wash-out; 2, rim-like enhancement with progressive centripetal filling; 3, weak or invisible enhancement with delay enhancement); (VII) peritumoral bile duct dilatation (0, absent; 1, present); (VIII) intratumor necrosis (0, absent; 1, present); (IX) intratumor stones (0, absent; 1, present); (X) intratumor vascularity (0, absent;

1,present); (XI) peritumoral vascular invasion: defined as vessel stenosis, occlusion or unclear boundary with lesion (0, absent; 1, present); (XII) intravascular tumor thrombus: defined as intravascular low density (0, absent; 1, present); (XIII) capsular retraction (0, absent; 1, present); (XIV) hepatic lobe atrophy (0, absent; 1, present). The largest lesion was analyzed when there were multiple lesions.

Histopathologic analysis

Another researcher (Zhen Zhang, two years of liver imaging experience) who did not take part in imaging evaluation independently searched the pathological database. The histologic parameters ordinarily included tumor differentiation, satellite lesions, lymph node metastasis, adjacent organ invasion, and cirrhosis. The lowest differentiation was recorded when there were different differentiation components in the tumor.

Tumors segmentation

As lymphatic invasion and tumor-infiltration might spatial heterogeneously distributed, three regions were considered for tumor segmentation to involve more infiltration features and better reflect tumor heterogeneity (29). Here, the tumor, peritumoral area (a 2-mm wide band of outer tissue adjacent to the outer tumor margin), and the tumor plus peritumoral area were delineated, respectively. Each area was segmented from the original image (5 mm AP and 5 mm PVP) and preprocessed image. For obtaining tumor regions of interest (ROI^{tumor}), tumors were manually segmented in three layers for each lesion, along the tumor contour at the largest section, up and down from this section 5 mm distance respectively, by using an open-source imaging platform (ITK-SNAP, version 3.6.0; www.itknap.org). The peritumoral regions of interest ($ROI^{peritumoral}$) were automatically reconstructed according to the ROI^{tumor} contours with an erosion and dilation algorithm and the process was performed by AK software (Artificial Intelligent Kit, V3.2.0, GE Healthcare). Additionally, the tumor plus peritumoral regions of interest ($ROI^{tumor+peritumoral}$) was computed based on the ROI^{tumor} and $ROI^{peritumoral}$. Radiomics features were extracted from each ROI, and the imaging preprocessing methods using AK software include the resampling with 1×1 pixels and noise reduction using classic Gaussian filter.

The ROIs segmentation was performed by two radiologists to assess the intra-observer reliability (Jun

Zhang, reader 1, fifteen years of liver imaging experience; Likun Cao, reader 2, three years of liver imaging experience) in a blinded fashion; both did not participate in imaging analysis and were blinded to pathological outcomes. The reader 1 and reader 2 performed segmentation of lesions in one week, respectively, the readout sessions were held within one month. The reliability was calculated by using the intraclass correlation coefficient, and the values greater than 0.8 were selected for later investigation.

Radiomics signature

A total of 396 radiomic features for each CT imaging modality were calculated for each ROI (Table S1). The extracted radiomic features were standardized using the zero-mean normalization method, and the equation is

$$X_n^* = \frac{x_n - \mu}{\sigma} \quad [1]$$

where x_n is the value of feature N, μ is the average of all feature N, and σ is the standard deviation.

The potential combination of the radiomic features with ICC and CHC was first assessed in training cohort by using a Mann-Whitney U test; and the least absolute shrinkage and selection operator (LASSO) method, which is suitable for regression of high-dimensional data, was used to select the most useful predictive features. A logistic regression model was developed based on radiomic features selected, and a radiomic score (Radscore) was calculated for each patient via the Radiomics model (R model).

Development and validation of models

For clinical factors and liver imaging data, the univariable analysis with a generalized linear model (GLM) was performed to predictor selection, respectively, and the clinical score (C-score) and LI-score were computed using the same method mentioned above. The multivariable analysis, association with Radscore, C-score, and LI-score, was then conducted to select the significant predictors using a backward step-down choice process with the Akaike information criterion (AIC). The final prediction model was selected from the combined model, which had the highest predictive ability and clinical benefit.

The development of the nomogram was based on proportionally converting regression coefficients of each predictor in the final model to a 0- to a 100-point scale. The effect of the variable with the highest \bar{a} coefficient

(absolute value) is assigned 100 points. The points are added across independent variables to derive total points, which are converted to predicted probabilities (Pi). The workflow of construction and validation of the models was shown in *Figure 2*.

Statistical analysis

Continuous variables were expressed as mean \pm standard deviation or median \pm Interquartile Range(IQR), compared by the *t*-test or Mann-Whitney *U* test as proper. Categorical variables were expressed as numbers and percentage, compared by the chi-square test or Fisher exact test. The inter-observer agreement was applied to assess the reliability of imaging analysis using the Kappa test; 0–0.2 represents slight, 0.21–0.40: fair, 0.41–0.60: moderate, 0.61–0.80: substantial, 0.81–1: excellent.

The discrimination performance of each prediction model was quantified by the receiver operating characteristic curve (ROC) and the area under the curve (AUC). The Delong's test was used to compare the statistical differences between any two AUCs. The decision curve analysis (DCA) was applied to evaluate the clinical usefulness and the net benefits of the models. The DCA calculates the net benefits of a model as the difference between the true-positive and false-positive rates, weighted by the odds of the selected threshold probability of risk. In addition, for nomograms, the calibration plots were generated to evaluate the calibration of the nomogram via bootstrapping resamples using the Hosmer-Lemeshow goodness-of-fit test.

Statistical analysis was performed using the R software (version 3.5.1, R Project for Statistical Computing, <http://www.r-project.org>). Two-sided $P < 0.05$ was considered statistically significant.

Results

Clinicopathologic characteristics

The 189 patients (ICC =123; CHC =66) were grouped into training cohort (n=132, ICC =86, CHC =46) and validation cohort (n=57, ICC =37, CHC =20) randomly. There was no significant difference in clinicopathologic characteristics between the training and validation cohorts (all $P > 0.05$) (*Table 1*). Gender, age, AFP, cirrhosis, and hepatitis B were

a significant difference between ICC and CHC based on univariable analysis (all $P < 0.05$).

Imaging analysis

Kappa test showed substantial performance for enhancement patterns of tumors and excellent for remain imaging features. No significant imaging features can be found between the training cohort and validation cohort (all $P > 0.05$) (*Table 2*). However, intratumor necrosis and intravascular tumor thrombus were a significant difference between ICC and CHC based on univariable analysis (all $P < 0.05$).

Development and validation of clinical model and liver imaging model

The clinical model (C model) contained gender [odds ratio (OR): 0.24; 95% confidence interval (95% CI): 0.07–0.8], AFP (OR: 0.06, 95% CI: 0.02–0.14), cirrhosis (OR: 0.44; 95% CI: 0.15–1.28) and hepatitis B (OR: 0.23; 95% CI: 0.07–0.72) in training cohort by multivariable analysis (all $P < 0.05$); furtherly be verified in the validation cohort. The liver imaging model (L model) contained intratumor necrosis (OR: 0.5; 95% CI: 0.24–1.06) and intravascular tumor thrombus (OR: 0.09; 95% CI: 0.01–0.81) in training cohort by multivariable analysis (all $P < 0.05$); also be verified in the validation cohort.

The AUC of the C model was 0.893 (95% CI: 0.831–0.954) and 0.858 (95% CI: 0.762–0.954), and the L model was 0.623 (95% CI: 0.532–0.715) and 0.671 (95% CI: 0.538–0.804) in training cohort and validation cohorts, respectively. The accuracy, sensitivity, and specificity of each model were shown in *Table 3*. The formulas for each model are shown in the Supplementary file.

Radiomics analysis

The AUC of each R model is close to or higher than 0.7 both in the training cohort and validation cohort (*Table S2*), indicating the radiomics features exacted from 12 ROIs all had high discrimination ability and accuracy (ACC). However, the ROC analysis demonstrated that AUCs were not significantly different among 12 R models ($P > 0.05$) (*Figure S1A,B*), suggesting imaging preprocessing did not

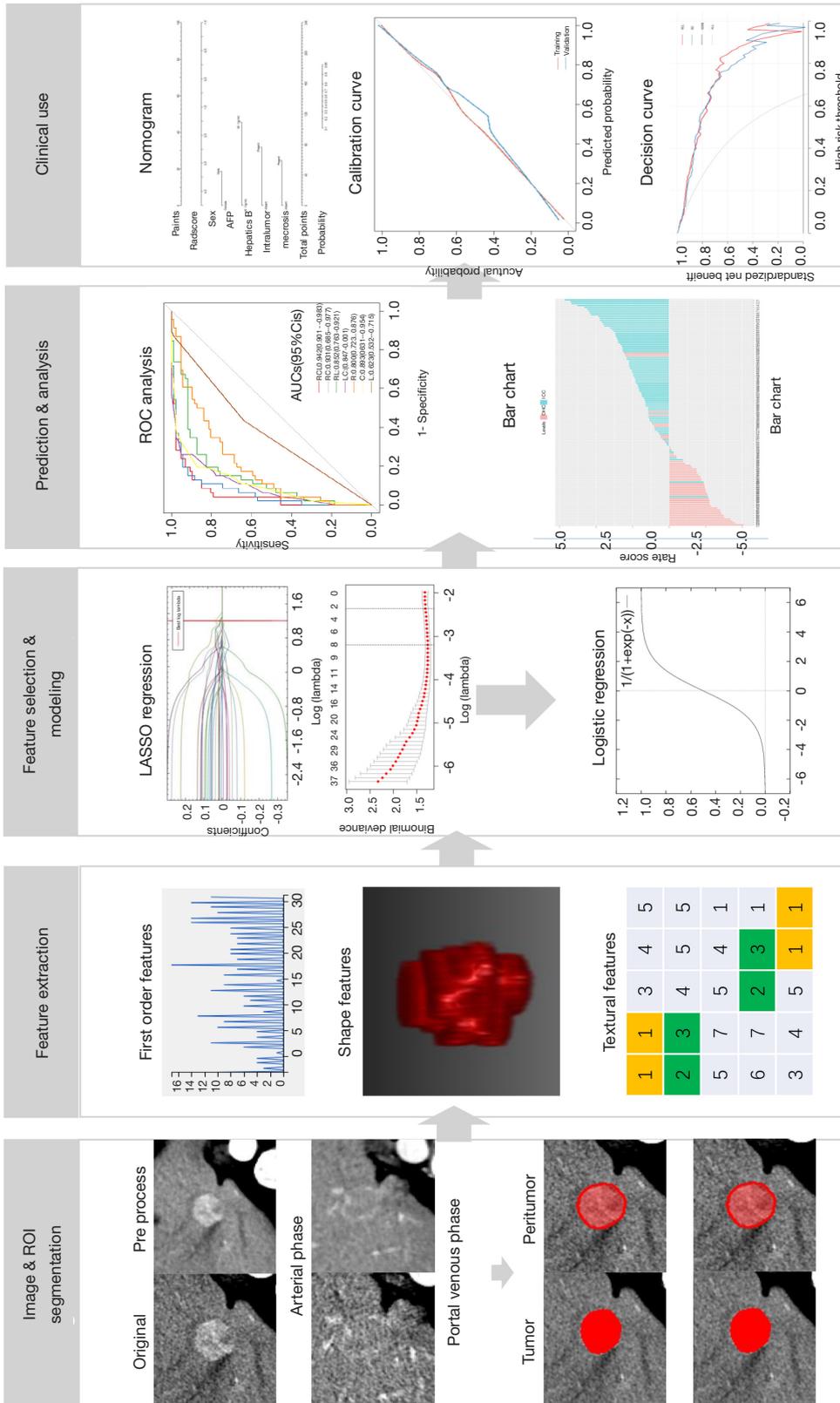


Figure 2 Workflow of construction and validation of the models.

Table 1 The clinicopathologic characteristics

Variable	ICC			CHC		
	Training (n=86)	Validation (n=37)	P value	Training (n=46)	Validation (n=20)	P value
Gender			0.2			0.4
Male	42 (48.8)	13 (35.1)		39 (84.8)	19 (95.0)	
Female	44 (51.2)	24 (64.9)		7 (15.2)	1 (5.0)	
Age	59.5±10.3	55.6±9.76	0.05	52.8±9.7	51.7±12.6	0.7
Tumors differentiations			0.1			0.9
Well	1 (1.2)	1 (2.7)		3 (6.5)	1 (5.0)	
Moderate	60 (69.8)	19 (51.4)		25 (54.3)	10 (50.0)	
Poor	25 (29.0)	17 (45.9)		18 (39.2)	9 (45.0)	
Satellite lesions			0.9			0.1
Absent	72 (83.7)	32 (86.5)		35 (76.1)	19 (95.0)	
Present	14 (16.3)	5 (13.5)		11 (23.9)	1 (5.0)	
Lymph node metastasis			0.7			1
Absent	65 (75.6)	30 (81.1)		41 (89.1)	18 (90.0)	
Present	21 (24.4)	7 (18.9)		5 (10.9)	2 (10.0)	
Adjacent organ invasion			1			0.4
Absent	83 (96.5)	35 (94.6)		42 (91.3)	20 (100.0)	
Present	3 (3.5)	2 (5.4)		4 (8.7)	0 (0)	
TB (umol/L)	12.2 (4.1–85.6)	11.9 (4–82.4)	0.9	11.4 (6.3–88.8)	17.1 (4.9–261)	0.2
ALT (IU/L)	23 (8–981)	24 (7–379)	0.5	34.5 (14–199)	37 (16–106)	0.9
AST (IU/L)	28 (13–765)	29 (14–367)	0.8	36 (19–165)	38 (19–114)	0.8
GGT (IU/L)	63 (16–772)	65 (14–423)	0.5	67 (19–472)	77 (20–274)	0.3
CEA (ng/mL)			0.05			0.07
≥3.4	36 (41.9)	8 (21.6)		19 (41.3)	3 (15.0)	
<3.4	50 (58.1)	29 (78.4)		27 (58.7)	17 (85.0)	
CA-199 (U/mL)			0.8			0.9
≥22	59 (68.6)	27 (73.0)		25 (54.3)	12 (60.0)	
<22	27 (31.4)	10 (27.0)		21 (45.7)	8 (40.0)	
AF (ng/mL)			0.4			0.2
≥8	10 (11.6)	7 (18.9)		32 (69.6)	10 (50.0)	
<8	76 (88.4)	30 (81.1)		14 (30.4)	10 (50.0)	
Cirrhosis			0.5			1
Absent	73 (84.9)	29 (78.4)		25 (54.3)	10 (50.0)	
Present	13 (15.1)	8 (21.6)		21 (45.7)	10 (50.0)	
Hepatitis B			1			1
Absent	54 (62.8)	24 (64.9)		7 (15.2)	3 (15.0)	
Present	32 (37.2)	13 (35.1)		39 (84.8)	17 (85.0)	

Data are shown as n (%) or median (IQR). CHC, combined hepatocellular and cholangiocarcinoma; ICC, intrahepatic cholangiocarcinoma; TB, total bilirubin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ -glutamyl transpeptidase; CEA, carcinoembryonic antigen; CA-199, cancer antigen 199.

Table 2 The CT imaging characteristics

Variable	ICC			CHC		
	Training (n=86)	Validation (n=37)	P value	Training (n=46)	Validation (n=20)	P value
Lobes with lesions involvement			0.5			0.5
One lobe	73 (84.9)	29 (78.4)		39 (84.5)	15 (75.0)	
More than one lobe	13 (15.1)	8 (21.6)		15 (15.4)	5 (25.0)	
Numbers of lesions			0.2			1
Single	64 (74.4)	32 (86.5)		29 (63.0)	13 (65.0)	
Multiple	22 (25.6)	5 (13.5)		17 (37.0)	7 (35.0)	
Maximum diameter of the lesions			0.6			0.6
≤5 cm	32 (37.2)	11 (29.7)		20 (54.1)	11 (55.0)	
<5 cm	54 (62.8)	26 (70.3)		26 (45.9)	9 (45.0)	
Lesions margin			0.2			0.7
Well-defined	32 (37.2)	19 (51.4)		12 (26.1)	7 (35.0)	
Ill-defined	54 (62.8)	18 (48.6)		34 (73.9)	13 (65.0)	
Hypodense halos			1			1
Absent	83 (96.5)	36 (97.3)		45 (97.8)	20 (100.0)	
Present	3 (3.4)	1 (2.7)		1 (2.2)	0 (0)	
Peritumoral bile duct dilatation			0.7			1
Absent	61 (73.5)	28 (75.7)		38 (82.6)	17 (85.0)	
Present	25 (26.5)	9 (24.3)		8 (17.4)	3 (15.0)	
Peritumoral vascular invasion			1			1
Absent	63 (75.9)	27 (73.0)		34 (73.9)	14 (70.0)	
Present	23 (24.1)	10 (27.0)		12 (26.1)	6 (30.0)	
Intratumor vascularity			0.5			0.4
Absent	54 (62.8)	20 (54.1)		25 (54.3)	14 (70.0)	
Present	32 (37.2)	17 (45.9)		21 (45.7)	6 (30.0)	
Enhancement patterns			0.4			0.6
I type	4 (4.7)	2 (5.4)		11 (23.9)	3 (15.0)	
II type	40 (46.5)	22 (59.5)		13 (28.3)	5 (25.0)	
III type	42 (48.8)	13 (35.1)		22 (47.8)	12 (60.0)	
Intratumor necrosis			1			0.6
Absent	56 (65.1)	25 (67.6)		21 (45.7)	7 (35.0)	
Present	30 (34.9)	12 (32.4)		25 (54.3)	13 (65.0)	
Intratumor stones			1			1
Absent	84 (97.7)	36 (97.3)		46 (100.0)	20 (100.0)	
Present	2 (2.3)	1 (2.7)		0 (0)	0 (0)	

Table 2 (continued)

Table 2 (continued)

Variable	ICC			CHC		
	Training (n=86)	Validation (n=37)	P value	Training (n=46)	Validation (n=20)	P value
Intravascular tumor thrombus			1			0.6
Absent	85 (98.8)	37 (100.0)		40 (87.0)	19 (95.0)	
Present	1 (1.2)	0 (0)		6 (13.0)	1 (5.0)	
Capsular retraction			0.3			0.4
Absent	69 (80.2)	26 (70.3)		39 (84.8)	19 (95.0)	
Present	17 (19.8)	11 (29.7)		7 (15.2)	1 (5.0)	
Hepatic lobe atrophy			1			1
Absent	78 (90.7)	34 (91.9)		45 (97.8)	20 (100.0)	
Present	8 (9.3)	3 (8.1)		1 (2.2)	0 (0.0)	

Data are the number of tumors, with percentages in parentheses. CHC, combined hepatocellular and cholangiocarcinoma; ICC, intrahepatic cholangiocarcinoma. Enhancement patterns; I type, wash-in and wash-out; II type, rim-like enhancement with progressive centripetal filling; III type, weak or invisible enhancement with delay enhancement.

Table 3 The results of model analysis

Models	Training cohort				Validation cohort				Cut-off values
	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	
Clinical	0.893	0.848	0.804	0.872	0.858	0.772	0.700	0.811	0.540
Liver imaging	0.623	0.614	0.565	0.640	0.671	0.667	0.650	0.676	0.665
Radiomics	0.800	0.742	0.739	0.744	0.789	0.754	0.700	0.784	0.590
RC	0.931	0.894	0.848	0.919	0.911	0.842	0.700	0.919	0.546
RL	0.852	0.818	0.804	0.826	0.841	0.789	0.750	0.811	0.572
LC	0.904	0.879	0.739	0.953	0.896	0.789	0.650	0.865	0.409
RCL	0.942	0.886	0.870	0.895	0.942	0.877	0.850	0.892	0.661

Lesser than cut-off value means combined hepatocellular and cholangiocarcinoma; ROC analysis demonstrated that the AUC of RCL model was statistically significant compared to the other models except for the RC model. RC, radiomics-clinical; RL, radiomics-liver imaging; LC, liver imaging-clinical; RCL, radiomics-clinical-liver imaging; AUC, area under the curve; ACC, accuracy; SEN, sensitivity; SPE, specificity.

exert influence on results of the radiomics analysis. The model constructed from ROI^{tumor} in AP with preprocessing was selected as the final R model due to the best performance. The formula for the R model is shown in the Supplementary file.

Performance of predictive models

The Radscore, C-score, and LI-score were identified as essential predictors by multivariate regression analysis (all

$P < 0.05$), and the single-factor models and combined models were established respectively. The AUC of the R model was slightly inferior to that of the C model, while the AUC of the L model was the lowest among single-factor models. The best performance of the model was the Radiomic-Clinical-Liver image model (RCL model) (Table 3), which reached the higher AUC [training cohort, 0.942 (95% CI, 0.901–0.983); validation cohort, 0.942 (95% CI, 0.886–0.998)]. The Delong's test illustrated that there was a significant difference between the RCL model and

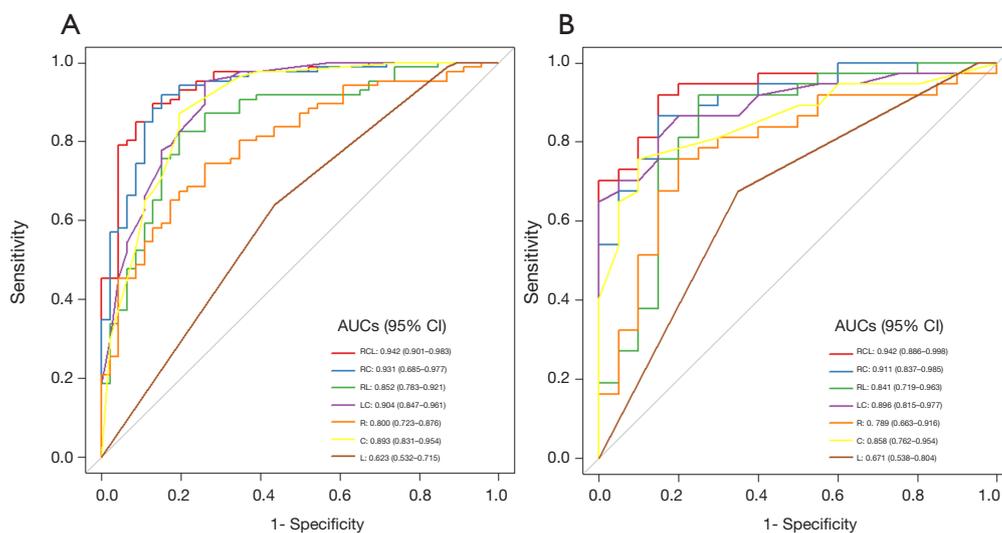


Figure 3 Receiver operating characteristic analysis for 4 combined models and 3 single-factor models. (A) Training cohort; (B) validation cohort. R, radiomics model. C, Clinical model. L, liver imaging model; RC, radiomics-clinical model; LC, liver imaging-clinical model; RL, radiomics-liver imaging model; RCL, radiomics-clinical-liver imaging model.

other single-factor models (all $P < 0.05$), while no significant difference can be found between RCL and Radiomics-clinical models (RC model) ($P = 0.2$) (Figure 3A,B).

Moreover, the DCA was used to estimate the net benefit for each combined model, respectively. The results showed that adding the LI-scores did not improve the net benefit of the RCL model compared with the RC model ($P = 0.7$) (Figure 4), while the AUC of the RCL model was the highest (Table 3). The final predictive model was named as the RCL model. The formula for RCL model was described as:

$$Y = 3.71 + 0.951 \times \text{Radscore} - 1.159 \times \text{Sex} - 2.813 \times \text{AFP} - 1.97 \times \text{Hepatitis B} - 1.523 \times \text{intratumor necrosis}.$$

The radiomics nomogram for visualization of the final predictive model is shown in Figure 5A. The calibration curve of the model showed excellent agreement between prediction and observation in both the training (C-index 0.935) and the validation cohort (C-index 0.931) (Figure 5B). The final predictive model was right in both data sets ($P = 0.2$ and 0.9).

Discussion

In our study, we developed and confirmed a radiomics-based model for differentiation CHC from ICC. The proposed model proved excellent discrimination ability and reliability both in the training cohort and the validation cohort. It may

play an essential role in clinical practice.

In the present study, we found that hepatitis B was an independent risk factor for CHC, inconsistent with the western study, which reported CHC is less related to hepatitis B (19). A reason for this discrepancy might be the different population characteristics between two studies, showing higher rates of hepatitis B virus infection in Chinese populations than Western counterparts. Moreover, we found that AFP was an independent predictor for CHC, which agrees with a study conducted by Tian *et al.* (30). However, CA-199 had no significant difference in our study. Our results were partly not consistent with previously published results, which found simultaneous elevation of both CA199 and AFP highly indicate CHC (31). The reason for this inconsistency may be low sensitivity for diagnosis CHC using the simultaneous elevation of both AFP and CA199 (32). Furthermore, we found that CHC occurs in males more frequent; our results are comparable with earlier studies (32,33).

Radiomics plays a vital role in oncology (27,34,35), so we tried to find CHC and ICC using radiomics features. Our results showed the radiomics features had a favorable ability for differentiating CHC from ICC in the training cohort and was verified in the validation cohort, suggesting the difference of intratumor heterogeneity between CHC and ICC. However, the AUC of the R model is slightly inferior to that of the C model, suggesting radiomics were weaker

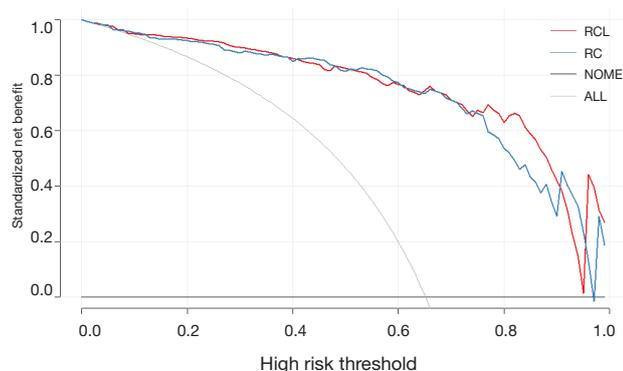


Figure 4 Decision curve analysis of combined models for predicting ICC and CHC. The y-axis measures the net benefits, wherein the red and blue line represents the RCL and RC model, respectively. The gray line represents the assumption that all patients have ICC. The black line represents the assumption that all patients have CHC. It shows that combined models are better than the treat-all-patients scheme, while if LI-scores are added, it does not result in more benefits for patients' discrimination (chi-squared =0.1, P=0.7). RCL, radiomics-clinical-liver imaging; RC, radiomics-clinical; CHC, combined hepatocellular and cholangiocarcinoma; ICC, intrahepatic cholangiocarcinoma.

compare to evidence-based clinical factors in the differential diagnosis of CHC and ICC. Our results were like studies conducted by Banerjee *et al.* (36) and Renzulli *et al.* (37). We speculated the reason for this might be the same existence component, which is biliary epithelial cell carcinoma, both in CHC and ICC, so that the radiomics features may be overlapping to a certain extent. It also may be a reason for hardly using radiographic imaging to differentiate CHC from ICC.

Besides, we also found radiomics features from $ROI^{peritumoral}$ and $ROI^{tumor+peritumoral}$ had favorable AUC and ACC in differentiation CHC from ICC, although no significant difference can be observed compared with radiomics features from ROI^{tumor} . The plausible reason may appear to tumor-infiltrating lymphocytes and tumor-associated macrophages around tumors (38); however, the density of lymphocytes and macrophages maybe not exactly similar between CHC and ICC, resulting in radiomics features difference.

We found that intratumor necrosis was an independent predictor for CHC, not consistent with the published papers (14-21), conclusions of which were also inconsistent. We

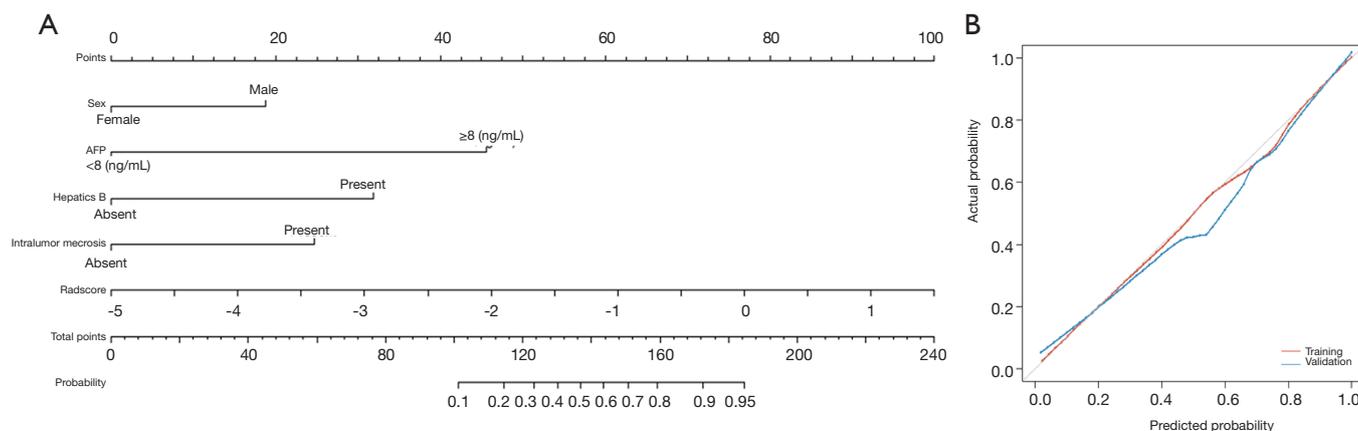


Figure 5 Nomogram of independent predictors of ICC/CHC and calibration curve of nomogram in training/validation cohorts. (A) The model presented with a nomogram scaled by the proportional regression coefficient of each predictor, the probability in nomogram describes the probability of a patient with CHC; (B) calibration curve for training and validation set. The Hosmer-Lemeshow test gave a chi-square of 2.87 (P=0.94) for the training cohort and validation cohort, indicating that the RCL model was appropriate in both data sets. RCL, radiomic-clinical-liver image model; CHC, combined hepatocellular and cholangiocarcinoma.

suggest two explanations of why radiographic imaging of CHC varies among researches. First, it may be population characteristics included in studies ranging from region to region. Second, it may be inconsistency with equipment and evaluation methods used for research. Moreover, the DCA demonstrated the net benefit of the RCL model did not be improved than the RC model, which also showing radiographic imaging plays a negligible effect on diagnosis CHC from another perspective.

Interestingly, we proved that imaging preprocessing did not influence on results of radiomics analysis. In general, the imaging in AP is superior to that in PVP as radiomics features extraction for differentiation CHC from ICC, although there is no significant difference ($P>0.05$). We guessed some reasons for these phenomena. First, imaging preprocessing may play a weak effect on imaging with thicker layers. Second, the reason for the imaging in AP slightly better than PVP may be linked to the different arterial blood flow between CHC and ICC. Further investigation is needed.

Several limitations must be considered in interpreting our results. First, a large number of patients who did not undergo surgery were excluded due to the retrospective nature of the study so that potential selection bias may exist. Second, our study was drawn from a single-center, and expanding results to other centers will need to confirm its reproducibility. Third, we only extracted two-dimensional radiomics features with a 5 mm layer from different equipment due to retrospective analysis, but it may be generalized. Fourth, we tried to extend 2 mm outside from tumor as peritumoral area, but it needs more researches considering the actual size. Fifth, delayed scanning did not use tracer technology, so the collected data may be biased. Finally, our participant sample was modest in size due to the rarity of CHC.

Conclusions

In this study, we proposed a radiomics-based model to help preoperative differentiation of CHC from ICC. By fully considering and combining the specific radiomics and clinical manifestations, the model may help clinicians to be more confident in inpatient management.

Acknowledgments

Funding: This research was supported by the Science and Technology Support Program of Sichuan Province, China

(Grant Number 2017SZ0003) and the 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University, China (Grant Number ZYJC18008).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Institutional Review Board approval (ID: 2016-297) was obtained before this retrospective study.

References

1. Jung DH, Hwang S, Hong SM, et al. Post-resection Prognosis of Combined Hepatocellular Carcinoma-Cholangiocarcinoma According to the 2010 WHO Classification. *World J Surg* 2017;41:1347-57.
2. Vienot A, Neuzillet C. Cholangiocarcinoma: the quest for a second-line systemic treatment. *Transl Cancer Res* 2019;8:S275-88.
3. Ogasawara S, Akiba J, Nakayama M, et al. Epithelial cell adhesion molecule-positive human hepatic neoplastic cells: development of combined hepatocellular-cholangiocarcinoma in mice. *J Gastroenterol Hepatol* 2015;30:413-20.
4. Hang H, Jeong S, Sha M, et al. Cholangiocarcinoma: anatomical location-dependent clinical, prognostic, and genetic disparities. *Ann Transl Med* 2019;7:744.
5. Connell LC, Harding JJ, Shia J, et al. Combined intrahepatic cholangiocarcinoma and hepatocellular carcinoma. *Chin Clin Oncol* 2016;5:66-74.
6. Petrick JL, Braunlin M, Laversanne M, et al. International trends in liver cancer incidence, overall and by histologic subtype, 1978-2007. *Int J Cancer* 2016;139:1534-45.
7. Cazals-Hatem D, Rebouissou S, Bioulac-Sage P, et al. Clinical and molecular analysis of combined hepatocellular-cholangiocarcinomas. *J Hepatol* 2004;41:292-8.
8. Jarnagin WR, Weber S, Tickoo SK, et al. Combined hepatocellular and cholangiocarcinoma: demographic, clinical, and prognostic factors. *Cancer* 2002;94:2040-6.
9. Lee JH, Chung GE, Yu SJ, et al. Long-term Prognosis of Combined Hepatocellular and Cholangiocarcinoma After Curative Resection Comparison With Hepatocellular

- Carcinoma and Cholangiocarcinoma. *J Clin Gastroenterol* 2011;45:69-75.
10. Song P, Midorikawa Y, Nakayama H, et al. Patients' prognosis of intrahepatic cholangiocarcinoma and combined hepatocellular-cholangiocarcinoma after resection. *Cancer Med* 2019;8: 5862-71.
 11. Futsukaichi Y, Tajiri K, Kobayashi S, et al. Combined hepatocellular-cholangiocarcinoma successfully treated with sorafenib: case report and review of the literature. *Clin J Gastroenterol* 2019;12:128-34.
 12. Sapisochin G, Facciuto M, Rubbia-Brandt L, et al. Liver transplantation for "very early" intrahepatic cholangiocarcinoma: International retrospective study supporting a prospective assessment. *Hepatology* 2016; 64:1178-88.
 13. Lee DD, Croome KP, Musto KR, et al. Liver transplantation for intrahepatic cholangiocarcinoma. *Liver Transpl* 2018;24:634-44.
 14. Machairas N, Stamopoulos P, Kostakis ID, et al. Mixed Hepatocellular Cholangiocarcinoma: A Review of Long-Term Outcomes Following Liver Transplantation. *Transplant Proc* 2019; 51:437-9.
 15. Saboo SS, Krajewski KM, Jagannathan JP, et al. Rapid progression of combined hepatocellular carcinoma and cholangiocarcinoma. *Cancer Imaging* 2011;11:37-41.
 16. Sagrini E, Iavarone M, Stefanini F, et al. Imaging of combined hepatocellular-cholangiocarcinoma in cirrhosis and risk of false diagnosis of hepatocellular carcinoma. *United European Gastroenterol J* 2019;7:69-77.
 17. Gera S, Ettl M, Acosta-Gonzalez G, et al. Clinical features, histology, and histogenesis of combined hepatocellular-cholangiocarcinoma. *World J Hepatol* 2017;9:300-9.
 18. Wang Y, Yang Q, Li S, et al. Imaging features of combined hepatocellular and cholangiocarcinoma compared with those of hepatocellular carcinoma and intrahepatic cholangiocellular carcinoma in a Chinese population. *Clin Radiol* 2019;74:407.e1-e10.
 19. Fowler KJ, Sheybani A, Parker RA, 3rd, et al. Combined hepatocellular and cholangiocarcinoma (biphenotypic) tumors: imaging features and diagnostic accuracy of contrast-enhanced CT and MRI. *AJR Am J Roentgenol* 2013;201:332-9.
 20. Ye J, Xie X, Lin Y, et al. Imaging features of combined hepatocellular-cholangiocarcinoma on contrast-enhanced ultrasound: correlation with clinicopathological findings. *Clin Radiol* 2018;73:237-43.
 21. Shetty AS, Fowler KJ, Brunt EM, et al. Combined hepatocellular-cholangiocarcinoma: what the radiologist needs to know about biphenotypic liver carcinoma. *Abdom Imaging* 2014;39:310-22.
 22. Li F, Han J, Han F, et al. Combined Hepatocellular Cholangiocarcinoma (Biphenotypic) Tumors: Potential Role of Contrast-Enhanced Ultrasound in Diagnosis. *AJR Am J Roentgenol* 2017;209:767-74.
 23. Potretzke TA, Tan BR, Doyle MB, et al. Imaging Features of Biphenotypic Primary Liver Carcinoma (Hepatocholangiocarcinoma) and the Potential to Mimic Hepatocellular Carcinoma: LI-RADS Analysis of CT and MRI Features in 61 Cases. *AJR Am J Roentgenol* 2016;207:25-31.
 24. Wells ML, Venkatesh SK, Chandan VS, et al. Biphenotypic hepatic tumors: imaging findings and review of literature. *Abdom Imaging* 2015;40:2293-305.
 25. Sammon J, Fischer S, Menezes R, et al. MRI features of combined hepatocellular-cholangiocarcinoma versus mass forming intrahepatic cholangiocarcinoma. *Cancer Imaging* 2018;18:8-16.
 26. Xu X, Zhang HL, Liu QP, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol* 2019;70:1133-44.
 27. Huang Y, Liu Z, He L, et al. Radiomics signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* 2016;281:947-57.
 28. Ji GW, Zhang YD, Zhang H, et al. Biliary Tract Cancer at CT: A Radiomics-based Model to Predict Lymph Node Metastasis and Survival Outcomes. *Radiology* 2019;290:90-8.
 29. Aishima S, Nishihara Y, Iguchi T, et al. Lymphatic spread is related to VEGF-C expression and D2-40-positive myofibroblasts in intrahepatic cholangiocarcinoma. *Mod Pathol* 2008; 21:256-64.
 30. Tian MX, He WJ, Liu WR, et al. A Novel Risk prediction Model for Patients with Combined Hepatocellular-Cholangiocarcinoma. *J Cancer* 2018;9:1025-32.
 31. Tang D, Nagano H, Nakamura M, et al. Clinical and pathological features of Allen's type C classification of resected combined hepatocellular and cholangiocarcinoma: a comparative study with hepatocellular carcinoma and cholangiocellular carcinoma. *J Gastrointest Surg* 2006;10:987-98.
 32. Li R, Yang D, Tang CL, et al. Combined hepatocellular carcinoma and cholangiocarcinoma (biphenotypic) tumors: clinical characteristics, imaging features of contrast-

- enhanced ultrasound and computed tomography. *BMC Cancer* 2016;16:158-68.
33. Kim SH, Park YN, Lim JH, et al. Characteristics of combined hepatocellular-cholangiocarcinoma and comparison with intrahepatic cholangiocarcinoma. *Eur J Surg Oncol* 2014;40:976-81.
 34. Kickingreder P, Gotz M, Muschelli J, et al. Large-scale Radiomic Profiling of Recurrent Glioblastoma Identifies an Imaging Predictor for Stratifying Anti-Angiogenic Treatment Response. *Clin Cancer Res* 2016;22:5765-71.
 35. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5:4006-14.
 36. Banerjee S, Wang DS, Kim HJ, et al. A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. *Hepatology* 2015;62:792-800.
 37. Renzulli M, Brocchi S, Cucchetti A, et al. Can current preoperative imaging be used to detect microvascular invasion of hepatocellular carcinoma? *Radiology* 2016;279:432-42.
 38. Beig N, Khorrani M, Alilou M, et al. Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas. *Radiology* 2019;290:783-92.

Cite this article as: Zhang J, Huang Z, Cao L, Zhang Z, Wei Y, Zhang X, Song B. Differentiation combined hepatocellular and cholangiocarcinoma from intrahepatic cholangiocarcinoma based on radiomics machine learning. *Ann Transl Med* 2020;8(4):119. doi: 10.21037/atm.2020.01.126

Supplementary

Table S1 The details of 396 radiomics features

Class	Numbers of features	3 representative features
Intensity histogram features	42	Energy, entropy, MaxIntensity
Shape	9	Sphericity, surface area, compactness 1
GLCM	154	The energy of GLCM, entropy of GLCM, inertia of GLCM
GLSZM	11	Small area emphasis, large area emphasis, grey level non-uniformity
GLRLM	180	Short run emphasis, long run emphasis, run length non-uniformity

All radiomic features in this study were described in the Image biomarker standardization initiative (IBSI). GLCM, grey level co-occurrence matrix-based features; GLSZM, grey level size zone matrix-based features; GLRLM, grey level run length matrix based features.

Table S2 The results of the radiomics analysis

Regions	Features selected	Training cohort				Validation cohort				Cut-off values
		AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	
AP										
<i>ROI^{tumor}</i>	9	0.794	0.709	0.625	0.870	0.761	0.678	0.632	0.762	0.71
<i>ROI^{peritumoral}</i>	7	0.747	0.735	0.791	0.630	0.716	0.684	0.757	0.550	0.69
<i>ROI^{plus}</i>	9	0.742	0.773	0.872	0.587	0.751	0.772	0.946	0.450	0.57
AP with preprocessing										
<i>ROI^{tumor}</i>	8	0.800	0.742	0.744	0.739	0.789	0.754	0.784	0.700	0.59
<i>ROI^{peritumoral}</i>	6	0.768	0.727	0.698	0.783	0.705	0.632	0.622	0.650	0.65
<i>ROI^{plus}</i>	6	0.763	0.644	0.512	0.891	0.761	0.696	0.459	0.850	0.74
PVP										
<i>ROI^{tumor}</i>	7	0.737	0.659	0.591	0.787	0.733	0.576	0.500	0.714	0.68
<i>ROI^{peritumoral}</i>	6	0.722	0.712	0.767	0.609	0.698	0.655	0.730	0.524	0.60
<i>ROI^{plus}</i>	4	0.757	0.682	0.570	0.891	0.752	0.707	0.649	0.810	0.70
PVP with preprocessing										
<i>ROI^{tumor}</i>	8	0.750	0.664	0.591	0.804	0.731	0.576	0.526	0.667	0.61
<i>ROI^{peritumoral}</i>	5	0.716	0.636	0.523	0.848	0.723	0.667	0.622	0.750	0.67
<i>ROI^{plus}</i>	4	0.743	0.682	0.651	0.739	0.704	0.667	0.730	0.550	0.64

ROC analysis demonstrated that AUCs were not significant difference among 12 radiomics models ($P > 0.05$); lesser than cut-off value means combined hepatocellular and cholangiocarcinoma. AP, arterial phase; PVP, portal vein phase; ROI, region of interest; plus, tumor plus peritumoral; AUC, area under the curve; ACC, accuracy; SEN, sensitivity; SPE, specificity.

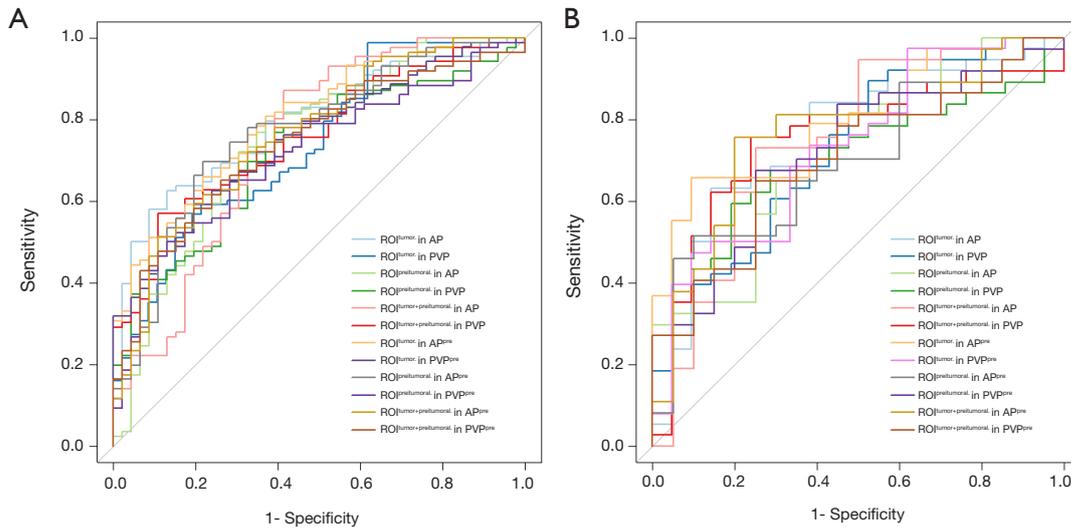


Figure S1 ROC analysis for differentiating between ICC and CHC patients in 12 R models. (A) Training cohort; (B) validation cohort. AP, arterial phase. PVP, portal vein phase; ROI, region of interest; Pre, imaging preprocessing; ICC, intrahepatic cholangiocarcinoma; CHC, combined hepatocellular and cholangiocarcinoma.

The formulas for clinical model, liver imaging model, and radiomics model

The formula for clinical model was described as:

$$Y = 3.81 - 1.413 \times \text{gender} (0, \text{male}; 1, \text{female}) - 2.772 \times \text{AFP} (0, <8; 1, \geq 8) - 0.817 \times \text{cirrhosis} (0, \text{absent}; 1, \text{present}) - 1.483 \times \text{hepatitis B} (0, \text{absent}; 1, \text{present}).$$

The formula for liver imaging model was described as:

$$Y = 1.052 - 0.695 \times \text{intratumor necrosis} (0, \text{absent}; 1, \text{present}) - 2.384 \times \text{intravascular tumor thrombus} (0, \text{absent};$$

1, present).

The formula for radiomics model was described as:

$$Y = 2.335 + 0.651 \times \text{skewness} - 0.169 \times \text{stdDeviation} + 0.844 \times \text{ClusterProminence_angle0_offset1} - 0.249 \times \text{ClusterShade_angle0_offset3} - 0.244 \times \text{GLCMEntropy_AllDirection_offset1_SD} + 7.273 \times \text{GLCMEntropy_AllDirection_offset3_SD} + 0.748 \times \text{HaralickCorrelation_AllDirection_offset1_SD} - 0.61 \times \text{Inertia_angle90_offset1}.$$