

Development and validation of magnetic resonance imaging-based nomograms for patients with perihilar cholangiocarcinoma after radical resection: Preoperative prediction of survival

Jian Zhao

Sichuan University West China Hospital

Wei Zhang

Sichuan University West China Hospital

Jun Zhang

Sichuan University West China Hospital

Wen-Jie Ma

Sichuan University West China Hospital

Fang Yuan

Sichuan University West China Hospital

Si-Yun Liu

GE healthcare

Fu-Yu Li

Sichuan University West China Hospital

Bin Song (✉ songlab_radiology@163.com)

Sichuan University West China Hospital

Research

Keywords: Perihilar cholangiocarcinoma, Overall survival, Curative intent resection, Magnetic resonance imaging, Radiomics, Nomogram

Posted Date: September 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-80701/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Survival status prediction for perihilar cholangiocarcinoma (pCCA) patients is essential for postoperative clinical decision making. This study aimed to develop and validate prediction models for overall survival (OS) evaluation in pCCA patients preoperatively.

Materials and Methods

A total of 184 patients who had curative resection for pCCA between January 2010 and December 2018 were enrolled. 110 patients were randomly selected for model development while other 74 patients for model testing. Imaging-derived radiomics signatures were developed. Independent preoperative clinical predictors were involved independently or in combination with radiomics signatures to construct different preoperative models through multivariable Cox proportional hazards method. The nomograms were constructed to predict OS, and the performance of which was evaluated by the discrimination ability, time-dependent receiver operating characteristic curve (ROC), calibration curve and decision curve.

Results

The clinical model (Model^{clinic}) was constructed based on three independent variables including preoperative CEA, cN stage and invasion of hepatic artery in images. The model with best performance (Model^{clinic&AP&PVP}) was build based on three independent variables, Signature^{AP} and Signature^{PVP}. In training and testing cohorts, the concordance indexes (C-indexes) of Model^{clinic} were 0.846 (95% CI, 0.735-0.957) and 0.755 (95% CI, 0.540-0.969), and Model^{clinic&AP&PVP} manifested favorable performance with C-indexes of 0.962 (95% CI, 0.905-1) and 0.814 (95% CI, 0.569-1), and both of them outperformed TNM staging system (C-indexes, 0.616, 95% CI, 0.522-0.711 and 0.599, 95% CI, 0.490-0.708). Good agreement was observed in the calibration curves, and favorable clinical utility was validated using the decision curve analysis both for Model^{clinic} and Model^{clinic&AP&PVP}.

Conclusions

Two preoperative nomograms were constructed to predict 1-, 3- and 5-years survival following surgery for individual pCCA patients. Such methods are easy to be performed which have clinical application potential for decision-making and patients stratification in randomized, controlled trials.

Background

Over the past four decades, the overall incidence of cholangiocarcinoma (CCA) has increased progressively worldwide [1]. CCA is a devastating malignancy, which derived from cells of the biliary tree and may occur anywhere in the biliary tree [2]. CCA is categorized according to anatomical location and divided into intrahepatic cholangiocarcinoma (iCCA), perihilar cholangiocarcinoma (pCCA), and distal

cholangiocarcinoma (dCCA) [3]. The most common type is pCCA, and pCCA is referred to as 'Klatskin' tumor sometimes [4]. pCCA arise in the perihilar region, the upper boundary of which is the second-order of bile ducts, the lower border is conjunction of the cystic duct and common hepatic duct [5]. Radical surgical resection is the only curative option for pCCA. Although accepted surgical resection, unfortunately, the prognosis is still poor along with high rate of local recurrence [6, 7]. Several recent studies indicated that adjuvant therapy could improve outcomes in resected pCCA [4, 5]. So, it is valuable to stratify patients who will have a short life expectancy and may benefit from oncologic treatment.

Oncologists and patients alike desire reliable prognostic tool tailored to the individual patient. The American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) staging system was mostly designated for predicting prognosis. The Amsterdam Medical Centre (AMC) Hilar Cholangiocarcinoma Group have proposed a nomogram which embrace post-operative tumor biological characteristics [8]. The Memorial Sloan-Kettering Cancer Centre (MSKCC) system, primarily used for assessment of potential resection, has been confirmed to be able to stratify patients' prognostic putatively [9]. But the accuracy of them were not satisfying [10]. There is an urgent need for an alternative tool to predict prognosis individually.

To achieve personalized medicine, personalized imaging is essential. Tumors are characterized by genetic and phenotype variation. Heterogeneity of malignancy is associated with cancer treatment failure and thus a poor prognosis [11]. Biomedical images contain information that reflects underlying pathophysiology. These relationships can be revealed via quantitative image analyses [12]. Radiomics could be used for extracting quantitative data which are undetectable at visual morphologic analysis by using dedicated algorithms [13]. The radiomics features have been reported to correlate with the prognosis of patients with malignancies [14–17]. And radiomics signature has also been involved to construct nomograms for individualized diagnosis [18, 19]. Nomograms can create a simple graphical representation of a statistical predictive model interpreting numerical probability of a clinical event, which is widely utilized in cancer area [20, 21]. The ability of nomogram to generate individualized prediction enables its use in clinic decision making. For many cancers, nomograms have been used for comparing to the traditional TNM staging systems favorably, and even have been proposed as an alternative or a new standard [22–25].

To our knowledge, radiomics-based nomogram for predicting prognosis in pCCA has not yet been established. Accordingly, the aim of this study is to develop and validate a decision support tool by incorporating preoperative clinical variables and MRI-based radiomics signature for pCCA patients in a high-volume center. In the ideal scenario, we hypothesized that the new developed tool could predict the prognosis of pCCA resected with curative intent accurately.

Materials And Methods

The current study was performed as a retrospective study. The ethics committee of West China Hospital, Sichuan University approved this retrospective analysis, and waived the requirement for informed consent

(2020 – 850). Our study was conducted in accordance with the Declaration of Helsinki, TRIPOD, and Guidelines of Luo et al [26, 27].

Patient demographic

All the patients who underwent radical surgical treatment for pCCA in West China Hospital, Sichuan University during January 2010 and December 2018 were identified. Patients with pCCA confirmed histologically were included in the study. Exclusion criterion: lack of preoperative imaging; only had CT images; examined at 3T MRI; movement artifacts and breathing artifacts; accepted preoperative transarterial chemoembolization; lost in follow-up since discharge; postoperative mortality within 90 days; incomplete pathologic reports or lack of preoperative serum test.

Standard patient demographic was collected. The admission notes, operation records, pathologic reports and radiologic findings were reviewed for each patient. Potential prognostic factors that might be associated with the prognosis of pCCA were collected: age; gender; maximum diameter of the tumor; preoperative blood glucose level; platelet count; white blood cell count; neutrophil count; lymphocyte count; monocyte count; albumin level; globulin level; ratio of albumin to globulin; total bilirubin (TBIL), direct bilirubin (DBIL), indirect bilirubin (IBIL), gamma-glutamyl transpeptidase (GGT), aspartate amino transferase (AST), alkaline phosphatase (ALP), alanine aminotransferase (ALT), carcinoembryonic antigen (CEA); carbohydrate antigen 19 – 9 (CA19-9); hepatitis B virus (HBV) infection; preoperative bile duct drainage; Bismuth type; resection margin status; histologic grade; T stage; operation details and survival status. The endpoint of this study was overall survival (OS), which was computed as the interval between the date of operation and the date of death or last follow-up. All laboratory indicators were examined within one week before surgery.

Treatment and follow-up

Surgical procedures were finally determined and conducted according to preoperative multidisciplinary team (MDT) discussion and intraoperative exploration. The operative technique and management were described in our previous study [28]. R1 was defined as microscopically positive. R0 was defined as no residual tumor. Routine histopathological workup was conducted for all resected pCCA by the Department of Pathology. The TNM stage of pCCA was mainly determined by operation and pathologic records.

Patients were followed-up until January 2020. Outpatient follow-up was executed every 2–3 months for the first year post operation, and every 3–6 months thereafter. At each visit, measurement of tumor markers and liver function, examination of CT and/or MRI were performed for assessment.

MRI Protocol and Conventional MRI Analysis

All enrolled patients underwent dynamic contrast-enhanced (DCE) MRI. All MR scans were performed at a 1.5 T MR scanner (Avanto, Siemens Healthcare, Erlangen, Germany) with an 8-channel body array coil. All patients were instructed to fasting for 6–8 h before the examination. Details regarding the MRI retrieval procedure and acquisition parameters are presented in Appendix E1 (online).

All data were transferred to a devoted workstation (Advantage Workstation 4.6, GE Healthcare) for analysis. Two radiologists with 16 (Wei Zhang, reader1) and 15 (Jun Zhang, reader2) years of experience in abdominal MRI diagnosis reviewed the imaging features of all patients independently. The radiologists were aware that the patients had pCCA, but they were blinded to the exact pathologic type and TNM stage. Differences in their findings were resolved by consensus. Conventional MRI findings were assessed focus on invasion of hepatic artery, invasion of portal vein, clinical lymph nodules status (cN stage) and growth patterns of the tumors. Invasion of hepatic artery and portal vein were considered present at following signs: (☒) overt tumor thrombus in the vessel. (☒) the vessel was deformed, stenosis or occluded. (☒) 50% or more of the circumference of vessel was surrounded by tumor. (☒) the border between the tumor and the relevant vessel was irregular. cN stage was considered according to the 8th edition of AJCC staging system. Stage cN0 represents free of metastatic regional LNs. Stage cN1 was defined as patients with 1–3 metastatic regional LNs. Stage cN2 was defined as >3 metastatic regional LNs [3]. Metastatic regional LNs was defined as: (☒) the lymph node has a short diameter of more than 1 cm. (☒) fusion of the lymph nodes. (☒) the occurrence of a centrally necrosis.

Univariate and multivariate COX regression analysis of preoperative variables

After univariate analysis, variables with $P < 0.10$ were inputted to multivariate analysis. Independent variables were test for the proportional hazards (PH) assumption and used for construction of model.

Image acquisition, Region-of-Interest segmentation

All images were stored in a Digital Imaging and Communications in Medicine (DICOM) format and anonymized. Prior to image analysis, the MR images at contrast enhanced artery phase (CE-AP) and portal vein phase (CE-PVP) were preprocessed by two steps as follows: (1) Intensity standardization (Z-score normalization) was conducted for all the MR images to reduce the variability across different image acquisitions and improve the radiomics feature reproducibility to some extent [29]. (2) Every standardized MR image was resampled to a uniform voxel dimension of $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ before region-of-interest (ROI) segmentation. Image registration for preoperative MR images at artery and portal vein phases was firstly performed with in-house software (Artificial Intelligence Kit, AK, *version 3.2.2*, GE Healthcare).

ROI segmentation was performed using ITK-SNAP software (*version 3.6.0*, open source software; <https://itk.org/>). ROI was delineated for the entire tumor on each slice of both axial CE-AP and CE-PVP images in each patient. To ensure intra- and inter-observer reproducibility, layer-by-layer images segmentation was performed along the tumor contour separately on CE-AP and CE-PVP MR images by two radiologists (twice by reader1 and once by reader2) for 30 patients selected randomly. The two radiologists segmented the images with double-blinded manner. The previously mentioned radiologist (Wei Zhang, reader1) then finished all tumor segmentation.

Radiomics feature extraction and selection, and model development

Radiomics features were extracted from CE-AP and CE-PVP images separately based on segmented tumor ROI by using in-house software (Artificial Intelligence Kit, AK, *version 3.2.2*, GE Healthcare). For respective CE-AP and CE-PVP images, 396 radiomics features were extracted, including 42 histogram features, 9 morphological features, 144 gray-level co-occurrence matrix (GLCM) features, 180 gray-level run length matrix (GLRLM) features, 11 gray-level size zone matrix (GLSZM) features and 10 Haralick features. The description of the radiomics features which are involved in the radiomics OS predicting models were described in Appendix E2.

Radiomics feature with missing values were firstly replaced with median values. Prior to feature selection, we used the intra- and interclass correlation coefficient (ICC) with the 95% confidence interval (CI) to assess the agreement for each extracted feature. After ICC calculation, features (ICC > 0.7, indicated good agreement) of CE-AP and CE-PVP were kept for further selection. Then the dataset was randomly divided by the ratio of 3:2 into training (110 patients) and testing (74 patients) cohorts, and the feature selection was performed in the training cohort.

The independent significant radiomics features for OS prediction were selected separately based on CE-AP and CE-PVP features. Following the Z-score normalization of radiomics feature in the training and testing cohorts, the redundant features in the training set were excluded by intra-correlation analysis at a cut value of 0.9, which means only one feature was retained when the correlation coefficient between two features larger than 0.9. Next, the univariate COX regression ($P < 0.05$) and the least absolute shrinkage and selection operator (LASSO) COX regression with 10-fold cross validation was used to select independent features for OS. Two radiomics signatures (Signature^{AP} and Signature^{PVP}) were built through a linear combination of selected features weighted by their corresponding coefficients generated from LASSO Cox regression models.

Finally, Signature^{AP}, Signature^{PVP} and independent clinical variables were used independently or in combination for construction of different models including: The clinical model (Model^{clinic}), the AP radiomics model (Model^{AP}), the PVP radiomics model (Model^{PVP}) and four combination model (Model^{clinic&AP}, Model^{clinic&PVP}, Model^{AP&PVP}, Model^{clinic&AP&PVP}).

Assessment of models

Then, we compared the performance of the models. The C-index was employed to assess the discrimination power of the models. The C-index was commonly used to evaluate the discriminative ability of prognostic models in survival analysis. The value of the C-index can range from 0.5 to 1.0 (0.5 indicates no discriminative ability and 1 indicates perfect performance). The best predictive model was determined by performance of comparison and selected for subsequent use. Meanwhile, comparison with the AJCC staging system was carried out using concordance probabilities.

Development of nomogram and validation of the prediction model

To provide a more visualized and individual predictive model, we drafted nomograms for the best predictable model and the clinical model. The discriminant accuracy of the model was further evaluated using the area under curves (AUC) of time-dependent receiver operating characteristic curve (ROC). Calibration curves were assessed by plotting the nomogram-predicted probabilities against the observed rates graphically via a bootstrap method with 1000-iteration resampling. Decision curve analysis (DCA) was employed to evaluate the clinical usefulness of models, through calculation of the net benefit for a range of threshold probabilities [30]. In addition, risk stratification using the constructed model was conducted. Kaplan-Meier survival curves were plotted, and the log-rank tests were used to compare OS between the low-risk and high-risk groups. The performance of the models was then validated in an independent testing set by using the formula generated from the training cohort.

Sample size estimation

As limited by the available sample size and no generally accepted methods to estimate the sample size for developing radiomics-based pCCA risk prediction models, we ensured that the number of positive outcome events in our study was near or larger than 10 times that of covariates (predictors) referring to the TRIPOD statement [26]. In the current study, there are 73 positive outcome events in the training set (110 patients in total) which meet the required condition. Meanwhile, there are 8, 2 and 3 independent predictors retained in the arterial-phase, portal-vein-phase radiomics and preoperative clinical OS predicting model, respectively.

Statistical analysis

All statistical analyses were performed using Statistical Package for Social Sciences software (SPSS, *version 25.0*, IBM, Armonk, NY, USA), R software (*Version: 3.5.3*, <https://www.r-project.org>). Continuous variables were presented as the mean and standard deviation (SD). The Mann-Whitney *U*-test or the *t*-test was performed to compare continuous variables. Cut-off values were determined by using the X-tile software (*Version: 3.6.1*, Yale University, New Haven, CT) [31]. Categorical variables were used Chi-squared (χ^2) test or Fisher's exact test to compare between groups. The Schoenfeld residual test was used to test the proportional hazards (PH) assumption for the selected clinical features [32]. All C-indexes and HRs were reported with 95% CIs. A two-tailed *P*-value < 0.05 was considered statistically significant. The following R packages were mainly used: "glmnet" for LASSO COX regression algorithm; "survival" for survival analyses including PH assumption of COX regression, COX regression analysis, Kaplan-Meier analysis and decision curve analysis; "tdROC" for time-dependent ROC analyses.

Results

Clinical characteristics and OS

The patients selecting workflow of the current study was shown in Fig. S1. A total of 184 patients underwent radical resection for pCCA were enrolled (115 male, 69 female; mean age 59.2 ± 10.3 years). The dataset was divided randomly into the training cohort (110 patients) and testing cohort (74 patients).

The detailed clinical characteristics and treatment parameters of patients with pCCA in the training and testing cohort are shown in Table 1. The estimated median OS time was 30.0 months (95% CI 19.5–40.5) for the training cohort, and 28.0 months (95% CI 21.5–34.5) for the testing cohort, with no significant difference between datasets ($P = 0.534$). The estimated median follow-up duration was 54.0 months (95% CI 40.7–67.3) for the training cohort, and 47.0 months (95% CI 23.6–70.4) for the testing cohort. At the last follow-up, 73 (66.4%) patients in the training cohort, and 49 (66.2%) in the testing cohort were died. The training and testing cohorts had similar baseline clinical characteristics (Table 1, $P > 0.05$).

Table 1
Clinicopathological features of pCCA patients in training and testing cohorts

| Variables | Training cohort (n/ mean ± SD) | Validation cohort (n/ mean ± SD) | P value |
|---|---|---|--------------------|
| Gender (male/female) | 63/47 | 52/22 | 0.074 |
| Age (years) | 59.100 ± 9.876 | 59.432 ± 10.923 | 0.779 |
| Tumor size (cm) | 2.450 ± 0.837 | 2.686 ± 1.112 | 0.466 |
| Preoperative laboratory indicators | | | |
| Blood glucose level (mmol/L) | 5.909 ± 2.143 | 5.706 ± 1.492 | 0.990 |
| Platelet count (10 ⁹ /L) | 218.630 ± 91.635 | 218.568 ± 75.306 | 0.604 |
| White blood cell count (10 ⁹ /L) | 6.634 ± 2.450 | 6.431 ± 2.176 | 0.625 |
| Neutrophile count (10 ⁹ /L) | 4.635 ± 2.246 | 4.342 ± 1.735 | 0.673 |
| Lymphocyte count (10 ⁹ /L) | 1.334 ± 0.573 | 1.373 ± 0.601 | 0.632 |
| Monocyte count (10 ⁹ /L) | 0.479 ± 0.202 | 0.487 ± 0.192 | 0.752 |
| Albumin level (g/L) | 37.022 ± 5.108 | 37.930 ± 4.681 | 0.154 |
| Globulin level (g/L) | 27.429 ± 4.866 | 26.797 ± 4.024 | 0.806 |
| Ratio of albumin to globulin | 1.388 ± 0.306 | 1.451 ± 0.296 | 0.223 |
| Ratio of neutrophile to lymphocyte | 4.332 ± 3.810 | 3.664 ± 2.350 | 0.565 |
| Ratio of monocyte to lymphocyte | 0.406 ± 0.237 | 0.403 ± 0.221 | 0.872 |
| Ratio of platelet to lymphocyte | 189.196 ± 105.969 | 182.441 ± 90.358 | 0.870 |
| TBIL(μmol/L) | 155.668 ± 115.786 | 158.803 ± 112.302 | 0.819 |
| DBIL μmol/L) | 133.723 ± 99.389 | 134.508 ± 97.247 | 0.959 |
| IBIL(μmol/L) | 21.945 ± 20.792 | 24.295 ± 20.515 | 0.176 |
| ALT(IU/L) | 184.572 ± 322.298 | 153.135 ± 176.244 | 0.649 |
| AST(IU/L) | 165.610 ± 286.653 | 145.392 ± 254.968 | 0.321 |
| ALP(IU/L) | 420.127 ± 348.749 | 434.189 ± 345.131 | 0.629 |

| Variables | Training cohort (n/ mean ± SD) | Validation cohort (n/ mean ± SD) | P value |
|--|---|---|--------------------|
| GGT(IU/L) | 547.500 ± 474.163 | 543.081 ± 472.628 | 0.924 |
| CEA (ng/mL) | 4.977 ± 7.932 | 5.626 ± 12.954 | 0.884 |
| CA199(U/mL) | 395.185 ± 370.609 | 389.879 ± 383.836 | 0.675 |
| Hepatitis B virus infection (present/absent) | 85/25 | 56/18 | 0.802 |
| Invasion of hepatic artery in images (present/absent) | 76/34 | 54/20 | 0.571 |
| Invasion of portal vein in images (present/absent) | 53/57 | 39/35 | 0.548 |
| Preoperative bile duct drainage (present/absent) | 28/82 | 16/58 | 0.550 |
| Type of resection | | | 0.965 |
| Left hemihepatectomy + caudate lobe resection | 48 | 34 | |
| Right hemihepatectomy + caudate lobe resection | 22 | 15 | |
| Left trisectionectomy + caudate lobe resection | 3 | 1 | |
| Right trisectionectomy + caudate lobe resection | 6 | 4 | |
| Mesohepatectomy + caudate lobe resection | 3 | 4 | |
| Bile duct resection with caudate lobe resection | 7 | 3 | |
| Bile duct resection only | 20 | 13 | |
| Liver transplantation | 1 | 0 | |
| Pathological differentiation | | | 0.270 |
| Well | 7 | 9 | |
| Moderate | 82 | 48 | |
| Poor | 21 | 17 | |
| Resection margins | | | 0.644 |
| R0 | 94 | 65 | |
| R1 | 16 | 9 | |
| Bismuth-Corlette classification | | | 0.222 |
| Type I | 19 | 6 | |

| Variables | Training cohort (n/ mean ± SD) | Validation cohort (n/ mean ± SD) | P value |
|---|-----------------------------------|-------------------------------------|---------|
| Type II | 38 | 26 | |
| Type IIIa | 15 | 8 | |
| Type IIIb | 30 | 23 | |
| Type IV | 8 | 11 | |
| Pattern of growth | | | 0.962 |
| Mass-forming type | 71 | 46 | |
| Periductal infiltrating type | 35 | 25 | |
| Intraductal growth type | 4 | 3 | |
| cN stage | | | 0.097 |
| N0 | 55 | 41 | |
| N1 | 54 | 29 | |
| N2 | 1 | 4 | |
| pTNM stage | | | 0.466 |
| I | 2 | 3 | |
| II | 44 | 23 | |
| IIIA | 17 | 18 | |
| IIIB | 11 | 9 | |
| IIIC | 28 | 14 | |
| IIIA | 3 | 4 | |
| IIIB | 5 | 3 | |
| Adjuvant chemotherapy (with/without) | 7/103 | 6/68 | 0.651 |
| Note: TBIL, Total bilirubin; DBIL, direct bilirubin; IBIL, indirect bilirubin; ALT, alanine aminotransferase; AST, aspartate amino transferase; ALP, alkaline phosphatase; GGT, gamma-glutamyl transpeptidase; CEA, carcinoembryonic antigen; CA19-9, carbohydrate antigen 19 - 9; SD, standard deviation | | | |

Conventional MRI Analysis

The tumor size was measured as the maximum diameter of the tumor. Maximum diameter of the tumor was 2.450 cm (SD, 0.837 cm) in training cohort and 2.686 cm (SD, 1.112 cm) in testing cohort. In training cohort, 55 patients were categorized as cN0, 54 patients as cN1 and one patient as cN2. In testing cohort,

cN0, cN1 and cN2 groups included 41, 29 and 4 patients. Invasion of hepatic artery were observed in 76 individuals and absent in 34 patients in training cohort. In testing cohort, 54 patients had invasion of hepatic artery and twenty without. Invasion of portal vein appeared in 53 patients and did not appear in 57 patients in training cohort, which of the testing cohort were 39 and 35 patients. Growth patterns of the tumors were classified as mass-forming (MF) type, periductal infiltrating (PI) type, and intraductal growth (IG) type via examination of preoperative MR-imaging appearance. Consequently, MF type, PI type and IG type involved 71, 35 and 4 patients in training cohort, and 46, 25 and 3 patients in testing cohort.

Univariate and multivariate COX regression analysis of preoperative clinical features

Cut-off values of preoperative variables were determined by X-tile software and were illustrated in Table S1. Univariate analysis exploring clinical factors found that preoperative white blood cell count, neutrophile count, lymphocyte count, CEA, CA19-9, cN stage, invasion of hepatic artery and portal vein in images were potentially associated with OS. Multivariate COX regression analysis identified CEA, cN stage and invasion of hepatic artery in images as independent variables (Table S1). The Schoenfeld residual test showed that the selected variables satisfied the proportional hazards assumption, and the constructed model as well (Fig. S2, $P = 0.766$).

Radiomic signature building

Figure 1 shows the radiomics workflow. After intra- and inter-observer ICC analyses, 340 and 252 features were kept respectively from CE-AP and CE-PVP feature sets. From these features, 117 features from CE-AP images and 85 features from CE-PVP images were selected after excluding redundancy by correlation analysis (cut-value: 0.9). After univariate COX regression, 9 features and 2 features from CE-AP and CE-PVP images were left. Applying Lasso-Cox regression analyses with 10-fold cross-validation, 8 features from CE-AP images (Fig. S3a-b), 2 features from CE-PVP images (Fig. S3c-d) were selected to build radiomic signatures (Signature^{AP}; Signature^{PVP}). The regression coefficients of constructing Signature^{AP} and Signature^{PVP} were shown in Table S2.

Models building, comparison and validation

Signature^{AP}, Signature^{PVP} and three independent clinical variables (preoperative CEA, cN stage and invasion of hepatic artery in images) were used to build Cox regression models with various combinations. For the training and testing cohorts, the C-indexes of these models for survival prediction were illustrated in Table 2.

Table 2
Performance of models for predicting overall survival

| Models | Training cohort ($n = 110$) | | | Testing cohort ($n = 74$) | | |
|---|-------------------------------|-------------|---------|-----------------------------|-------------|---------|
| | C-index | 95%CI | AIC | C-index | 95%CI | AIC |
| Clinic | 0.846 | 0.735–0.957 | 555.758 | 0.755 | 0.540–0.969 | 334.429 |
| AP | 0.871 | 0.771–0.971 | 559.568 | 0.774 | 0.637–0.910 | 340.335 |
| PVP | 0.709 | 0.504–0.914 | 575.983 | 0.678 | 0.485–0.871 | 340.663 |
| Clinic&AP | 0.895 | 0.814–0.976 | 542.661 | 0.744 | 0.558–0.931 | 333.188 |
| Clinic&PVP | 0.967 | 0.922-1 | 541.641 | 0.648 | 0.401–0.894 | 339.728 |
| AP&PVP | 0.877 | 0.774–0.979 | 559.892 | 0.756 | 0.615–0.897 | 334.846 |
| Clinic&AP&PVP | 0.962 | 0.905-1 | 544.501 | 0.814 | 0.569-1 | 336.677 |
| TNM | 0.616 | 0.522–0.711 | 580.421 | 0.599 | 0.490–0.708 | 340.305 |
| Note: models were build based on Signature ^{AP} , Signature ^{PVP} and three independent clinical variables (preoperative CEA, cN stage and invasion of hepatic artery in images), with various combinations. C-index, Harrell's concordance indices; CI, confidence interval; AIC, Akaike's information criterion. | | | | | | |

Model^{clinic} was built based on three independent clinical variables without the addition of radiomics signatures. Model^{clinic} had good performance and the C-indexes were 0.846 (95% CI, 0.735–0.957) and 0.755 (95% CI, 0.540–0.969) in training and testing cohorts. When adding radiomics signature into the models, Model^{clinic&AP&PVP} manifested greatest performance, with C-indexes of 0.962 (95% CI, 0.905-1) and 0.814 (95% CI, 0.569-1) in training and testing cohorts. Then, nomograms was constructed for Model^{clinic} (Fig. 2a) and Model^{clinic&AP&PVP} (Fig. 2b). Furthermore, Time dependent-ROC analysis was applied to validate Model^{clinic} and Model^{clinic&AP&PVP} in predicting OS, and gained satisfactory results (Fig. 3a-d).

The nomograms showed a significant improvement in predicting survival over the TNM staging system. In the training cohort, the Model^{clinic} showed a better prognostic capability than TNM staging system (C-index: 0.846 vs. 0.616). In the testing cohort, compared to TNM staging system, a higher C-index value was observed (C-index: 0.755 vs. 0.599). In the training cohort, Model^{clinic&AP&PVP} showed a better prognostic capability than TNM staging system (C-index: 0.962 vs. 0.616). In the testing cohort, a better prognostic ability was also observed in the novel nomogram (C-index: 0.814 vs. 0.599). The time-dependent ROC curves manifested that the Model^{clinic} (Fig. 3a and Fig. 3b), Model^{clinic&AP&PVP} (Fig. 3c and Fig. 3d) outperformed TNM staging system (Fig. 3e and Fig. 3f) in the training cohort and testing cohort in predicting both short- and long-term survival.

The calibration curves were plotted for Model^{clinic} (Fig. 4a-b) and Model^{clinic&AP&PVP} (Fig. 4c-d). Good agreement was observed in the calibration curves between the nomograms estimated 1-, 3- and 5-years OS rate and the observed OS rate in both Model^{clinic} and Model^{clinic&AP&PVP}. But less satisfied results were observed in TNM staging system (Fig. 4e-f).

Subgroup survival analysis stratified by the nomograms

By Model^{clinic}, we calculated the risk score for each patient in training and testing cohorts. The formula was as follow: $\text{risk score}^{\text{clinic}} = 0.7643512 \times \text{CEA} + 0.6273337 \times \text{cN} + 1.4823995 \times \text{invasion of hepatic artery in images}$. The optimal cutoff value of 0.627 was calculated by R function “surv_cutpoint” which is an outcome-oriented method. Based on this cutoff value, the patients were appointed to low-risk and high-risk groups. In training cohort, the risk score of the high risk subgroup was 1.898 (range, 1.788 to 2.008), and that of low risk group was 0.255 (range, 0.142 to 0.368). The risk scores of the high risk subgroup and low risk subgroup were 1.915 (range, 1.777 to 2.053) and 0.148 (range, 0.007 to 0.289) in testing cohort. Patients in the high-risk subgroup of the training cohort showed a poorer survival rate than low-risk subgroup in the subgroup survival analysis (1-year SR: 80.5% vs. 93.8%; 3-year SR: 31.6% vs. 78.8%; 5-year SR: 6.5% vs. 57.3%; $P < 0.001$, Fig. 5a). Similar result was observed in the testing cohort between high-risk subgroup and low-risk subgroup (1-year SR: 78.9% vs. 100.0%; 3-year SR: 27.8% vs. 62.6%; 5-year SR: 10.0% vs. 22.7%; $P = 0.007$, Fig. 5b).

Similarly, risk score for each patient was acquired by Model^{clinic&AP&PVP}. The formula was as follow: $\text{risk score}^{\text{clinic&AP&PVP}} = 0.7137881 \times \text{Signature}^{\text{AP}} + 0.4683287 \times \text{Signature}^{\text{PVP}} + 0.9269602 \times \text{CEA} + 0.4928044 \times \text{cN} + 1.1506710 \times \text{invasion of hepatic artery in images}$. The optimal cutoff value was 0.198 which stratified the patients into low-risk and high-risk groups. In training cohort, the risk score of the high risk subgroup was 1.566 (range, 1.422 to 1.711), and that of low risk group was -0.423 (range, -0.719 to -0.128). The risk scores of the high risk subgroup and low risk subgroup were 1.444 (range, 1.288 to 1.600) and -0.446 (range, -0.889 to -0.004) in testing cohort. Patients in the high-risk subgroup of the training cohort showed a poorer survival rate than low-risk subgroup in the subgroup survival analysis (1-year SR: 80.0% vs. 100.0%; 3-year SR: 31.6% vs. 94.4%; 5-year SR: 7.4% vs. 84.8%; $P < 0.001$, Fig. 5c). Similar result was observed in the testing cohort between high-risk subgroup and low-risk subgroup (1-year SR: 79.0% vs. 91.6%; 3-year SR: 26.7% vs. 91.6%; 5-year SR: 13.0% vs. 36.7%; $P = 0.025$, Fig. 5d).

Clinical Use

The decision curve analysis for Model^{clinic}, Model^{clinic&AP&PVP} and TNM staging system were presented in Fig. S4a-d. According to DCA, when the threshold probability for a patient was within the range of 20–80%, the nomogram added more net benefit than the “treat all” or “treat none” strategies in training cohort in predicting both 3-year survival (Fig. S4a) and 5-year survival (Fig. S4b) in Model^{clinic}. Model^{clinic&AP&PVP} had similar performance (Fig. S4a and Fig. S4b for 3-year and 5-year survival) in training cohort. In testing cohort, good result was observed in predicting 3-year survival (Fig. S4c), but fair result in predicting 5-year survival (Fig. S4d) in Model^{clinic}. Model^{clinic&AP&PVP} had similar performance (Fig. S4c

and Fig. S4d for 3-year and 5-year survival). Both Model^{clinic} and Model^{clinic&AP&PVP} had more favorable clinical utility than TNM staging system.

Discussion

In recent years, gratifying achievements have been made in a few study of cholangiocarcinoma. Radiomics nomogram based on radiomics features and clinical characteristics was shown to predict early recurrence of iCCA after resection preoperatively [33]. Radiomics signatures based on ultrasound (US) medicine images have been proven to predict the biological behaviors of iCCA and had moderate efficiency [34]. In a study by Ji et al. [35], a radiomics model was established for predicting LNs metastasis of iCCA. Another study also confirmed the applicability of such method [36]. Radiomics model was also suggested to predict differentiation degree and lymph node metastases of extrahepatic cholangiocarcinoma [37]. To our knowledge, no study was performed to investigate individual prediction of survival of pCCA using MRI features. In this study, several MRI based prognostic models were derived and validated to predict OS after curative intent resection of pCCA. The models were based on Signature^{AP}, Signature^{PVP} and three independent clinical variables (CEA, cN stage and invasion of hepatic artery in images). Model^{clinic} and Model^{clinic&AP&PVP} which were presented as nomogram showed good discrimination. Calibration curves were good in the training and testing cohort. Both of them outperformed AJCC staging system, and can be used to assist clinical decision-making, inform patients' prognosis and stratify patients in randomized, controlled trials.

The performance of the AJCC TNM staging system to predict long-term outcome was poor (< 0.7) both in current training cohort (C-index, 0.616) and testing cohort (C-index, 0.599). Ruzzenente et al. [38] demonstrated that the 8th edition was superior to 7th edition, with C-index of 0.624. A study from South Korea confirmed this result with C-index of 0.621 [39]. The C-index of 8th edition was 0.67 in another study from Netherlands in patients who underwent a curative-intent resection [40]. Our results were similar to those literature, and Model^{clinic} and Model^{clinic&AP&PVP} had an improved performance in predicting OS compared with TNM staging system.

The current clinical model incorporated preoperative CEA, cN stage and invasion of hepatic artery in images had a C-index of 0.846 in training cohort, and 0.755 in testing cohort. When adding radiomics signature to the model, discrimination was further improved (C-indexes, 0.962 and 0.814 in training and testing cohorts). In a report by van der Gaag and colleagues [8], they developed a nomogram including 175 patients with extrahepatic cholangiocarcinoma in Academic Medical Center in Amsterdam (AMC), based on resection margins status, lymph node status and tumor differentiation. That nomogram was demonstrated to better than TNM staging system (C-index, 0.66 vs. 0.63). In a study by Bird et al. [41], they confirmed the AMC nomogram (C-index, 0.62) had an improved prognostic capability compared to the AJCC TNM staging system (C-index, 0.59). Their new developed multivariate models demonstrated significantly improved prognostic capability (C-index, 0.71). Another study developed a nomogram which demonstrated to better than TNM system, with a C-index of 0.68 in primary cohort and 0.65 in validation cohort [42]. A study incorporated variables including age, surgery, SEER historic stage and lymph node

metastases to establish a nomogram, and the C-index for OS prediction was 0.651 [43]. Whereas validation cohort was absent in their study. Groot Koerkamp et al. [44] proposed a nomogram based on data from the Memorial Sloan Kettering Cancer Center (MSKCC) for patients with resected pCCA. The concordance index was 0.73 in primary cohort and 0.72 in the validation cohort. Buettner et al. [10] demonstrated that the proposed nomogram by MSKCC for pCCA performed poorly, with C-index of 0.587. In that study, a revised nomogram was constructed, but the performance still less satisfied (C-index, 0.682). Most of these studies only highlight the clinicopathological characteristics postoperatively. Our nomograms were driven from preoperative variables and had a higher C-index. In clinical practice, when radiomics feature could not gain conveniently, this clinical model could act as an effective substitute.

The calibration curves of Model^{clinic} and Model^{clinic&AP&PVP} indicated that the constructed nomograms were reliable and robust predictive models. Few patients could survival five years after surgery, thus the 3-year survival rate was the most important factor affecting clinical decision-making. So, although the decision curve analysis manifested fair result in predicting 5-year survival in testing cohort, the clinical utility will not affected. Therefore, the proposed MRI-based models can be applied directly in clinical practice. The decision curve analyses also confirmed that our nomograms had better clinical usefulness compared to 8th edition of AJCC TNM staging system.

The current study has several limitations. Firstly, our study is limited by its retrospective nature, selection bias may exist. Secondly, this study was performed in a single institute and single equipment, this may impede application of propagation, further multi-national and multi-institutional studies are required to confirm this result. Thirdly, regarding ROI segmentation, development of robust and reliable methods for auto-segmentation may be necessary in future [45]. Fourthly, this study only included CE-AP and CE-PVP sequences, whereas did not embrace diffusion weighed imaging (DWI), non-enhanced T1WI and T2WI. The reason for this selection is that the tumor is generally small and not conducive to delineation in non-enhanced images. Furthermore, the cohort of this study were conducted in patients who had curative-intent resection, thus its use is limited to the operative situation. At last, the knowledge regarding different molecular markers (gene- and protein expression) in bile duct tumors is rapidly developing. Introducing those markers into the model may further increase its accuracy.

Conclusion

We constructed two nomograms which are potential for easy day-to-day clinical use for individual patient. The nomograms provide tools to preoperatively predict 1-, 3- and 5-years survival for pCCA following surgery. It's valuable for assisting clinical decision-making and stratifying patients in randomized, controlled trials.

Abbreviations

CCA: Cholangiocarcinoma; CEA: Carcinoembryonic antigen; CA19-9: Carbohydrate antigen 19 – 9; IQR: Inter Quartile Range; LNs: Lymph nodes; OS: Overall survival; pCCA: Perihilar cholangiocarcinoma; AJCC: The

American Joint Committee on Cancer; TNM:Tumor-nodes-metastasis; HR:Hazard ratio; 95% CI:95% confidence interval; SD:Standard deviation

Declarations

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author (E-mail: songlab_radiology@163.com) on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Consent to publish

Not applicable.

Ethics approval and consent to participate

Ethical approval was obtained from respective institutional review boards (IRB) of West China Hospital, Sichuan University, and waived the requirement for informed consent (2020 – 850).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

J Z1 contributed to interpretation of data and manuscript drafting; W Z reviewed the literature and contributed to manuscript drafting; J Z2 contributed to collecting the admission notes, operation records, pathologic reports, and radiologic findings and revising manuscript; F Y and SY L contributed to statistics analysis and revising manuscript; WJ M completed the follow-up of the patients and revising manuscript; FY L and B S contributions to the conception or design of the work and responsible for the revise of the manuscript for important intellectual content. All authors have read and approved the manuscript.

Acknowledgements

We thank all the surgeons from Department of Biliary Surgery, West China Hospital, Sichuan University who performed the surgery in this study.

References

1. Saha SK, Zhu AX, Fuchs CS, Brooks GA. Forty-Year Trends in Cholangiocarcinoma Incidence in the U.S.: Intrahepatic Disease on the Rise. *Oncologist*. 2016;21:594–9.
2. Forner A, Vidili G, Rengo M, Bujanda L, Ponz-Sarvise M, Lamarca A. Clinical presentation, diagnosis and staging of cholangiocarcinoma. *Liver Int*. 2019;39(Suppl 1):98–107.
3. Chun YS, Pawlik TM, Vauthey JN. 8th Edition of the AJCC Cancer Staging Manual: Pancreas and Hepatobiliary Cancers. *Ann Surg Oncol*. 2018;25:845–7.
4. Esnaola NF, Meyer JE, Karachristos A, Maranki JL, Camp ER, Denlinger CS. Evaluation and management of intrahepatic and extrahepatic cholangiocarcinoma. *Cancer*. 2016;122:1349–69.
5. Rizvi S, Khan SA, Hallemeier CL, Kelley RK, Gores GJ. Cholangiocarcinoma - evolving concepts and therapeutic strategies. *Nat Rev Clin Oncol*. 2018;15:95–111.
6. Razumilava N, Gores GJ. Cholangiocarcinoma *The Lancet*. 2014;383:2168–79.
7. Ahn DH, Bekaii-Saab T. Biliary cancer. intrahepatic cholangiocarcinoma vs. extrahepatic cholangiocarcinoma vs. gallbladder cancers: classification and therapeutic implications. *J Gastrointest Oncol*. 2017;8:293–301.
8. van der Gaag NA, Kloek JJ, de Bakker JK, Musters B, Geskus RB, Busch OR, et al. Survival analysis and prognostic nomogram for patients undergoing resection of extrahepatic cholangiocarcinoma. *Ann Oncol*. 2012;23:2642–9.
9. Zaydfudim VM, Clark CJ, Kendrick ML, Que FG, Reid-Lombardo KM, Donohue JH, et al. Correlation of staging systems to survival in patients with resected hilar cholangiocarcinoma. *Am J Surg*. 2013;206:159–65.
10. Buettner S, van Vugt JL, Gani F, Groot Koerkamp B, Margonis GA, Ethun CG, et al. A Comparison of Prognostic Schemes for Perihilar Cholangiocarcinoma. *J Gastrointest Surg*. 2016;20:1716–24.
11. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 2013;501:338–45.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278:563–77.
13. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749–62.
14. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V. Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology*. 2013;266:177–84.
15. Cassinotto C, Chong J, Zogopoulos G, Reinhold C, Chiche L, Lafourcade JP, et al. Resectable pancreatic adenocarcinoma: Role of CT quantitative imaging biomarkers for predicting pathology

- and patient outcomes. *Eur J Radiol.* 2017;90:152–8.
16. Zhang B, He X, Ouyang F, Gu D, Dong Y, Zhang L, et al. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. *Cancer Lett.* 2017;403:21–7.
 17. Akai H, Yasaka K, Kunimatsu A, Nojima M, Kokudo T, Kokudo N, et al. Predicting prognosis of resected hepatocellular carcinoma by radiomics analysis with random survival forest. *Diagn Interv Imaging.* 2018;99:643–51.
 18. Zhang Z, Jiang H, Chen J, Wei Y, Cao L, Ye Z, et al. Hepatocellular carcinoma: radiomics nomogram on gadoxetic acid-enhanced MR imaging for early postoperative recurrence prediction. *Cancer Imaging.* 2019;19:22.
 19. Ji GW, Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, et al. Biliary Tract Cancer at CT: A Radiomics-based Model to Predict Lymph Node Metastasis and Survival Outcomes. *Radiology.* 2019;290:90–8.
 20. Shariat SF, Karakiewicz PI, Suardi N, Kattan MW. Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature. *Clin Cancer Res.* 2008;14:4400–7.
 21. Vickers AJ, Kattan MW, Daniel S. Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials.* 2007;8:14.
 22. Song KY, Park YG, Jeon HM, Park CH. A nomogram for predicting individual survival of patients with gastric cancer who underwent radical surgery with extended lymph node dissection. *Gastric Cancer.* 2014;17:287–93.
 23. Chen S, Lai Y, He Z, Li J, He X, Shen R, et al. Establishment and validation of a predictive nomogram model for non-small cell lung cancer patients with chronic hepatitis B viral infection. *J Transl Med.* 2018;16:116.
 24. Wen J, Yang Y, Liu P, Ye F, Tang H, Huang X, et al. Development and validation of a nomogram for predicting survival on the base of modified lymph node ratio in breast cancer patients. *Breast.* 2017;33:14–22.
 25. Fang C, Wang W, Feng X, Sun J, Zhang Y, Zeng Y, et al. Nomogram individually predicts the overall survival of patients with gastroenteropancreatic neuroendocrine neoplasms. *Br J Cancer.* 2017;117:1544–50.
 26. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162:W1–73.
 27. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *J Med Internet Res.* 2016;18:e323.
 28. Zhao J, Zhang W, Zhang J, Zhang Y, Ma WJ, Liu SY, et al. Survival analysis of patients with stage T2a and T2b perihilar cholangiocarcinoma treated with radical resection. *BMC Cancer.* 2020;20:849.
 29. Schwier M, van Griethuysen J, Vangel MG, Pieper S, Peled S, Tempany C, et al. Repeatability of Multiparametric Prostate MRI Radiomics Features. *Sci Rep.* 2019;9:9441.

30. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565–74.
31. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. 2004;10:7252–9.
32. Quantin C, Moreau T, Asselain B, Maccario J, Lellouch J. A regression survival model for testing the proportional hazards hypothesis. *Biometrics*. 1996;52:874–85.
33. Liang W, Xu L, Yang P, Zhang L, Wan D, Huang Q, et al. Novel Nomogram for Preoperative Prediction of Early Recurrence in Intrahepatic Cholangiocarcinoma. *Front Oncol*. 2018;8:360.
34. Peng YT, Zhou CY, Lin P, Wen DY, Wang XD, Zhong XZ, et al. Preoperative Ultrasound Radiomics Signatures for Noninvasive Evaluation of Biological Characteristics of Intrahepatic Cholangiocarcinoma. *Acad Radiol*. 2020;27:785–97.
35. Ji GW, Zhu FP, Zhang YD, Liu XS, Wu FY, Wang K, et al. A radiomics approach to predict lymph node metastasis and clinical outcome of intrahepatic cholangiocarcinoma. *Eur Radiol*. 2019;29:3725–35.
36. Xu L, Yang P, Liang W, Liu W, Wang W, Luo C, et al. A radiomics approach based on support vector machine using MR images for preoperative lymph node status evaluation in intrahepatic cholangiocarcinoma. *Theranostics*. 2019;9:5374–85.
37. Yang C, Huang M, Li S, Chen J, Yang Y, Qin N, et al. Radiomics model of magnetic resonance imaging for predicting pathological grading and lymph node metastases of extrahepatic cholangiocarcinoma. *Cancer Lett*. 2020;470:1–7.
38. Ruzzenente A, Bagante F, Ardito F, Campagnaro T, Scoleri I, Conci S, et al. Comparison of the 7th and 8th editions of the American Joint Committee on Cancer Staging Systems for perihilar cholangiocarcinoma. *Surgery*. 2018;164:244–50.
39. Lee JW, Lee JH, Park Y, Lee W, Kwon J, Song KB, et al. Prognostic Predictability of American Joint Committee on Cancer 8th Staging System for Perihilar Cholangiocarcinoma: Limited Improvement Compared with the 7th Staging System. *Cancer Res Treat*. 2020;52:886–95.
40. Gaspersz MP, Buettner S, van Vugt JLA, de Jonge J, Polak WG, Doukas M, et al. Evaluation of the New American Joint Committee on Cancer Staging Manual 8th Edition for Perihilar Cholangiocarcinoma. *J Gastrointest Surg*. 2020;24:1612–8.
41. Bird NTE, Manu N, Quinn L, Needham A, Jones R, Fenwick S, et al. Evaluation of the utility of prognostic models for patients with resected hilar cholangiocarcinoma. *HPB (Oxford)*. 2019;21:1376–84.
42. Chen P, Li B, Zhu Y, Chen W, Liu X, Li M, et al. Establishment and validation of a prognostic nomogram for patients with resectable perihilar cholangiocarcinoma. *Oncotarget*. 2016;7:37319–30.
43. Qi F, Zhou B, Xia J. Nomograms predict survival outcome of Klatskin tumors patients. *PeerJ*. 2020;8:e8570.
44. Groot Koerkamp B, Wiggers JK, Allen PJ, Busch OR, D'Angelica MI, DeMatteo RP, et al. American Joint Committee on Cancer staging for resected perihilar cholangiocarcinoma: a comparison of the 6th and 7th editions. *HPB (Oxford)*. 2014; 16:1074–1082.

45. Lin YC, Lin CH, Lu HY, Chiang HJ, Wang HK, Huang YT, et al. Deep learning for fully automated tumor segmentation and extraction of magnetic resonance radiomics features in cervical cancer. *Eur Radiol.* 2020;30:1297–305.

Figures

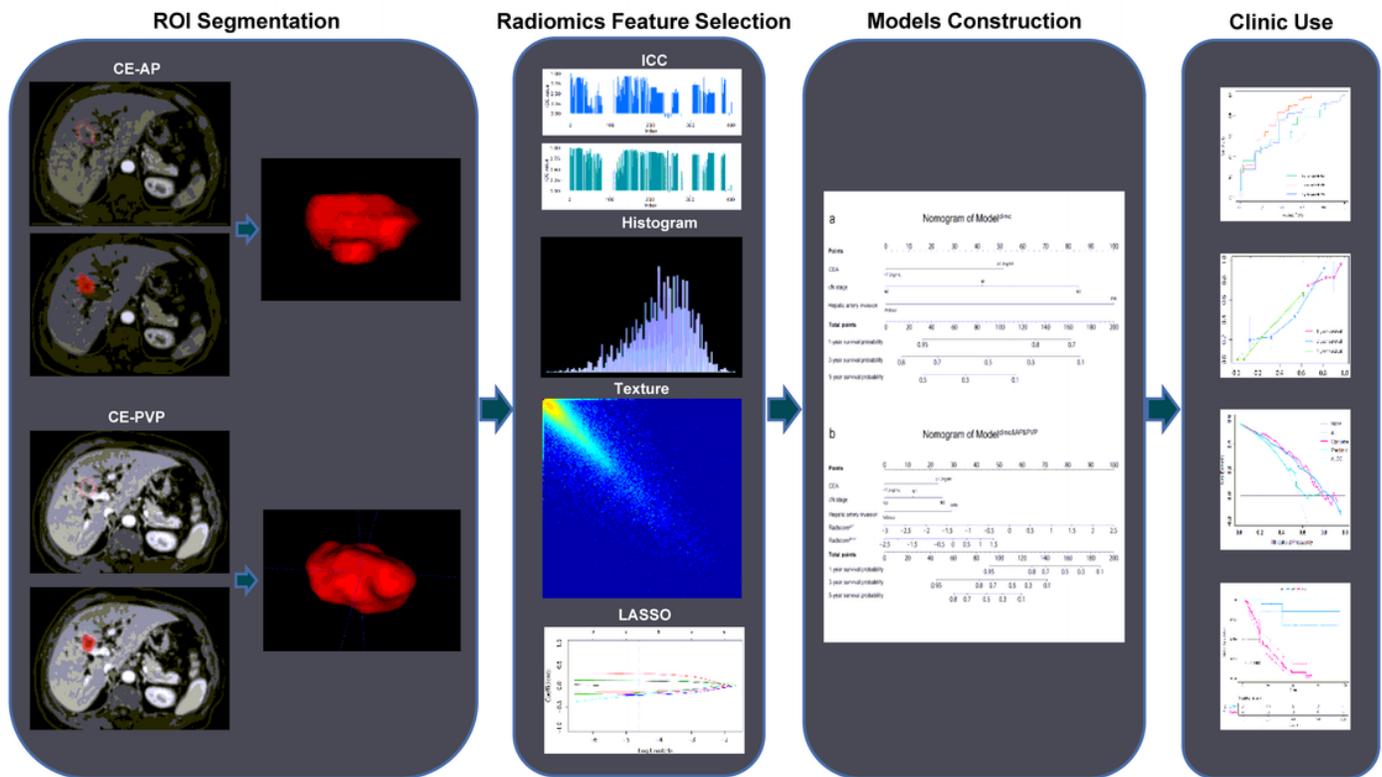


Figure 1

Workflow of necessary steps in current study. Tumors were segmented manually on axial hepatic artery and portal venous phase MRI section. Radiomic features were extracted from within the defined tumor contours on MRI images to quantify tumor intensity, shape, and texture. For feature selection, two successive steps were applied to the extracted features. At first, with inter- and intra-observer reliability assessment, and then univariate and multivariate LASSO-Cox regression. To provide a more understandable outcome measure, nomograms were built for individualized evaluation. The performance of the model was assessed and followed by decision curve analysis and survival prediction.

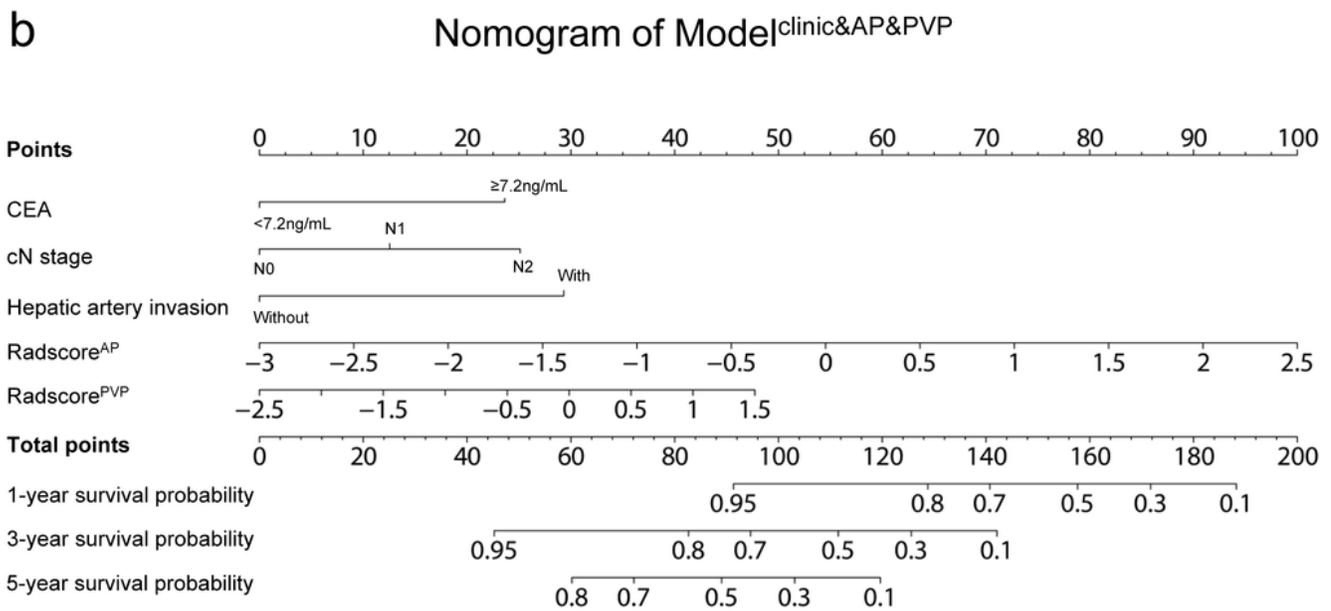
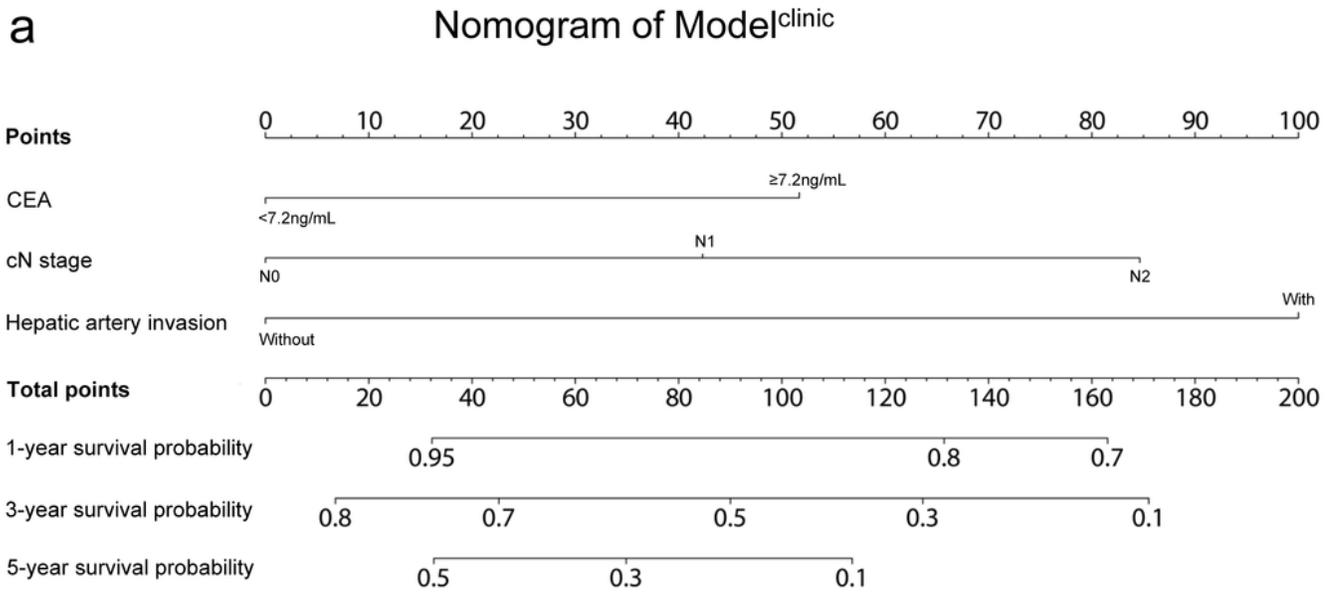


Figure 2

Nomograms to predict the 1-, 3-, 5-survival rate in patients with pCCA after radical resection. (a) Nomogram of Model^{clinic}; (b) Nomogram of Model^{clinic&AP&PVP}. To use the nomogram, please follow the steps below: (1) finding the individual patient's value located on each variable axis. (2) drawing a line upward to determine the risk score for each variant. (3) The sum of these scores and located on the total

points axis. (⊠) drawing a line downward to the survival axes to determine the probability of 1-, 3- and 5-year survival.

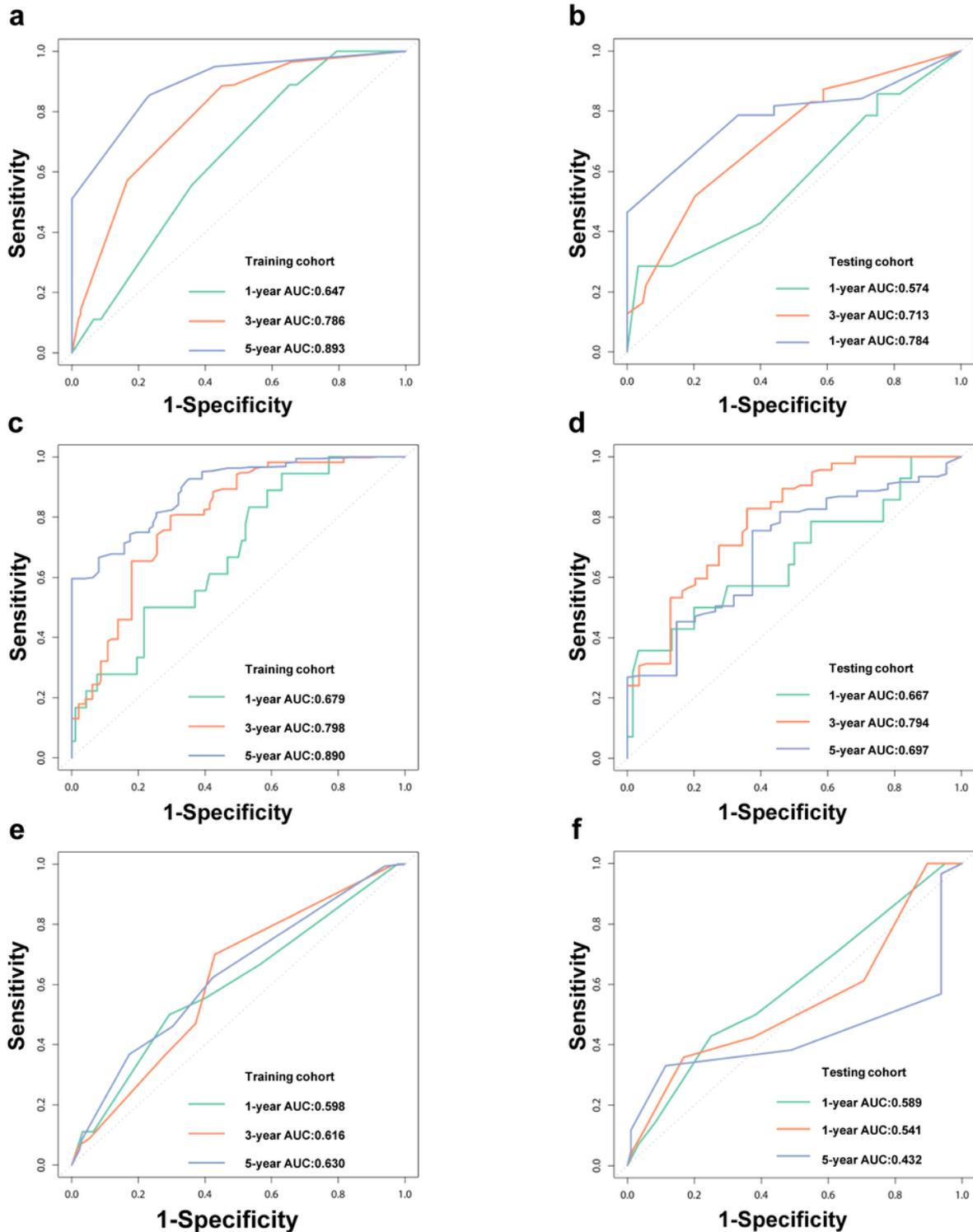


Figure 3

Time-dependent ROC curves. The performance of Modelclinic manifested good in training cohort (a) and testing cohort (b). The performance of Modelclinic&AP&PVP manifested satisfying in training cohort (c)

and testing cohort (d). AJCC TNM staging system had an unsatisfied performance in training cohort (e) and testing cohort (f).

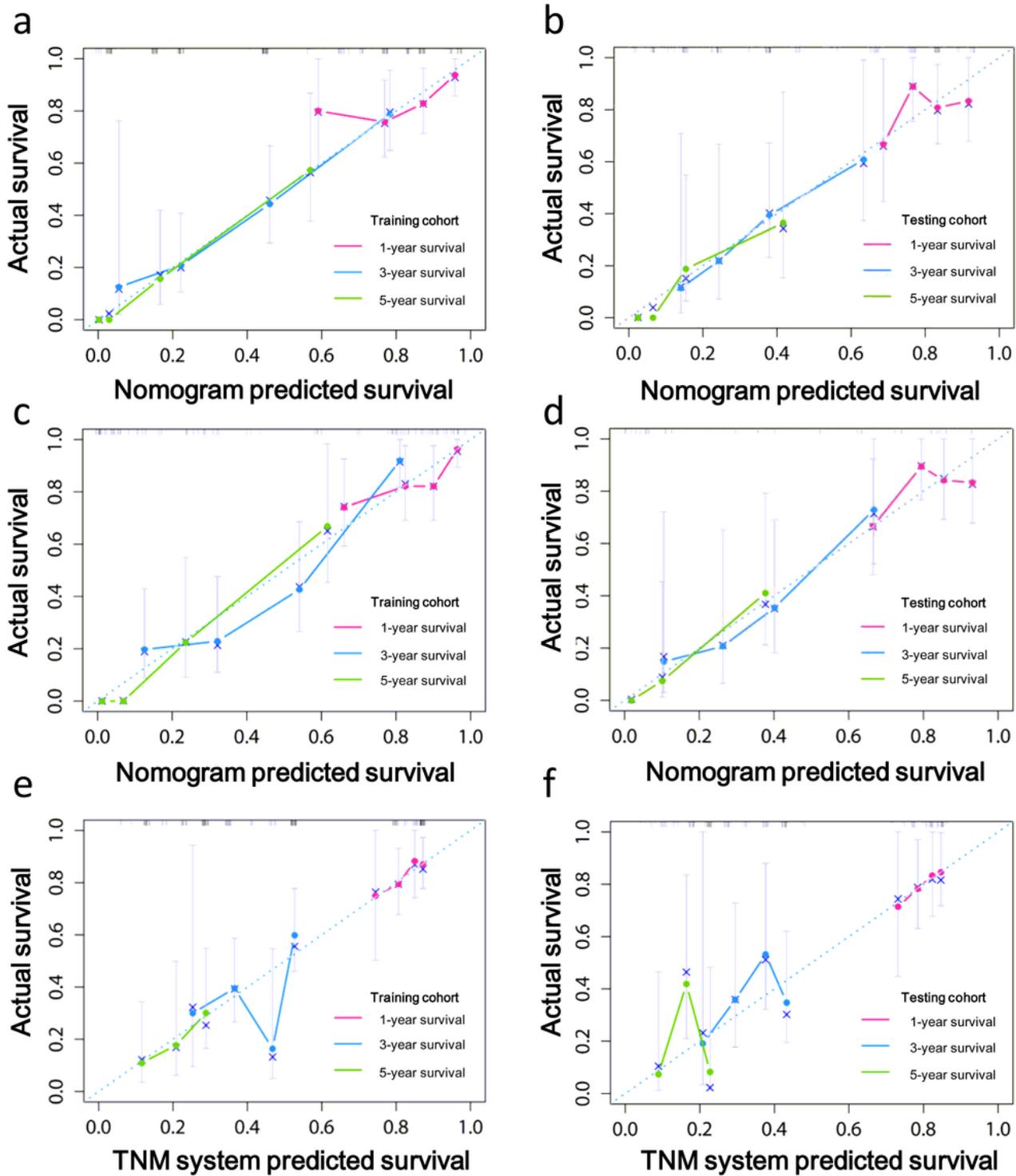


Figure 4

Calibration curves of models to predict survival probability. The curves of Modelclinic showed good agreement between the predicted and the actual 1-, 3- and 5-year OS in the training cohort (a) and testing cohort (b). The curves of Modelclinic&AP&PVP showed favorable agreement between the predicted and

the actual 1-, 3- and 5-year OS in the training cohort (c) and testing cohort (d). The curves of AJCC TNM staging system showed fair agreement between the predicted and the actual 1-, 3- and 5-year OS in the training cohort (e) and testing cohort (f). Model predicted OS probability was plotted on the x-axis. The actual OS rate was plotted on the y-axis. The closeness of the lines indicated absence of systematic bias. All CIs lie over the 45° dotted line of perfect calibration in Modelclinic and Modelclinic&AP&PVP.

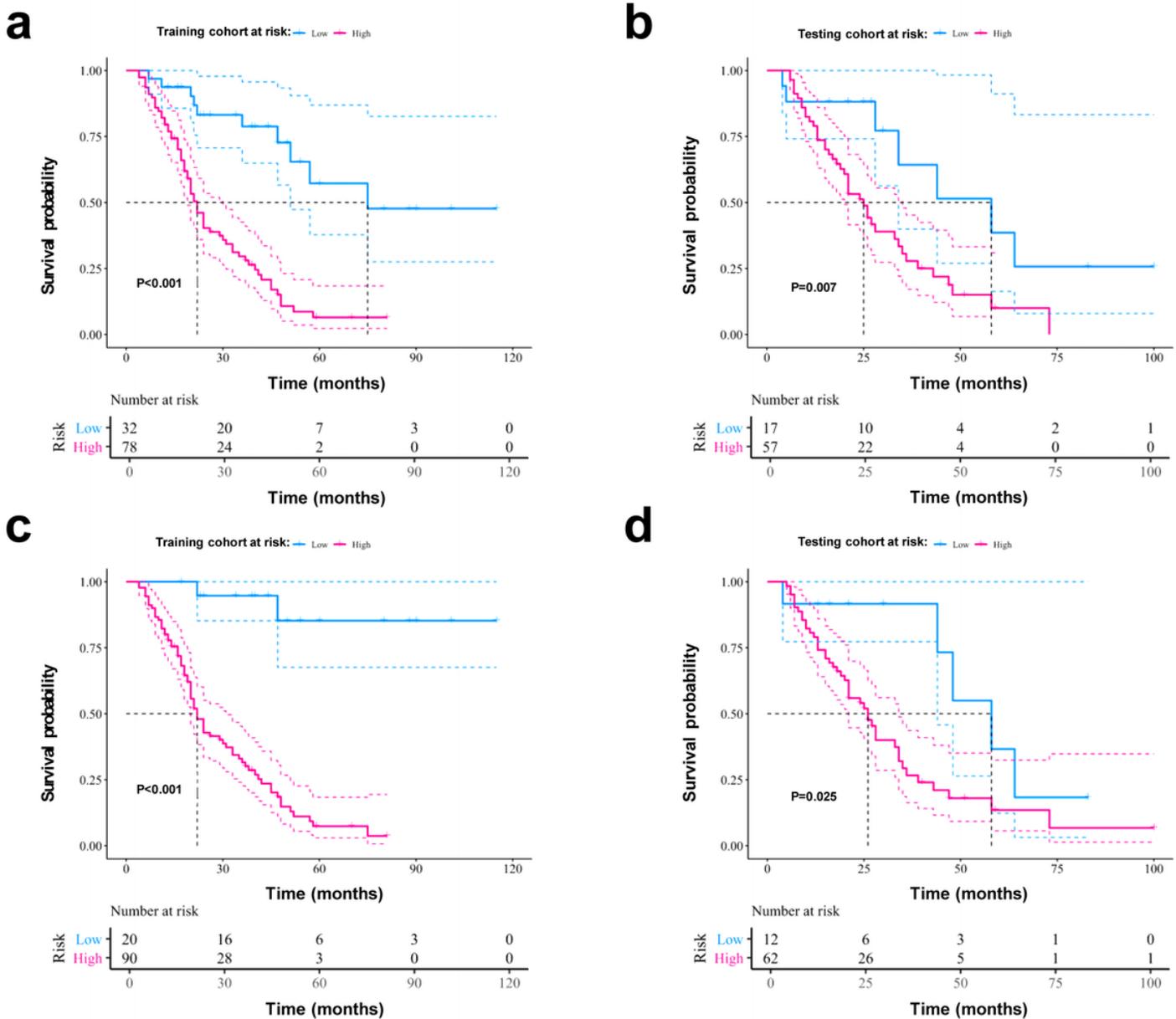


Figure 5

Kaplan-Meier curves of survival between high-risk and low-risk groups stratified by models. (a) Training cohort stratified by Modelclinic; (b) Testing cohort stratified by Modelclinic; (c) Training cohort stratified by Modelclinic&AP&PVP; (d) Testing cohort stratified by Modelclinic&AP&PVP. P values were calculated using the log-rank test.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigS4.DCA.tif](#)
- [FigS4.DCA.tif](#)
- [FigS3.Lasso.tif](#)
- [FigS3.Lasso.tif](#)
- [FigS2.SchoenfeldTest.tif](#)
- [FigS2.SchoenfeldTest.tif](#)
- [Appendix1radiomicsfeaturedescription.docx](#)
- [Appendix1radiomicsfeaturedescription.docx](#)
- [FigS1.Patientrecruitmentprocess.tif](#)
- [FigS1.Patientrecruitmentprocess.tif](#)