Journal of Hepatocellular Carcinoma

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

A Machine Learning Model Based on Counterfactual Theory for Treatment Decision of Hepatocellular Carcinoma Patients

Xiaoqin Wei[1,*], Fang Wang[2,*], Ying Liu[3], Zeyong Li[4], Zhong Xue[2], Mingyue Tang[5], Xiaowen Chen[1]

[1]School of Medical Imaging, North Sichuan Medical College, Nanchong City, Sichuan Province, People's Republic of China; [2]Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd, Shanghai, People's Republic of China; [3]Department of Radiology, The First Affiliated Hospital of Chengdu Medical College, Chengdu City, Sichuan Province, People's Republic of China; [4]Department of Radiology, Bishan Hospital of Chongqing Medical University, ChongQing, People's Republic of China; [5]Department of Physics, School of Basic Medicine, North Sichuan Medical College, Nanchong, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xiaowen Chen, School of Medical Imaging, North Sichuan Medical College, 234 Fujiang Road, Nanchong City, Sichuan Province, People's Republic of China, 637001, Email 56833804@qq.com

Purpose: To predict the efficacy of patients treated with hepatectomy and transarterial chemoembolization (TACE) based on machine learning models using clinical and radiomics features.

Patients and Methods: Patients with HCC whose first treatment was hepatectomy or TACE from June 2016 to July 2021 were collected in the retrospective cohort study. To ensure a causal effect of treatment effect and treatment modality, perfectly matched patients were obtained according to the principle of propensity score matching and used as an independent test cohort. Inverse probability of treatment weighting was used to control bias for unmatched patients, and the weighted results were used as the training cohort. Clinical characteristics were selected by univariate and multivariate analysis of cox proportional hazards regression, and radiomics features were selected using correlation analysis and random survival forest. The machine learning models ($Death_{hepatectomy}$ and $Death_{TACE}$) were constructed to predict the probability of patient death after treatment (hepatectomy and TACE) by combining clinical and radiomics features, and an optimal treatment regimen was recommended. In addition, a prognostic model was constructed to predict the survival time of all patients.

Results: A total of 418 patients with HCC who received either hepatectomy (n=267, mean age, 58 years ± 11 [standard deviation]; 228 men) or TACE (n=151, mean age, 59 years ± 13 [standard deviation]; 127 men) were recruited. After constructing the machine learning models $Death_{hepatectomy}$ and $Death_{TACE}$, patients were divided into the hepatectomy-preferred and TACE-preferred groups. In the hepatectomy-preferred group, hepatectomy had a significantly prolonged survival time than TACE (training cohort: $P < 0.001$; testing cohort: $P < 0.001$), and vise versa for the TACE-preferred group. In addition, the prognostic model yielded high predictive capability for overall survival.

Conclusion: The machine learning models could predict the outcomes difference between hepatectomy and TACE, and prognostic models could predict the overall survival for HCC patients.

Keywords: radiomics, hepatocellular carcinoma, prognosis, hepatectomy

Introduction

Different treatment guidelines were proposed to provide appropriate treatment options for hepatocellular carcinoma (HCC) patients, such as China Liver Cancer (CNLC) staging[1] and Barcelona Clinic Liver Cancer (BCLC) staging.[2] However, they might be inadequate for clinical decision-making. First, there is disagreement among treatment guidelines over the recommended treatment regimens. In CNLC, hepatectomy is recommended for patients with Ia, Ib, and IIa stages, and transarterial chemoembolization (TACE) is recommended for patients with IIb, IIIa, and some IIIb stages.[1] In BCLC, patients with Ia and Ib stages are defined as 0 or A stage, IIa and IIb are defined as B stage, and IIIa and IIIb are

Received: 25 March 2024
Accepted: 13 August 2024
Published: 30 August 2024

defined as C stage.[2] Hepatectomy, TACE, and systemic treatments are recommended for different stages respectively.[2] Second, there exists heterogeneity among patients at the same stage, and survival times for patients even with the same treatment regimen are different.[3,4] Finally, in clinical practice, 50% of patients have deviated from their original recommendations based on the treatment guidelines.[4,5] Therefore, the treatment regimens for HCC patients are still unclear.

Several studies[6–10] have found that TACE and hepatectomy have similar survival rates for patients with early-stage HCC. The question of which treatment is superior remains unresolved. Previous studies comparing the efficacy of hepatectomy versus TACE in patients with single HCC have shown that patients in the hepatectomy group had higher survival rates at 1, 3, and 5 years than those in the TACE-treated group.[11]. However, the study of Lee et al[12] presented contradictory results, with similar 1-, 3-, and 5-year survival rates for hepatectomy and TACE. In addition, TACE treatment was considered the standard treatment option for patients with stage BCLC-B in 2012, with the highest level of recommendation being 1A.[13] When the guidelines were proposed, there was insufficient data to conclusively support the efficacy differences between TACE treatment and other treatments such as hepatectomy. In addition, hepatectomy has been shown to extend the survival of some patients with intermediate-stage HCC.[6,14–18] Hepatectomy is the treatment of choice for patients with type I/II portal vein tumor thrombus when the lesion is resectable.[19,20] In patients with advanced-stage HCC, TACE is relatively safe and has significantly better survival rates.[21–24] Thus, it is controversial which patients could benefit from hepatectomy or TACE.

To measure the causal effect of hepatectomy over TACE, the outcomes between the two treatments in the same individual would have to be compared.[25,26] So, each patient with liver cancer had two outcomes: the outcome after receiving the actual treatment (observed) and the counterfactual outcome if the patient had received a different treatment (unobserved).[26] However, counterfactual outcomes are difficult to observe and counterfactual cases may be missed even in randomized controlled trials.[26] This study attempted to estimate the counterfactual outcome of patients treated with transhepatic resection and TACE by machine learning techniques. Therefore, a machine learning model built upon the counterfactual theory can effectively compare the efficacy of hepatectomy with TACE based on retrospective data.

In this study, we first estimate the counterfactual outcomes of HCC patients treated with hepatectomy and TACE using retrospective data based on machine learning model, and construct a prognostic model to predict the overall survival of patients.

Material and Methods

Patients

This study was approved by the Ethics Review Committee of the Affiliated Hospital of North Sichuan Medical College (Number: 2022ER059-2), and the informed consent of the participants was waived for this retrospective study. All the patient identity information has been anonymized. The research was conducted in accordance with the principles of the Declaration of Helsinki.

The workflow of the study is shown in Figure 1. Patients with HCC whose first treatment option was hepatectomy or TACE at the Affiliated Hospital of North Sichuan Medical College from June 2016 to July 2021 were retrospectively collected in the cohort study. The inclusion and exclusion criteria are described in supplementary text 1.

We created a standardized form for collecting the clinical variables (supplementary text 2). The initial treatment was decided based on CNLC guidelines and the patient's willingness.

To ensure a causal effect of efficacy and treatment regimen, patients with similar baseline characteristics were screened by propensity score matching (PSM) and used as an independent testing cohort (supplementary text 3). The supplementary table 1 shows the correlation coefficients of baseline characteristics with the treatment regimen, survival time, and survival status. For the remaining unmatched patients, the inverse probability of treatment weighting (IPTW) was used to balance the confounders between treatment groups, and the weighted patients were used as the training cohort (supplementary text 3).

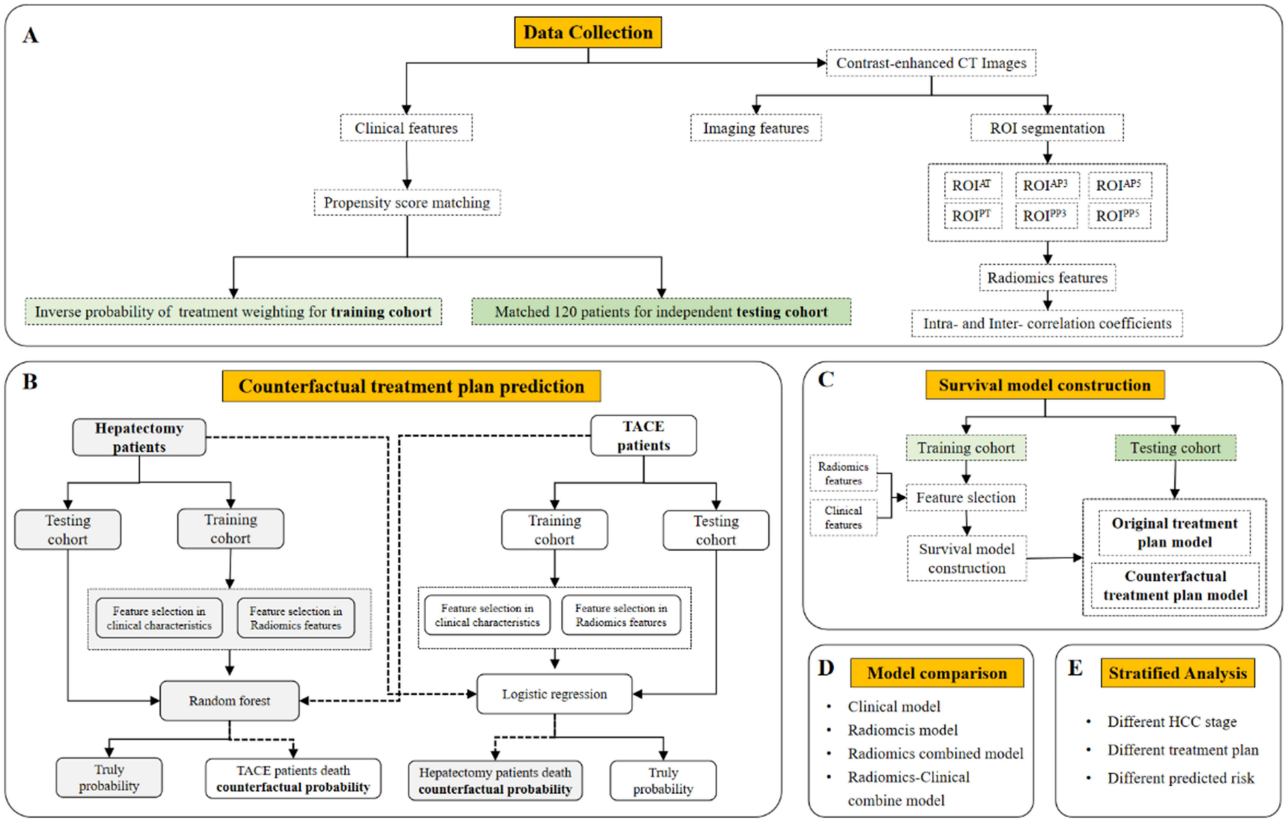

Figure 1 The workflow of the study. (**A**) Collecting clinical features and extracting the radiomics features from CT images. (**B**) Predicting the patients undergoing the hepatectomy or TACE death counterfactual probability. (**C**) Building prognostic model to predicting the prognosis of HCC patients. (**D**) Comparing the performance of the models. (**E**) Analyzing the predictive ability of the model in different subgroups.
Abbreviations: ROI, region of Interest; TACE, transarterial chemoembolization; HCC, hepatocellular carcinoma.

The primary endpoint was the overall survival (OS). The definition of OS and the follow-up program is described in supplementary Text 4. As of the last visit, 36 patients were lost to follow-up. Patients lost to follow-up were censored for survival.

Pre-Processing, Segmentation, and Extraction of Radiomics Features

All the enhanced CT examinations were completed within one month prior to treatment. The scanning parameters and acquisition protocols are shown in supplementary table 2 and supplementary text 5. Supplementary text 6 shows the image pre-processing information.

Two radiologists manually delineated the regions of interest (ROIs) along each tumor margin and peritumor regions using 3D Slicer (http://www.slicer.org). Details of annotation and its reproducibility assessment are described in supplementary text 6. Six ROIs were obtained for each subject, including tumoral ROI at the arterial phase (A), tumoral ROI at the portal phase (P), peritumoral regions within 3mm/5mm distances around the tumor at the arterial phase (A3/A5), and peritumoral regions within 3mm/5mm distances around the tumor at the portal phases (P3/P5).

Radiomics features were extracted from the pre-processed images based on these annotated ROIs by using the Pyradiomics package (https://pyradiomics.readthedocs.io/en/v3.0.1/), see supplementary text 7 for details. Z-score normalization was conducted for the radiomics features to reduce the deviance of dimension.

Treatment Efficacy Prediction

For patients in two treatment regimens (hepatectomy and TACE group), two survival prediction models ($Death_{hepatectomy}$ and $Death_{TACE}$) were constructed respectively, so that patients' survival status can be predicted from their clinical

characteristics and imaging radiomics characteristics. Supplementary text 8 shows the details of the model construction. The $Death_{hepatectomy}$ and $Death_{TACE}$ models were then used to predict the survival (probability of death) and the counterfactual survival for each patient assuming their treatment regimens were switched. For patient *l* in the hepatectomy group, the probability of death predicted by the $Death_{hepatectomy}$ model ($\hat{p}_l$) was regarded as the proper or fitting prediction, whereas the probability of death predicted by applying the $Death_{TACE}$ model ($\hat{p}'_l$) was considered as the counterfactual prediction. $\hat{p}_l$ and $\hat{p}'_l$ were calculated as follows:

$$\hat{p}_l = p(Y_{i=1}|T_{i=1}, X_{i=1}), \# \quad (1)$$

$$\hat{p}'_l = p(Y_{i=1}|T_{i=0}, X_{i=0}), \# \quad (2)$$

$$Z_{i=1} = \hat{p}_l - \hat{p}'_l, \# \quad (3)$$

where T_i is the treatment regimen of group *i* (i=1 means hepatectomy group, i=0 represents TACE group), Y_i is the survival status of group *i*, $_i$ represents predictive variable (features) for group *i*. Z_i is the difference of the probabilities of death between the two predictions. If Z_i is greater than a prescribed constant value C ($C=0.1$),[27] the patient can be considered to be more favorable for survival with the TACE, otherwise the favorable treatment remains the same.

Subsequently, we compared the prognosis of the TACE group using the same machine learning method, and the favorable treatment for each patient can be re-determined. Finally, all the patients were regrouped into hepatectomy- preferred and TACE-preferred groups. Then, we used stratified analysis to verify effectiveness of the new grouping, and the clinical characteristics of each group are more distinguishable and interpretable and are also more consistent with the guidelines.

Feature Selection of Prognostic Model

Clinical characteristics were selected by univariate and multivariate analysis of cox proportional hazards regression (COX). Then, prognostic models were constructed based on these selected clinical characteristics, which exhibited *p*-values below 0.05 in multivariate Cox analysis.

Correlation analysis and random survival forest (RSF) were used to select radiomics features from CT images. The former was used to eliminate variables with collinearity, and the latter was applied to rank the importance of the remaining non-collinearity features. Top 30 features were selected for each RSF operation, and the RSF feature selecting was performed 100 times, so that the top features appeared 80 times and above were finally maintained. The weighting of each feature was assigned based on its ranking of importance, and Radscores were calculated by linear multiplication of the features and weights (Radscore_A, Radscore_P, Radscore_A3, Radscore_P3, Radscore_A5, and Radscore_P5).

Prognostic Model Construction

We used the accelerate failure time model (AFT) to estimate the effect of covariates on patient survival time (supplementary text 9). The standard semiparametric AFT model correlates covariates to the logarithm of survival time through the following regression model:

$$Log(t_l) = \beta^T \varphi_l + \varepsilon, \# \quad (4)$$

where t_l represents the survival time of patient *l*, β is the parameter vector, φ_1 is the covariate vector, and ε is a constant (8.455 in this study). Minimization of mean squared error can be used to estimate the regression parameter given a number of patient data.

Statistical Analysis

Open-source R (version 4.1.1) software was used for data analysis. Confounding factors between the two treatment groups were determined by correlation analysis. In the description of the characteristic distribution, the Kolmogorov–Smirnov test was used to test the normal distribution of the continuous variables. For normally distributed variables,

t-tests were used and expressed as mean ± SD (standard deviation). Otherwise, the Mann–Whitney *U*-test was used and expressed as the median (inter-quartile range). For categorical variables, the chi-square test or Fisher test was used and expressed as counts (%). $p < 0.05$ was considered statistically significant. In addition, stratified analysis using clinical factors was performed to verify the predictive performance of machine learning models and the prognostic performance of the AFT model.

So that other researchers can independently replicate our findings, all experimental steps have been described in detail and completely. All the experimental analyses were performed with open-source software, including R software (http://www.r-project.org), 3D slicer (http://www.slicer.org), and Pyradiomics package (https://pyradiomics.readthedocs.io/en/v3.0.1/). And the data and code analyzed during the study are publicly available through the Treatment-efficacy-prediction GitHub repository (https://github.com/lalala111-hai/Treatment-efficacy-prediction).

Results

Patients

A total of 418 patients were enrolled in the study, of whom 267 (mean age, 58 years ± 11 [standard deviation]; 228 men) were treated with hepatectomy, and 151 (mean age, 59 years ± 13 [standard deviation]; 127 men) were treated with TACE. The flow chart of patient recruitment is shown in the supplementary figure 1. The median of the follow-up intervals was 864 days for the entire cohort (inter-quartile range [IQR]: 657–1004), 919 days for the training cohort (IQR: 643–1037), and 856 days for the testing cohort (IQR: 616–1253). The clinical characteristics of the patients are shown in the supplementary table 3.

One hundred and twenty patients were matched by using PSM to form an independent testing cohort (TACE group, 60 patients; hepatectomy group, 60 patients). The remaining 298 patients (TACE group, 91 patients; hepatectomy group, 207 patients) were used as the training cohort, and confounding variables were treated by IPTW. The PSM and IPTW results demonstrated there is no significant difference in clinical characteristics between the two treatment groups ($P > 0.05$) (supplementary table 4, supplementary table 5).

Treatment Efficacy Prediction

Supplementary text 8 and supplementary table 6 shows the details of the model construction.Table 1 shows the performance of different models in the hepatectomy and TACE groups. The result shows that the RFS model ($Death_{hepatectomy)}$ built by combining clinical characteristics, Radscore_A3 and Radscore_P has the best performance (training cohort: AUC=0.839 [0.781–0.896]; testing cohort: AUC=0.838 [0.727–0.948]). And the logistic regression model ($Death_{TACE}$) developed by combining clinical characteristics and Radscore_A5 was effective in predicting the probability of death (training cohort: AUC=0.790 [0.661–0.919]; testing cohort: AUC=0.803 [0.690–0.916]).

When applying the $Death_{TACE}$ model to predict the probability of death for the hepatectomy (counterfactual treatment) and the TACE group (fitting group), Figure 2 shows the results. Nineteen patients in the hepatectomy group were counterfactually predicted to be treated with TACE; 89 patients in the TACE group were suggested to be treated with hepatectomy (supplementary table 7). Patients were divided into the hepatectomy-preferred group and TACE-preferred group based on the results predicted by the machine learning model. In the hepatectomy-preferred group, hepatectomy had a prolonged survival time than TACE in both the training cohort ($P < 0.001$) and the testing cohort ($P < 0.001$) (Figure 3A–C). In the TACE-preferred group, patients treated with TACE had longer overall survival than hepatectomy in both the training cohort ($P = 0.011$) and the testing cohort ($P < 0.001$), indicating that TACE might be a better treatment option than hepatectomy (Figure 3B–D). In addition, in the hepatectomy group, hepatectomy-preferred patients had a prolonged survival time than TACE-preferred patients in both the training cohort ($P < 0.001$) and the testing cohort ($P < 0.001$) supplementary Figure 2 A–C). There were similar results in TACE treatment (supplementary figure 2 B–D).

Table 1 Performances of the Different Models in the Hepatectomy and TACE Groups

Treatment Group	Models	Group	AUC (95% CI)	F1 Score	Precision	Sensitivity	Specificity	Accuracy
Hepatectomy	Clinical	train	0.850(0.794–0.906)	0.732	0.677	0.798	0.740	0.763
Hepatectomy		test	0.777(0.641–0.913)	0.843	0.795	0.897	0.571	0.783
Hepatectomy	Radscore_A	train	0.726(0.655–0.797)	0.562	0.592	0.536	0.748	0.662
Hepatectomy		test	0.574(0.428–0.720)	0.533	0.762	0.410	0.762	0.533
Hepatectomy	Radscore_A3	train	0.704(0.631–0.776)	0.615	0.546	0.702	0.602	0.643
Hepatectomy		test	0.635(0.489–0.781)	0.703	0.743	0.667	0.571	0.633
Hepatectomy	Radscore_A5	train	0.723(0.654–0.791)	0.593	0.496	0.738	0.488	0.589
Hepatectomy		test	0.590(0.435–0.744)	0.711	0.730	0.692	0.524	0.633
Hepatectomy	Radscore_P	train	0.740(0.672–0.807)	0.632	0.560	0.726	0.610	0.657
Hepatectomy		test	0.736(0.603–0.869)	0.773	0.806	0.744	0.667	0.717
Hepatectomy	Radscore_P3	train	0.699(0.626–0.772)	0.507	0.597	0.44	0.797	0.652
Hepatectomy		test	0.594(0.451–0.737)	0.580	0.667	0.513	0.524	0.517
Hepatectomy	Radscore_P5	train	0.692(0.619–0.766)	0.542	0.549	0.536	0.699	0.633
Hepatectomy		test	0.687(0.548–0.826)	0.636	0.778	0.538	0.714	0.600
Hepatectomy	Clinical+A3	train	0.857(0.802–0.911)	0.732	0.677	0.798	0.740	0.763
Hepatectomy		test	0.816(0.695–0.937)	0.84	0.810	0.872	0.619	0.783
Hepatectomy	Clinical+P	train	0.860(0.807–0.913)	0.724	0.700	0.75	0.780	0.768
Hepatectomy		test	0.805(0.681–0.928)	0.854	0.814	0.897	0.619	0.800
Hepatectomy	Clinical+A3+P	train	0.839(0.781–0.896)	0.740	0.719	0.762	0.797	0.783
Hepatectomy		test	0.838(0.727–0.948)	0.800	0.833	0.769	0.714	0.750
TACE	Clinical	train	0.772(0.640–0.904)	0.833	0.615	0.802	0.929	0.878
TACE		test	0.724(0.587–0.861)	0.618	0.731	0.667	0.750	0.678
TACE	Radscore_A	train	0.569(0.429–0.709)	0.654	0.231	0.593	0.836	0.734
TACE		test	0.507(0.356–0.657)	0.206	0.808	0.467	0.583	0.304
TACE	Radscore_A3	train	0.621(0.415–0.828)	0.038	1.000	0.176	1.000	0.073
TACE		test	0.672(0.535–0.809)	0.147	1.000	0.517	1.000	0.256
TACE	Radscore_A5	train	0.564(0.378–0.751)	0.064	1.000	0.198	1.000	0.120
TACE		test	0.648(0.506–0.791)	0.147	0.962	0.500	0.833	0.250
TACE	Radscore_P	train	0.567(0.426–0.708)	0.654	0.231	0.593	0.836	0.734
TACE		test	0.509(0.358–0.660)	0.206	0.808	0.467	0.583	0.304
TACE	Radscore_P3	train	0.447(0.280–0.614)	0.949	0.077	0.824	0.860	0.902
TACE		test	0.521(0.373–0.670)	0.029	0.962	0.433	0.500	0.055
TACE	Radscore_P5	train	0.592(0.441–0.743)	0.949	0.000	0.813	0.851	0.897
TACE		test	0.601(0.454–0.748)	0.206	0.846	0.483	0.636	0.311
TACE	Clinical+A	train	0.792(0.663–0.921)	0.833	0.615	0.802	0.929	0.878
TACE		test	0.714(0.575–0.852)	0.794	0.577	0.700	0.711	0.750
TACE	Clinical+A3	train	0.770(0.633–0.908)	0.808	0.615	0.78	0.926	0.863
TACE		test	0.792(0.673–0.910)	0.794	0.615	0.717	0.73	0.761
TACE	Clinical+A5	train	0.790(0.661–0.919)	0.821	0.615	0.791	0.928	0.871
TACE		test	0.803(0.690–0.916)	0.794	0.654	0.733	0.750	0.771
TACE	Clinical+P	train	0.791(0.661–0.920)	0.833	0.615	0.802	0.929	0.878
TACE		test	0.713(0.574–0.852)	0.794	0.577	0.700	0.711	0.750
TACE	Clinical+P3	train	0.784(0.651–0.917)	0.808	0.615	0.780	0.926	0.863
TACE		test	0.757(0.633–0.881)	0.794	0.577	0.700	0.711	0.750
TACE	Clinical+P5	train	0.806(0.688–0.923)	0.859	0.538	0.813	0.918	0.888
TACE		test	0.740(0.609–0.871)	0.765	0.692	0.733	0.765	0.765
TACE	Clinical+A3+P5	train	0.800(0.676–0.924)	0.808	0.615	0.780	0.926	0.863
TACE		test	0.790(0.676–0.904)	0.794	0.654	0.733	0.750	0.771

Abbreviations: TACE, transarterial chemoembolization; Radscore_A, radiomics signature of tumoral ROI of the arterial phases; Radscore_P, radiomics signature of tumoral ROI of the portal phases; Radscore_A3, radiomics signature of peritumoral 3 mm ROI of the arterial phases; Radscore_P3, radiomics signature of peritumoral 3 mm ROI of the portal phases; Radscore_A5, radiomics signature of peritumoral 5 mm ROI of the arterial phases; Radscore_P5, radiomics signature of peritumoral 5 mm ROI of the portal phase.

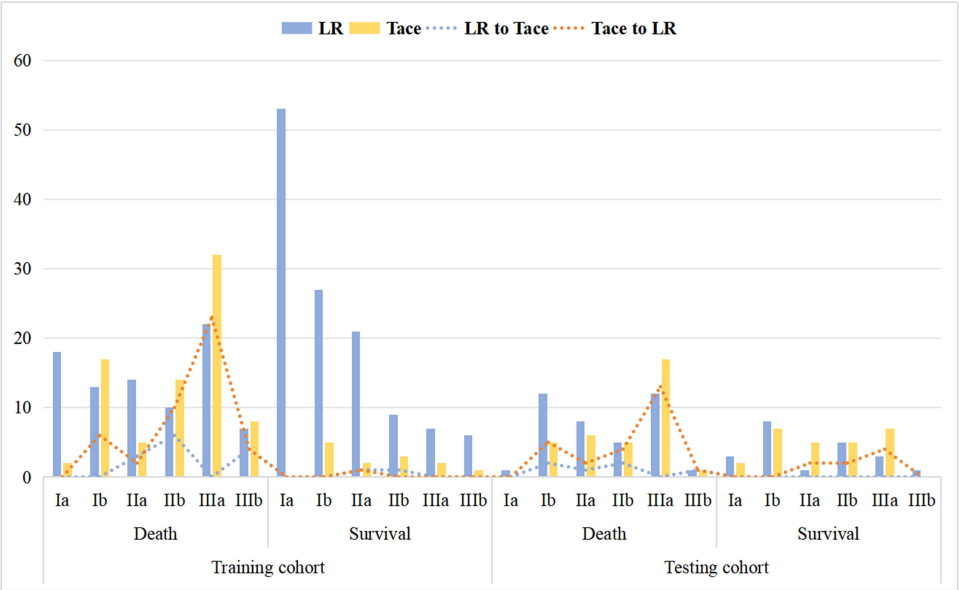

Figure 2 Real and counterfactual treatment regimens. The bar plot shows the real treatment regimens, and the dotted lines indicates the counterfactual treatment regimens.
Abbreviations: LR, Liver resection; Tace, transarterial chemoembolization.

Feature Selection of Prognostic Model

Univariate and multivariate COX regression results showed (supplementary table 8) that treatment regimen (p=0.033, HR=1.440, 95% CI=1.030–2.010), ALBI grade (p=0.045, HR=1.420, 95% CI= 0.998–2.030), Tumor number (p= 0.018, HR=1.230, 95% CI= 1.040–1.460) and GGT (p=0.029, HR=1.230, 95% CI=1.020–1.480) were significantly associated with survival status in HCC patients. Covariance analysis and RFS were used for radiomics features selection, and the results of selected features for the six ROIs are shown in supplementary figure 3.

Prognostic Model Construction

Using the minimum Akaike information criteria (AIC) as the model selection criterion, the AFT model built by combining clinical characteristics, Radscore_A and Radscore_P5 had the best performance (AIC=73.391). Using the results of the machine learning model as a treatment regimen, the AIC value of the model was reduced to 69.753. The specific variables of the prognostic model are shown in Table 2 (all variables $p<=0.05$). In the validation cohort, the time-ROC curves (Figure 4) show that compared to the actual treatment regimen (Figure 4A and B), the counterfactual treatment regimen (Figure 4C and D) improved the 1–5 years AUC by 7.6% (0.728 vs 0.652), 7.7% (0.814 vs 0.737), 8.2% (0.789 vs 0.707), 4.3% (0.811 vs 0.768) and 4.3% (0.811 vs 0.768). The integrated Brier scores of the prognostic model are 0.167 in the training cohort and 0.168 in the testing cohort.

The median risk score, 6.748, predicted by the prognostic model was used as a threshold to classify HCC patients into high-risk and low-risk groups. The cumulative mortality at 1- to 3-year for high-risk and low-risk patients were 57.02%, 67.03%, 67.03% and 16.78%, 30.87%, 34.9%, respectively. The Kaplan-Meier survival curve of the OS is shown in Figure 5.

Stratified Analysis

Stratified analysis of patients by Cirrhosis (normal or abnormal), CNLC (I or II/ III), and Child-Pugh classification (grade A or B) showed that surgical treatment had significantly longer survival than TACE treatment in the hepatectomy-preferred group with $p < 0.001$ (supplementary figures 4–6). In contrast, for the TACE-preferred group, TACE treatment was associated with significantly longer survival ($p < 0.05$, supplementary figures 4–6). The prognostic model enables for risk stratification in different populations, with survival times significantly lower in the high-risk group than in the low-risk group ($p\leq0.001$, supplementary figure 7).

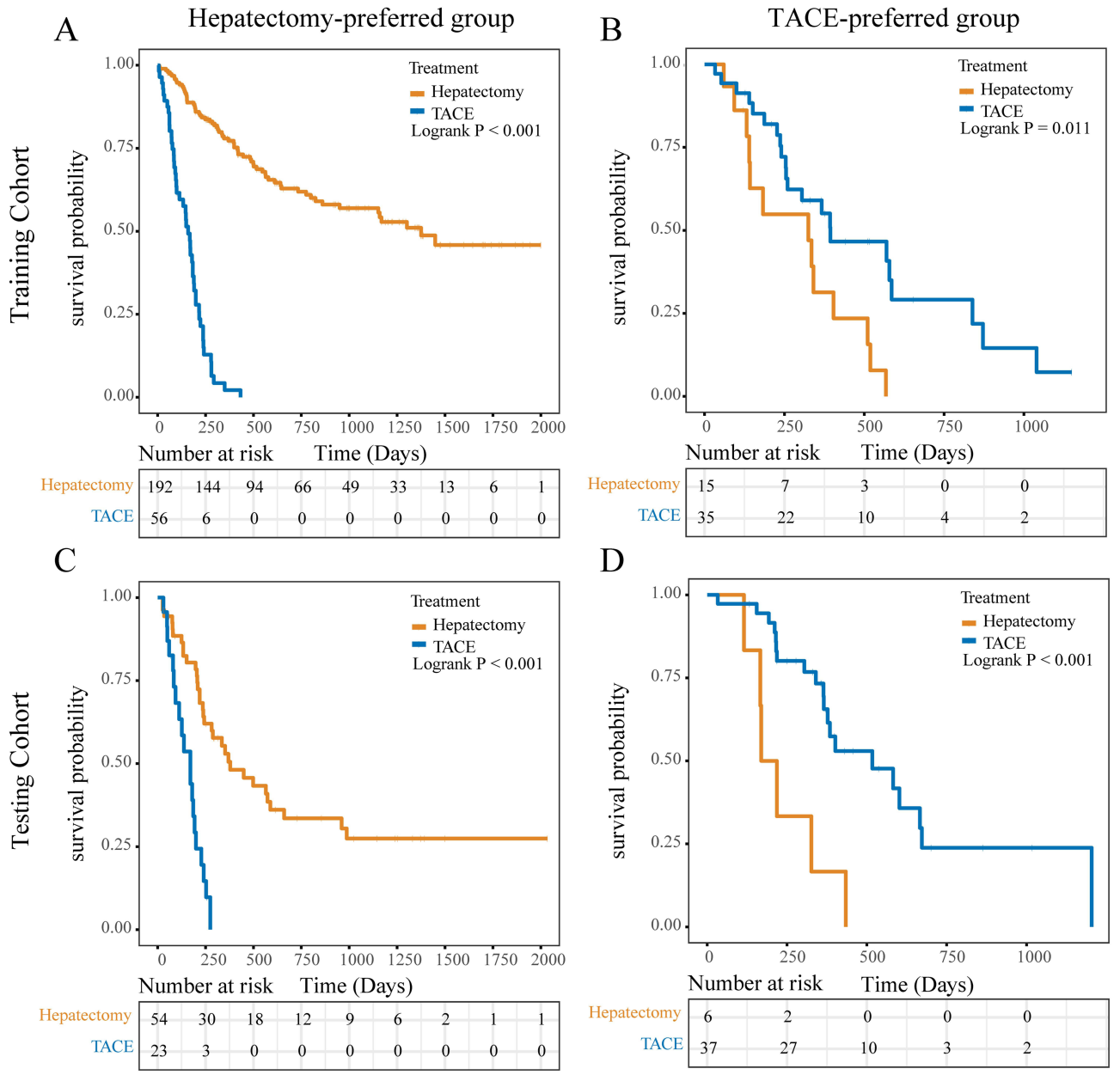

Figure 3 Kaplan-Meier curves of prognostic differences between hepatectomy and TACE treatment. Comparisons of the hepatectomy and TACE groups in the Hepatectomy-preferred group (**A**), and the TACE-preferred group (**B**) in the training cohort. Comparisons of the hepatectomy and TACE groups in the Hepatectomy-preferred group (**C**), and TACE-preferred group (**D**) in the testing cohort. TACE, transarterial chemoembolization.
Abbreviation: TACE, transarterial chemoembolization.

Discussion

In this study, a predictive model for hepatectomy versus TACE treatments in HCC patients was developed based on counterfactual theory. The prognosis of patients with real treatment regimens versus counterfactual treatment regimens was compared. Our machine learning models estimated patients who might be beneficial from the preferred treatment regimen (hepatectomy and TACE), and thus recommend the treatment option for the patients that may be of greater benefit. In addition, a prognostic model for predicting the overall survival time of HCC patients was constructed, which showed high discrimination in both the training and testing cohorts.

Traditional decision support systems such as the CNLC staging and the BCLC staging classify patients into different grades and recommend corresponding treatment strategies. Although clinical guidelines are important for prognosis and

Table 2 Variables of the Prognostic Model

Factors	Coefficients	Stand error	z value	p value
Radscore_A	−1.562	0.394	−3.96	<0.001
Radscore_P5	−1.173	0.525	−2.24	0.025
Cirrhosis	−0.53	0.116	−4.59	<0.001
ALBI grade	−0.314	0.085	−3.69	<0.001
Treatments	−0.246	0.126	−1.96	0.05
Tumor number	−0.109	0.025	−4.31	<0.001
GGT	−0.001	0.000	−2.99	0.003

Abbreviations: ALB, serum albumin; GGT, Glutamyl transpeptidase.

treatment, the decision-making capacity may be inadequate.[2,15,28–30] Thus, an increasing number of studies have compared the efficacy of hepatectomy versus TACE treatment in patients with the same stage, but even randomized controlled trials would not immune from the causal inference problem.[25] Although previous radiomics or deep learning

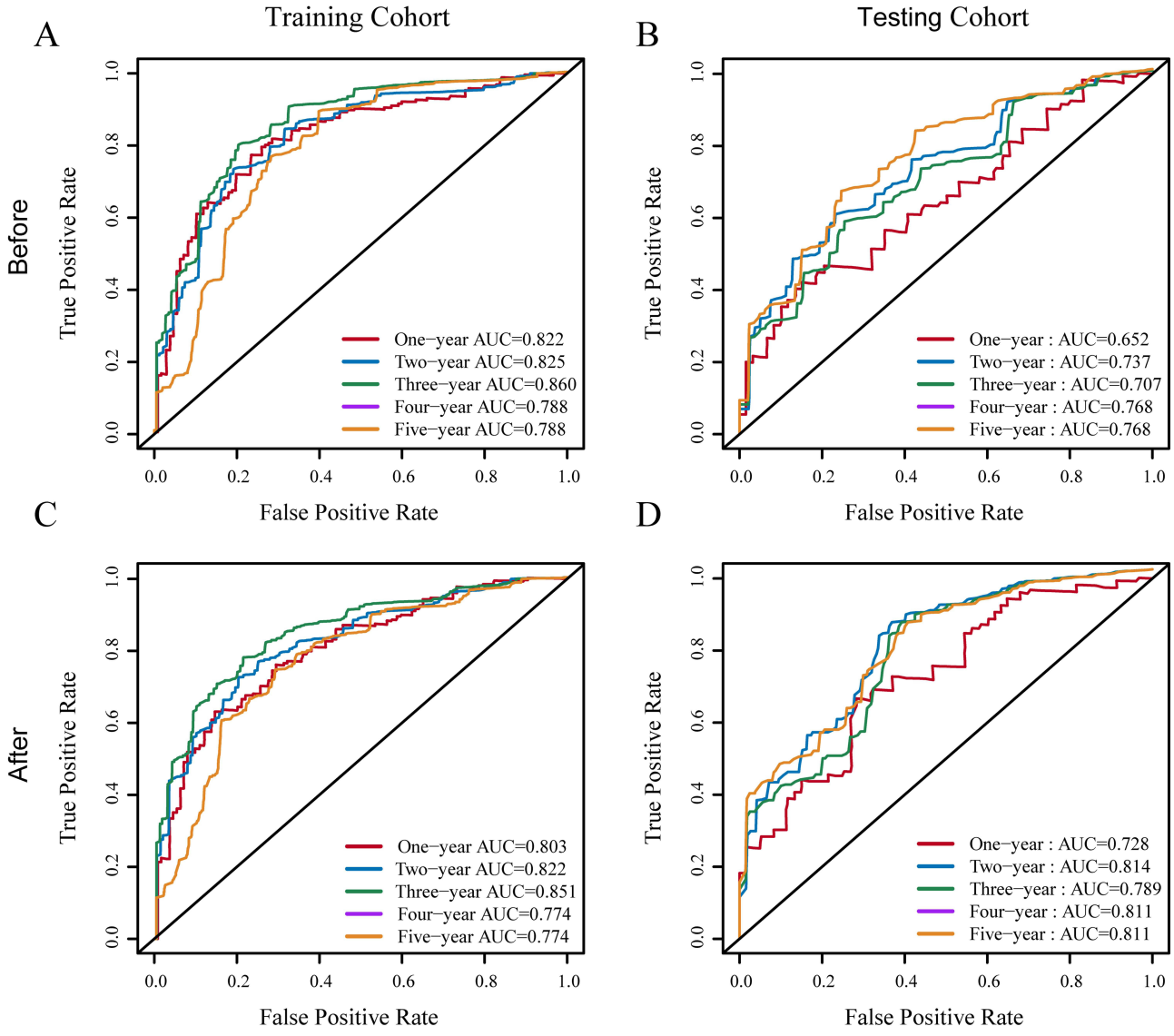

Figure 4 Time-ROC curves of the prognostic model. The time-ROC curves for the prognostic model of the actual treatment regimen (**A**), and the counterfactual treatment regimen (**B**) in the training cohort. The time-ROC curves for the prognostic model of the actual treatment regimen (**C**), and the counterfactual treatment regimen (**D**) in the testing cohort. **Abbreviation**: AUC, areas under the receiver operating characteristic curves.

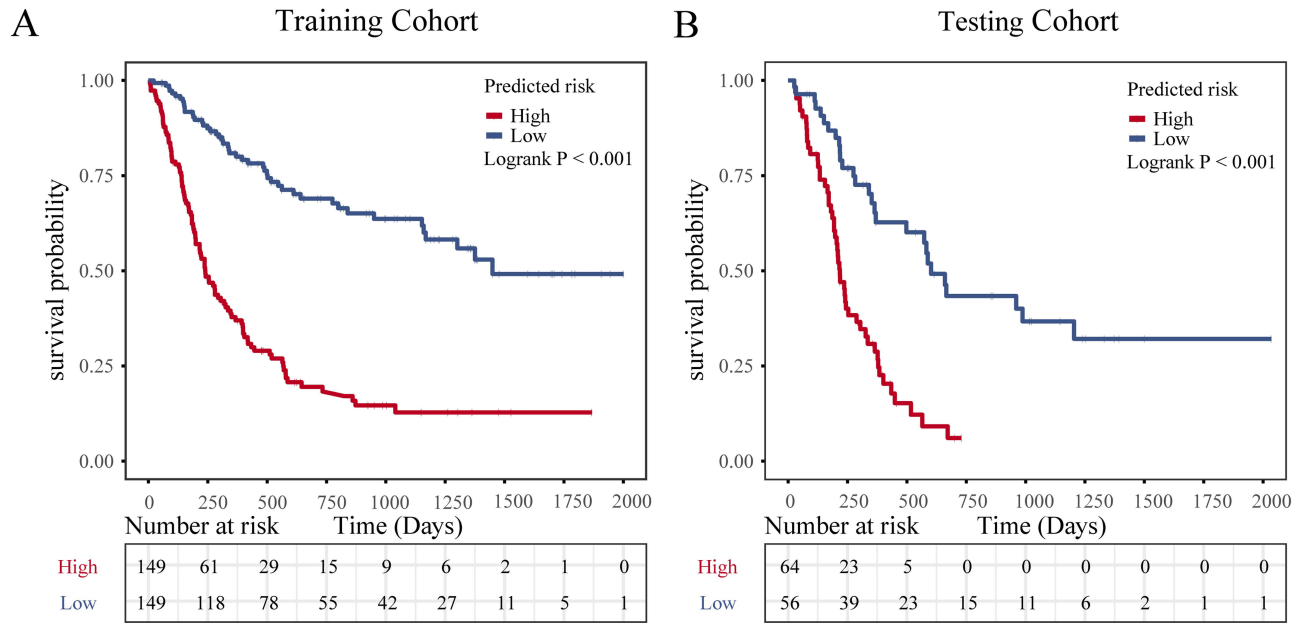

Figure 5 Kaplan-Meier curves of high-risk and low-risk patients. Kaplan-Meier curves of overall survival in the training cohort (**A**) and the test cohort (**B**).

studies on HCC have built prognostic models, they have not taken into account the impact of treatment regimens.[31,32] In contrast, Cucchetti et al[33] considered the effect of treatment regimens based on the counterfactual principle, but only estimated the mean effect of the treatments. The study only evaluated the overall efficacy of different treatments, so it could not give individualized treatment recommendations according to the patient's actual situation. The above studies may only provide limited help to clinicians in making treatment decisions. Therefore, this study inferred the prognosis of HCC patients taking different treatments based on the counterfactual principle. In addition, we provided a prognostic model that took into account treatment factors, which not only had high prognostic ability but also allowed for risk stratification. There is the first study designed to assess the efficacy and the prognosis of HCC patients in different treatments.

In this study, the probability of death was calculated for each patient after hepatectomy and TACE treatment. The results showed that 89 patients underwent TACE but the counterfactual predicted that hepatectomy was recommended, while only 19 patients who underwent surgery differed from the counterfactual results. As a whole, there may be a greater survival benefit with hepatectomy. This study confirmed previous findings, a large multicenter study comparing the benefits of hepatectomy at different stages of BCLC showed that hepatectomy was more beneficial to HCC patient survival in patients without liver dysfunction, regardless of the patient's BCLC stage.[28] Similarly, Cucchetti et al[33] took a counterfactual approach to assess the difference in efficacy between hepatectomy and TACE, showing that patients treated with hepatectomy rather than TACE would have increased survival time from 18.3 months to 38.0 months and survival rates by more than 50%. In addition, the difference in efficacy between hepatectomy and radiofrequency ablation was considered in the study by Cucchetti et al,[33] but was not discussed in our study. In subsequent studies, we will focus on the other treatment options, such as ablation, immunotherapy and chemotherapy.

In this study, there were patients with early, intermediate and advanced HCC whose actual surgical procedures differed from the counterfactual recommended treatment, suggesting that there is a large heterogeneity among patients with the same clinical stage and that a reasonable treatment should be selected for each patient. Similar conclusions were reached in a study by Golfieri et al[4] which showed that patients at the same clinical stage exhibited different tumor burden and liver function but were still recommended for the same treatment. The large variation in survival of patients at the same stage suggests that clinical staging guidelines do not fulfill the role of dividing patients into subgroups with significantly different prognosis.[4] In our study, each patient was further analyzed, combining clinical factors and quantitative imaging features to provide a more personalized treatment opinion. Specifically, when confronted with patients who have difficulties with decision-making choices, the probability

of death for each patient treated with hepatectomy and TACE was calculated and compared using the model from this study. If the probability of death is similar for those treated with hepatectomy and TACE, then TACE is recommended for pain control. Conversely, whichever treatment received had a lower probability of death was recommended. The predictive value of the probability of death provided in this study may provide a reference for physicians when making treatment decisions. In our study, a prognostic model to predict the OS of patients with HCC was also constructed. The prognostic model not only has high predictive ability but also could classify HCC patients into subgroups with significant differences in survival time. For patients in the high-risk group, a more rigorous follow-up strategy or consideration of additional adjuvant therapy should be adopted after surgery. For patients in the low-risk group, a more rational follow-up protocol could be adopted considering the risk-to-benefit ratio.

However, there are still some limitations of our study. First, the study only considered hepatectomy and TACE, and other treatment modalities should be included for comparison in future studies. Second, this retrospective study was observational and relied on statistical extrapolation to address the issue of causal inference, our results require further validation with other hospitals to check for the generalizability. Third, the sample size of the single-center study might be relatively small. Considering the complexity of the causal inference and treatment selection issues, the subsequent studies with the larger database to validate the results is needed. Finally, the initial treatment regimen was a carefully considered decision by the clinicians, and we used IPTW and PSM to control for baseline differences between patients to minimize the influence of the subjective judgment of clinicians and patients on the treatment regimen. But the possibility of selection bias cannot be completely eliminated.

Conclusion

In conclusion, we predicted the counterfactual outcomes in HCC patients treated by hepatectomy and TACE based on the counterfactual principle and predict the overall survival of HCC patients. This study might provide a new direction for individualized patient management and the selection of potentially appropriate treatments for patient.

Data Sharing Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

This study was approved by the Ethics Review Committee of the Affiliated Hospital of North Sichuan Medical College (Number: 2022ER059-2), and the informed consent of the participants was waived for this retrospective study.

Consent for Publication

The informed consent of the participants was waived for this retrospective study by the Ethics Review Committee of the Affiliated Hospital of North Sichuan Medical College.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study has received funding by the Natural Science Foundation of Science and Technology Department of Sichuan Province (2023NSFSC0646).

Disclosure

The authors declare that they have no competing interests.

References

1. Xie DY, Ren ZG, Zhou J, Fan J, Gao Q. Chinese clinical guidelines for the management of hepatocellular carcinoma: updates and insights. *Hepatobiliary Surg Nutr*. 2020;9(4):452–463. doi:10.21037/hbsn-20-480
2. Reig M, Forner A, Rimola J, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. *J Hepatol*. 2022;76(3):681–693. doi:10.1016/j.jhep.2021.11.018
3. Bolondi L, Burroughs A, Dufour JF, et al. Heterogeneity of patients with intermediate (BCLC B) Hepatocellular Carcinoma: proposal for a subclassification to facilitate treatment decisions. *Semin Liver Dis*. 2012;32(4):348–359. doi:10.1055/s-0032-1329906
4. Golfieri R, Bargellini I, Spreafico C, Trevisani F. Patients with Barcelona Clinic Liver Cancer Stages B and C Hepatocellular Carcinoma: time for a Subclassification. *Liver Cancer*. 2019;8(2):78–91. doi:10.1159/000489791
5. Matsumoto MM, Mouli S, Saxena P, et al. Comparing Real World, Personalized, Multidisciplinary Tumor Board Recommendations with BCLC Algorithm: 321-Patient Analysis. *Cardiovasc Intervent Radiol*. 2021;44(7):1070–1080. doi:10.1007/s00270-021-02810-8
6. Zhao YN, Zhang YQ, Ye JZ, et al. Hepatic resection versus transarterial chemoembolization for patients with Barcelona Clinic Liver Cancer intermediate stage Child-Pugh A hepatocellular carcinoma. *Exp Ther Med*. 2016;12(6):3813–3819. doi:10.3892/etm.2016.3810
7. Hsu KF, Chu CH, Chan DC, et al. Superselective transarterial chemoembolization vs hepatic resection for resectable early-stage hepatocellular carcinoma in patients with Child-Pugh class a liver function. *Eur J Radiol*. 2012;81(3):466–471. doi:10.1016/j.ejrad.2010.12.058
8. Golfieri R, Cappelli A, Cucchetti A, et al. Efficacy of selective transarterial chemoembolization in inducing tumor necrosis in small (<5 cm) hepatocellular carcinomas. *Hepatology*. 2011;53(5):1580–1589. doi:10.1002/hep.24246
9. Bargellini I, Sacco R, Bozzi E, et al. Transarterial chemoembolization in very early and early-stage hepatocellular carcinoma patients excluded from curative treatment: a prospective cohort study. *Eur J Radiol*. 2012;81(6):1173–1178. doi:10.1016/j.ejrad.2011.03.046
10. Hsu CY, Huang YH, Chiou YY, et al. Comparison of radiofrequency ablation and transarterial chemoembolization for hepatocellular carcinoma within the Milan criteria: a propensity score analysis. *Liver Transpl*. 2011;17(5):556–566. doi:10.1002/lt.22273
11. Liu PH, Su CW, Hsu CY, et al. Solitary Large Hepatocellular Carcinoma: staging and Treatment Strategy. *PLoS One*. 2016;11(5):e0155588. doi:10.1371/journal.pone.0155588
12. Lee YB, Lee DH, Cho Y, et al. Comparison of transarterial chemoembolization and hepatic resection for large solitary hepatocellular carcinoma: a propensity score analysis. *J Vasc Interv Radiol*. 2015;26(5):651–659. doi:10.1016/j.jvir.2015.02.004
13. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol Apr*. 2012;56(4):908–943.
14. Yang B, Zheng B, Yang M, et al. Liver resection versus transarterial chemoembolization for the initial treatment of Barcelona Clinic Liver Cancer stage B hepatocellular carcinoma. *Hepatol Int*. 2018;12(5):417–428. doi:10.1007/s12072-018-9888-4
15. Yin L, Li H, Li AJ, et al. Partial hepatectomy vs. transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan Criteria: a RCT. *J Hepatol*. 2014;61(1):82–88. doi:10.1016/j.jhep.2014.03.012
16. Tsilimigras DI, Bagante F, Moris D, et al. Recurrence Patterns and Outcomes after Resection of Hepatocellular Carcinoma within and beyond the Barcelona Clinic Liver Cancer Criteria. *Ann Surg Oncol*. 2020;27(7):2321–2331. doi:10.1245/s10434-020-08452-3
17. Labgaa I, Taffé P, Martin D, et al. Comparison of Partial Hepatectomy and Transarterial Chemoembolization in Intermediate-Stage Hepatocellular Carcinoma: a Systematic Review and Meta-Analysis. *Liver Cancer*. 2020;9(2):138–147. doi:10.1159/000505093
18. Kawaguchi Y, Hasegawa K, Hagiwara Y, et al. Effect of Diameter and Number of Hepatocellular Carcinomas on Survival After Resection, Transarterial Chemoembolization, and Ablation. *Am J Gastroenterol*. 2021;116(8):1698–1708. doi:10.14309/ajg.0000000000001256
19. Jiang JF, Lao YC, Yuan BH, et al. Treatment of hepatocellular carcinoma with portal vein tumor thrombus: advances and challenges. *Oncotarget*. 2017;8(20):33911–33921. doi:10.18632/oncotarget.15411
20. Cheng S, Chen M, Cai J, et al. Chinese Expert Consensus on Multidisciplinary Diagnosis and Treatment of Hepatocellular Carcinoma with Portal Vein Tumor Thrombus (2018 Edition). *Liver Cancer*. 2020;9(1):28–40. doi:10.1159/000503685
21. Niu ZJ, Ma YL, Kang P, et al. Transarterial chemoembolization compared with conservative treatment for advanced hepatocellular carcinoma with portal vein tumor thrombus: using a new classification. *Med Oncol*. 2012;29(4):2992–2997. doi:10.1007/s12032-011-0145-0
22. Luo J, Guo RP, Lai EC, et al. Transarterial chemoembolization for unresectable hepatocellular carcinoma with portal vein tumor thrombosis: a prospective comparative study. *Ann Surg Oncol*. 2011;18(2):413–420. doi:10.1245/s10434-010-1321-8
23. Xue TC, Xie XY, Zhang L, Yin X, Zhang BH, Ren ZG. Transarterial chemoembolization for hepatocellular carcinoma with portal vein tumor thrombus: a meta-analysis. *BMC Gastroenterol*. 2013;13:60. doi:10.1186/1471-230X-13-60
24. Llovet JM, Bruix J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: chemoembolization improves survival. *Hepatology*. 2003;37(2):429–442. doi:10.1053/jhep.2003.50047
25. Hernán MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health*. 2004;58(4):265–271. doi:10.1136/jech.2002.006361
26. Cucchetti A, Serenari M. Resection or ablation for very early hepatocellular carcinoma and the fundamental problem of causal inference. *Hepatobiliary Surg Nutr*. 2017;6(4):272–273. doi:10.21037/hbsn.2017.05.03
27. Foster JC, Taylor JM, Ruberg SJ. Subgroup identification from randomized clinical trial data. *Stat Med*. 2011;30(24):2867–2880. doi:10.1002/sim.4322
28. Vitale A, Burra P, Frigo AC, et al. Survival benefit of liver resection for patients with hepatocellular carcinoma across different Barcelona Clinic Liver Cancer stages: a multicentre study. *J Hepatol*. 2015;62(3):617–624. doi:10.1016/j.jhep.2014.10.037
29. Kokudo T, Hasegawa K, Matsuyama Y, et al. Liver resection for hepatocellular carcinoma associated with hepatic vein invasion: a Japanese nationwide survey. *Hepatology*. 2017;66(2):510–517. doi:10.1002/hep.29225
30. Zhu K, Chen J, Lai L, et al. Hepatocellular carcinoma with portal vein tumor thrombus: treatment with transarterial chemoembolization combined with sorafenib--a retrospective controlled study. *Radiology*. 2014;272(1):284–293. doi:10.1148/radiol.14131946
31. Kim S, Shin J, Kim DY, Choi GH, Kim MJ, Choi JY. Radiomics on Gadoxetic Acid-Enhanced Magnetic Resonance Imaging for Prediction of Postoperative Early and Late Recurrence of Single Hepatocellular Carcinoma. *Clin Cancer Res*. 2019;25(13):3847–3855. doi:10.1158/1078-0432.CCR-18-2861
32. Chen M, Cao J, Hu J, et al. Clinical-Radiomic Analysis for Pretreatment Prediction of Objective Response to First Transarterial Chemoembolization in Hepatocellular Carcinoma. *Liver Cancer*. 2021;10(1):38–51. doi:10.1159/000512028
33. Cucchetti A, Mazzaferro V, Pinna AD, et al. Average treatment effect of hepatic resection versus locoregional therapies for hepatocellular carcinoma. *Br J Surg*. 2017;104(12):1704–1712. doi:10.1002/bjs.10613