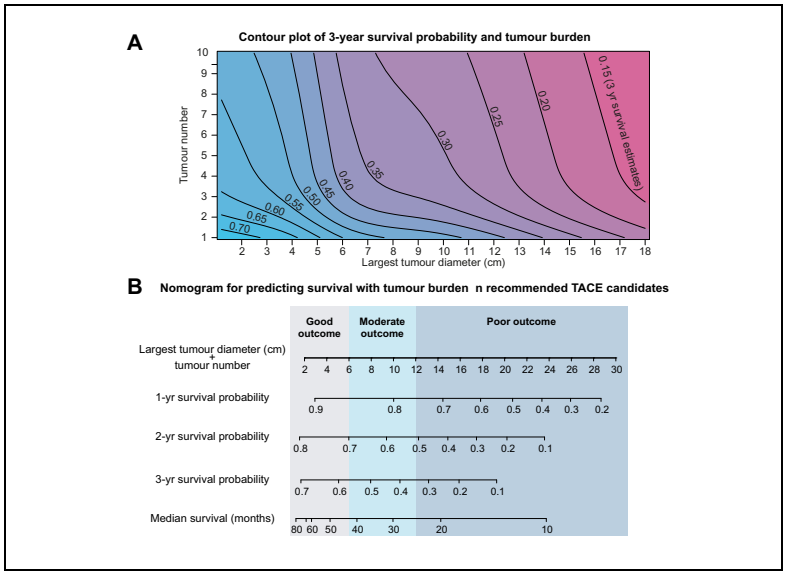
Research Article
Cancer

JOURNAL
OF HEPATOLOGY

Development of a prognostic score for recommended TACE candidates with hepatocellular carcinoma: A multicentre observational study

Graphical abstract



Highlights

- First prognostic model specifically developed for ideal TACE candidates.
- The individualized prediction score is presented as the sum of tumour size (cm) and number.
- With cut-offs of 6 and 12, the score can stratify these patients into 3 prognostic strata.
- The score outperformed other available models in performance and discrimination.
- The score is advantageous for easy-to-use and individualized prediction.

Authors

Qiuhe Wang, Dongdong Xia, Wei Bai, ..., Jielai Xia, Daiming Fan, Guohong Han

Correspondence

hangh@fmmu.edu.cn
(G. Han)

Lay summary

There is currently no prognostic model specifically developed for recommended or ideal transarterial chemoembolization (TACE) candidates with hepatocellular carcinoma, despite these patients being frequently identified as the best target population in pivotal randomized controlled trials. The six-and-twelve score provides patient survival prediction, especially in ideal candidates of TACE, outperforming other currently available models in both training and validation sets, as well as different subgroups. With cut-off values of 6 and 12, the score can stratify ideal TACE candidates into 3 strata with significantly different outcomes and may shed light on risk stratification of these patients in clinical practice as well as in clinical trials.

http://dx.doi.org/10.1016/j.jhep.2019.01.013
 J. Hepatol. 2019, 70, 893–903

Research Article
Cancer

JOURNAL OF HEPATOLOGY

Development of a prognostic score for recommended TACE candidates with hepatocellular carcinoma: A multicentre observational study

Qiuhe Wang[1,†], Dongdong Xia[1,†], Wei Bai[1,†], Enxin Wang[1,†], Junhui Sun[2,†], Ming Huang[3,†], Wei Mu[4], Guowen Yin[5], Hailiang Li[6], Hui Zhao[7], Jing Li[8], Chunqing Zhang[9], Xiaoli Zhu[10], Jianbing Wu[11], Jiaping Li[12], Weidong Gong[13], Zixiang Li[14], Zhengyu Lin[15], Xingnan Pan[16], Haibin Shi[17], Guoliang Shao[18], Jueshi Liu[19], Shufa Yang[20], Yanbo Zheng[21], Jian Xu[22], Jinlong Song[23], Wenhui Wang[24], Zhexuan Wang[1], Yuelin Zhang[2], Rong Ding[3], Hui Zhang[4], Hui Yu[5], Lin Zheng[6], Weiwei Gu[7], Nan You[8], Guangchuan Wang[9], Shuai Zhang[10], Long Feng[11], Lin Liu[12], Peng Zhang[13], Xueda Li[14], Jian Chen[15], Tao Xu[16], Weizhong Zhou[17], Hui Zeng[18], Yongjin Zhang[19], Wukui Huang[20], Wenjin Jiang[21], Wen Zhang[22], Wenbo Shao[23], Lei Li[24], Jing Niu[1], Jie Yuan[1], Xiaomei Li[1], Yong Lv[1], Kai Li[1], Zhanxin Yin[1], Jielai Xia[25], Daiming Fan[26], Guohong Han[1,*], on behalf of China HCC-TACE Study Group

[1]*Department of Liver Disease and Digestive Interventional Radiology, National Clinical Research Centre for Digestive Disease and Xijing Hospital of Digestive Diseases, Fourth Military Medical University, Xi'an, China;* [2]*Department of Hepatobiliary and Pancreatic Interventional Cancer, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China;* [3]*Department of Minimally Invasive International Therapy, The Third Affiliated Hospital of Kunming University, Tumour Hospital of Yunnan Province, Kunming, China;* [4]*Department of Radiology, The Southwest Hospital, Third Military Medical University, Chongqing, China;* [5]*Department of Interventional Radiology, Jiangsu Provincial Cancer Hospital, The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China;* [6]*Department of Interventional Radiology, Henan Cancer Hospital, The Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, China;* [7]*Department of Interventional Radiology, The Affiliated Hospital of Nantong University, Nantong, China;* [8]*Department of Hepatobiliary Surgery, Xinqiao Hospital, Third Military Medical University, Chongqing, China;* [9]*Department of Gastroenterology and Hepatology, Shandong Province Hospital Affiliated to Shandong University, Jinan, China;* [10]*Department of Interventional Radiology, The First Affiliated Hospital of Soochow University, Suzhou, China;* [11]*Department of Oncology, The Second Affiliated Hospital of Nanchang University, Nanchang, China;* [12]*Department of Interventional Radiology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China;* [13]*Department of Interventional Radiology, Tangdu Hospital, Fourth Military Medical University, Xi'an, China;* [14]*Interventional Medical Centre of the Affiliated Hospital of Qingdao University, Qingdao, China;* [15]*Department of Interventional Radiology, First Affiliated Hospital of Fujian Medical University, Fuzhou, China;* [16]*Clinical Liver Diseases Research Centre, Nanjing Military Command, 180th Hospital of PLA, Quanzhou, China;* [17]*Department of Interventional Radiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China;* [18]*Department of Radiology, Zhejiang Cancer Hospital, Hangzhou, China;* [19]*Department of Interventional Radiology and Vascular Surgery, Hunan Provincial People's Hospital, Changsha, China;* [20]*Department of Interventional Radiology, The Affiliated Tumour Hospital of Xinjiang Medical University, Urumqi, China;* [21]*Department of Interventional Radiology, Yantai Yuhuangding Hospital, Yantai, China;* [22]*Department of Medical Imaging, Nanjing General Hospital of the Nanjing Military Command, Nanjing, China;* [23]*Department of Interventional Therapy, Shandong Tumour Hospital, Jinan, China;* [24]*Department of Interventional Medicine, The First Affiliated Hospital of Lanzhou University, Lanzhou, China;* [25]*Department of Health Statistics, Fourth Military Medical University, Xi'an, China;* [26]*State Key Laboratory of Cancer Biology, National Clinical Research Centre for Digestive Disease and Xijing Hospital of Digestive Diseases, Fourth Military Medical University, Xi'an, China*

Background & Aims: Previous prognostic scores for transarterial chemoembolization (TACE) were mainly derived from real-world settings, which are beyond guideline recommendations. A robust model for outcome prediction and risk stratification of recommended TACE candidates is lacking. We aimed to develop an easy-to-use tool specifically for these patients.

Methods: Between January 2010 and May 2016, 1,604 treatment-naïve patients with unresectable hepatocellular carcinoma (HCC), Child-Pugh A5-B7 and performance status 0 undergoing TACE were included from 24 tertiary centres. Patients were randomly divided into training (n = 807) and validation (n = 797) cohorts. A prognostic model was developed and subsequently validated. Predictive performance and discrimination were further evaluated and compared with other prognostic models.

Results: The final presentation of the model was "linear predictor = largest tumour diameter (cm) + tumour number", which consistently outperformed other currently available models in both training and validation datasets as well as in different subgroups. The thirtieth percentile and the third quartile of the linear predictor, namely 6 and 12, were further selected as cut-off values, leading to the "six-and-twelve" score which could divide

Keywords: Transarterial chemoembolization; Hepatocellular carcinoma; Risk stratification; Individual prediction.
Received 16 June 2018; received in revised form 27 December 2018; accepted 4 January 2019; available online 18 January 2019

* Corresponding author. Address: Department of Liver Disease and Digestive Interventional Radiology, National Clinical Research Centre for Digestive Disease and Xijing Hospital of Digestive Diseases, Fourth Military Medical University, 15 West Changle Road, Xi'an, China. Tel.: +86 29 84771528, fax: +86 29 82539041.
E-mail address: hangh@fmmu.edu.cn (G. Han).
† These authors share the co-first authorship.

patients into 3 strata with the sum of tumour size and number ≤6, >6 but ≤12, and >12 presenting significantly different median survival of 49.1 (95% CI 43.7–59.4) months, 32.0 (95% CI 29.9–37.5) months, and 15.8 (95% CI 14.1–17.7) months, respectively.

Conclusions: The six-and-twelve score may prove an easy-to-use tool to stratify recommended TACE candidates (Barcelona Clinic Liver Cancer stage-A/B) and predict individual survival with favourable performance and discrimination. Moreover, the score could stratify these patients in clinical practice as well as help design clinical trials with comparable criteria involving these patients. Further external validation of the score is required.

Lay summary: There is currently no prognostic model specifically developed for recommended or ideal transarterial chemoembolization (TACE) candidates with hepatocellular carcinoma, despite these patients being frequently identified as the best target population in pivotal randomized controlled trials. The six-and-twelve score provides patient survival prediction, especially in ideal candidates of TACE, outperforming other currently available models in both training and validation sets, as well as different subgroups. With cut-off values of 6 and 12, the score can stratify ideal TACE candidates into 3 strata with significantly different outcomes and may shed light on risk stratification of these patients in clinical practice as well as in clinical trials.

Introduction

According to the guidelines of the American Association for the Study of Liver Disease (AASLD) and the European Association for the Study of Liver (EASL), transarterial chemoembolization (TACE) is currently the only recommended treatment option for patients with intermediate stage hepatocellular carcinoma (HCC),[1–3] with well-preserved liver function and performance status.[4–6] These patients, as well as those at early stage but considered unresectable due to tumour size, location, patient age, and suggestions of stage migration are considered "recommended" or "ideal" TACE candidates, *i.e.*, the best target population for TACE, and have frequently been set as the target population in pivotal studies (Table S1A-1B).[7–12] However, this population is rather heterogeneous with a variable median overall survival of 13–43 months,[13] rendering it crucial to develop a risk stratification tool.[14] Indeed, the necessity of stratifying risk in these patients has also been underlined by the most recent guidelines.[14] More importantly, a pre-procedure prognostic model providing survival estimates after TACE as a reference may enable outcome comparisons with other treatments and thus could aid clinical decision making.[15]

Nevertheless, existing models such as hepatoma arterial-embolization prognostic (HAP) score are mostly derived from patients receiving TACE in a broader setting beyond guideline recommendations rather than the best target population for TACE (Table S1C)[16] and although their performance and discrimination have been validated in other settings, it remains unknown whether a consistent result could be observed, especially in these recommended TACE candidates. A model established especially for this population is therefore urgently needed and validation of previous models is crucial.

Fundamental and informative prognostic indicators for developing a model specific to patients with HCC include tumour burden profiles, liver function, performance status, and biomarkers such as alpha-fetoprotein (AFP). Among these parameters, tumour burden profiles seemed particularly important, not only due to their well-perceived negative correlation with survival[17,18] and response rates,[19,20] but also because of the disparate tumour load varying from a solitary tumour to multiple tumours with different diameters in these ideal TACE candidates,[21–23] which might contribute significantly to the heterogeneity of the population. Meanwhile, liver function and performance status have also been established as significant predictors in prognostic models for patients with HCC.[24]

On the other hand, individualized prediction has been perceived as another requisite for an optimal prognostic model. Well-recognized categorized scores and models such as "four-and-seven" criteria, HAP score, and Barcelona Clinical Liver Cancer (BCLC) intermediate stage sub-classification can be useful for stratifying the level of risk but may not be sufficiently informative for individual outcome evaluation, calling for an individualized model displaying continuous data.

Therefore, the aim of this study was to develop a specific model that presents continuous data and can be used to predict individual survival and stratify patient risk in ideal candidates of TACE.

Patients and methods

Study population

Between January 2010 and May 2016, a total of 3,819 consecutive patients with unresectable HCC receiving conventional TACE from 24 tertiary Chinese centres were retrospectively screened. HCC was diagnosed by either imaging or histological evaluation according to the AASLD or EASL guidelines. The target population is "recommended" or "ideal" candidates of TACE, defined by unresectable BCLC-A patients identified by comprehensive assessment and BCLC-B patients. On the basis of preserved liver function, good performance status, and no vascular invasion or extrahepatic spread, the former, according to the criteria of BCLC staging systems, mainly includes single nodule >2 cm or up to 3 nodules ≤3 cm whereas the latter refers to multinodular HCC beyond BCLC-A.[25] Notably, although TACE procedure has been performed in patients with performance status score 1 and impaired liver function, these patients belong to BCLC C stage where TACE is not recommended as first-line therapy, and thus were not considered "ideal candidates". Specifically, the inclusion criteria included: i) treatment-naïve patients with unresectable HCC receiving TACE; ii) Child-Pugh score A5-B7; and iii) at least 1 measurable lesion >1 cm. Patients were excluded based on the following criteria: i) vascular invasion or extrahepatic spread; ii) spontaneous tumour rupture; iii) comorbidity with other malignancies; iv) decompensated liver cirrhosis (gastrointestinal bleeding, ascites, jaundice, or encephalopathy); v) performance status score >0; vi) treatment with any systemic or loco-regional therapies; and vii) absence of baseline imaging information. Patients within each centre were randomly assigned to training or validation datasets by computer-generated randomized number. Diameter of the largest nodule (tumour size, hereinafter measured in centimetres) and tumour number were evaluated by 2 independent investigators (Bai W. and Xia D.) using either multiphasic computer

tomography (CT) or dynamic contrast-enhanced magnetic resonance imaging (MRI).

Treatment procedures

During TACE procedure, an emulsion of mixtures of lipiodol (3–30 ml) and chemotherapeutic drugs was injected. Doxorubicin (10–50 mg), cisplatin (10–110 mg), epirubicin (10–50 mg), or oxaliplatin (100–200 mg) were selected according to the practice of each centre. Afterwards, either gelatine sponge or polyvinyl alcohol foam (PVA) particles were introduced, and the embolization was monitored until the tumour arterial flow was reduced as observed on angiography. Tumour-feeding vessels were embolized either selectively or super-selectively,[26] and "on demand" TACE procedures are scheduled at an interval of 6 to 12 weeks upon the demonstration of viable tumours or intrahepatic recurrences by CT/MRI in patients with favourable clinical and laboratory findings (performance status, liver function, *etc.*), as well as the absence of extrahepatic spread or vascular invasion. All procedures were performed by investigators with at least 8 years of experience.

Statistical analysis

Quantitative variables were presented as median with interquartile range (IQR) and compared by Student's *t* test or non-parametric Mann-Whitney *U* test, whereas categorical variables were presented as counts with percentages and compared by Chi-squared test or Fisher's exact test. Overall Survival (OS) was defined as the time interval between initial TACE and all-cause death. Patients who survived at last follow-up date (December 15th, 2017) or lost to follow-up were censored. Survival curves were estimated using the Kaplan-Meier method and compared by log-rank test. Univariate and multivariate Cox regression analyses were applied to the training cohort to identify prognostic factors after multiple imputation had been performed with 5 independent draws for missing values.

Continuous variables identified as predictors of survival were analysed with 5-knot restricted cubic splines to evaluate their correlations with the relative hazard of death and to test non-linearity. A contour plot was used to depict survival estimates according to tumour burden. The interactions between predictors were also tested. Possible models for risk stratification were then developed based on the above analyses, and discrimination, performance, and calibration were measured by C-index, likelihood ratio Chi-square, area under time-dependent receiving operator characteristic curve, and calibration curves, respectively. Akaike information criterion (AIC) was also calculated to compare the loss of information for different models. Bootstrapping with 1,000 samples was used for model validation in subgroups with sample sizes less than 500. The final model was compared with prognostic models including up-to-seven criteria,[27] four-and-seven criteria,[28] HAP score,[16] modified HAP-II (mHAP-II) score,[29] modified HAP-III (mHAP-III) score,[15] intermediate HCC (BCLC-B) sub-classification,[21] and albumin-bilirubin (ALBI) score,[30] and with prognostic systems including BCLC staging system,[25] Hong Kong Liver Cancer Staging System (HKLC),[31] TNM staging system by Liver Cancer Study Group of Japan (TNM-LCSGJ),[32] the 8th edition of American Joint Committee on Cancer TNM staging (AJCC TNM-8),[33] Japan Integrated Staging (JIS),[32] Tokyo score,[34] and the most recent Italian Liver Cancer (ITA.LI.CA) prognostic system[35] in both training and validation sets, as well as in different subgroups.

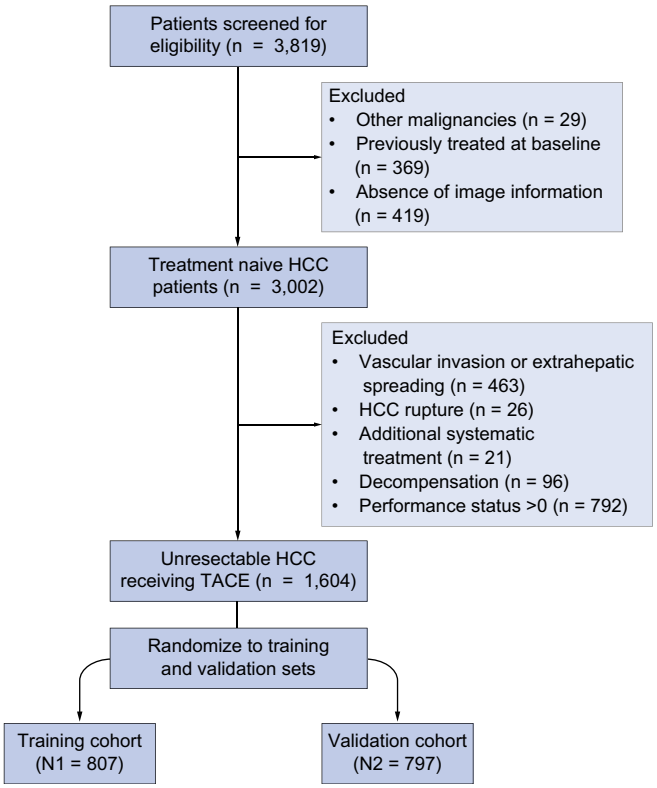


Fig. 1. Flow chart of study design. HCC, hepatocellular carcinoma; TACE, transarterial chemoembolization.

It was difficult to calculate sample size beforehand due to scarce evidence in developing a risk stratification model for recommended TACE candidates. However, the total number of events (all-cause death) reached 811, and a ratio of 10 events per variable[36] was exceeded given that the expected number of candidate variables was no more than 20, indicating sufficient accuracy of estimation.

Differences were considered statistically significant when corresponding *p* values were less than 0.05. All statistical analyses were performed using R version 3.3.2 with packages mice,[37] rms,[38] timeROC,[39] and boot.[40] All authors had access to the study data and reviewed and approved the final manuscript.

Ethics

The study protocol was approved by each participating centre's ethics committee. Permission to use the corresponding data was obtained from all patients by written consent.

For further details regarding the materials used, please refer to the CTAT table and supplementary information.

Results

Baseline characteristics

A total of 1,604 patients were finally included and randomly divided into the training (N1 = 807) and validation (N2 = 797) datasets (Fig. 1). Baseline characteristics were comparable between the 2 datasets (Table 1). Hepatitis B virus (HBV) was the main aetiology of HCC (85.2%). Among all the patients, 1,493 had detectable HBV-DNA or hepatitis C virus (HCV)-RNA

Table 1. Baseline demographics and clinical characteristics in 1,604 patients.

Baseline characteristics	Number (%)/Median (IQR)*			p value
	Entire cohort (N = 1,604)	Training cohort (N1 = 807)	Validation cohort (N2 = 797)	
Gender				0.695
Male	1,390 (86.7)	702 (87.0)	688 (86.3)	
Female	214 (13.3)	105 (13.0)	109 (13.7)	
Age (yr)	57 (48–65)	57 (48–65)	58 (48–65)	0.738
Aetiology				0.700
HBV	1,366 (85.2)	690 (85.5)	676 (84.8)	
Others	238 (14.8)	117 (14.5)	121 (15.2)	
Largest tumour diameter, cm	6.1 (3.8–9.8)	6.1 (3.6–10.0)	6.0 (4.0–9.55)	0.783
≤3	262 (16.3)	144 (17.8)	118 (14.8)	
>3–≤7	674 (42.0)	318 (39.4)	356 (44.7)	
>7–≤10	302 (18.8)	153 (19.0)	149 (18.7)	
>10	366 (22.8)	192 (23.8)	174 (21.8)	
Tumour number	1 (1–2)	1 (1–2)	1 (1–2)	0.427
1	919 (57.3)	457 (56.6)	462 (58.0)	
2	346 (21.6)	174 (21.6)	172 (21.6)	
3	158 (9.9)	86 (10.6)	72 (9.0)	
>3	181 (11.3)	90 (11.2)	91 (11.4)	
AFP, ng/ml				0.408
<400	970 (60.5)	478 (59.2)	492 (61.7)	
≥400	591 (36.8)	304 (37.7)	287 (36.0)	
Current BCLC staging**				0.539
A	982 (61.2)	488 (60.4)	494 (62.0)	
B	622 (38.8)	319 (39.6)	303 (38.0)	
Previous BCLC staging**				0.688
A	408 (25.4)	209 (25.9)	199 (25.0)	
B	1,196 (74.6)	598 (74.1)	598 (75.0)	
Child-Pugh score				1.000
A5	1,239 (77.2)	623 (77.2)	616 (77.3)	
A6	289 (18.0)	149 (18.5)	140 (17.6)	
B7	76 (4.7)	35 (4.2)	41 (5.1)	
ALBI grade				1.000
1	799 (49.8)	409 (50.7)	390 (48.9)	
2	782 (48.8)	390 (48.3)	392 (49.2)	
3	22 (1.4)	7 (0.9)	15 (1.9)	
WBC, 10^9/L	5.30 (4.03–6.61)	5.28 (4.00–6.69)	5.30 (4.10–6.57)	0.369
RBC, 10^{12}/L	4.41 (4.00–4.83)	4.44 (4.08–4.86)	4.40 (3.93–4.81)	0.988
PLT, 10^9/L	134 (87–186)	136 (85–188)	133 (89–183.5)	0.148
INR	1.06 (1.00–1.13)	1.06 (1.00–1.14)	1.06 (1.00–1.15)	0.101
ALT, IU/L	39.00 (26.00–60.00)	38.00 (27.00–59.00)	40.00 (26.00–61.50)	0.177
AST, IU/L	44.00 (31.00–65.00)	44.00 (31.00–66.00)	44.00 (31.00–64.00)	0.162
ALB, g/L	39.70 (36.00–43.40)	40.00 (36.10–43.30)	39.40 (35.80–43.45)	0.422
TBIL, μmol/L	15.60 (11.40–21.70)	15.60 (11.40–22.00)	15.50 (11.35–21.35)	0.814
BUN, mg/dl	5.40 (4.40–6.40)	5.25 (4.30–6.32)	5.50 (4.51–6.50)	0.102
Creatinine, μmol/L	71 (61–82)	71 (61–83)	71 (62–82)	0.528
TACE sessions	3 (2–4)	3 (2–4)	3 (2–4)	0.057
Radiological assessment				0.756
Multiphasic CT	1,281 (79.9%)	642 (79.6%)	639 (80.2%)	
Dynamic contrast–enhanced MRI	323 (20.1%)	165 (20.4%)	158 (19.8%)	
Follow-up duration	22.0 (11.9–34.1)	22.3 (12.3–34.2)	21.5 (11.2–33.5)	0.091

AFP, alpha-fetoprotein; ALB, albumin; ABLI, albumin-bilirubin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancer; BUN, blood urea nitrogen; CT, computed tomography; HBV, hepatitis B virus; INR, international normalized ratio; IQR, interquartile range; MRI, magnetic resonance imaging; PLT, platelet; PS, performance status; RBC, red blood cell; TACE, transarterial chemoembolization; TBIL, total bilirubin; WBC, white blood cell.

* Median with interquartile range are shown for quantitative variables, whereas counts with proportions are shown for categorical variables.

** With PS0, reserved liver function, and the absence of metastasis or macrovascular invasion, current BCLC staging considers single tumour >2 cm as stage A, whereas previous BCLC stage considers single tumour ≤5 cm or no more than 3 tumours with maximum diameter ≤3 cm as stage A. Staging of patients under these 2 criteria is listed in the table.

and received antiviral treatment, and virologic response (as defined by recent guidelines[41,42]) was achieved in 1,374 (92.0%) patients during the study period (681 in training set and 693 in validations set).

Overall survival

The median follow-up was 22.0 months (IQR 11.9–34.1 months) and 22.3 months (IQR 12.3–34.2 months) in training and validation sets, respectively. A total of 183 (11.4%) patients were lost to follow-up (84 in the training cohort; 99 in the validation cohort). The median survival of the entire cohort was 32.9 (95% CI 30.4–35.4) months, with 1-year, 2-year, and 3-year survival being 81.2%, 60.9%, and 46.5%, respectively. No significant difference was observed between training and validation sets (median OS 32.6 [95% CI 28.6–37.0] months *vs.* 32.9 [95% CI 30.1–36.7] months, $p = 0.537$), (Fig. 2A-B).

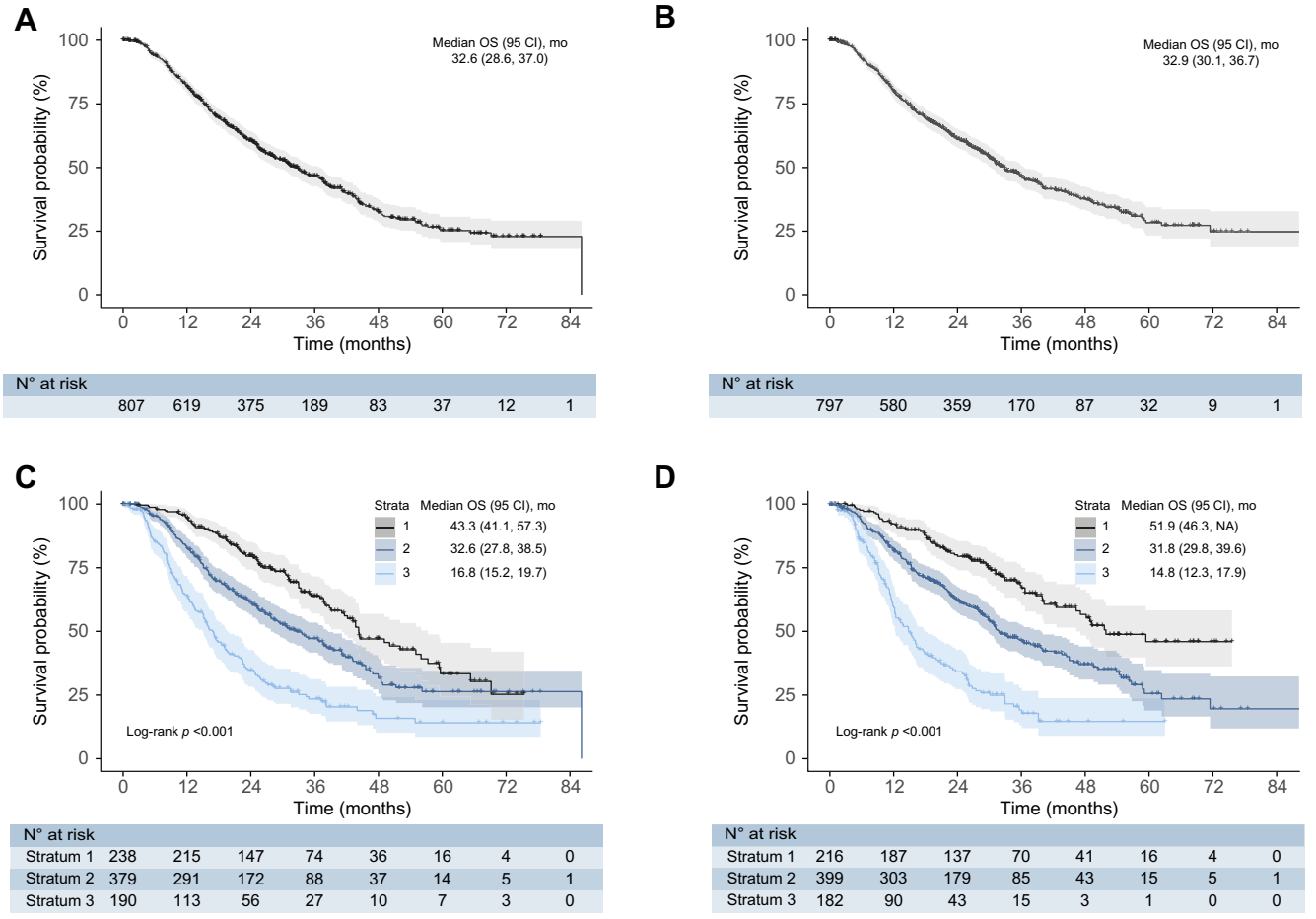


Fig. 2. Treatment outcome according to the current model. (A) Unstratified Kaplan-Meier curve of training set; (B) Unstratified Kaplan-Meier curve of validation set; (C) Kaplan-Meier curve of stratified survival in training set; (D) Kaplan-Meier curve of stratified survival in validation set. OS, overall survival.

Univariate and multivariate analysis

Percentages and patterns of missing values are shown in Fig. S1. The multivariate Cox regression analysis after multiple imputation suggested that largest tumour diameter (tumour size), tumour number, and AFP were independent prognostic factors (Table S2) that will be considered for model development.

Development of the prognostic model

AFP (≤400 ng/ml *vs.* >400 ng/ml) was encoded as a categorical factor, whereas tumour size and number were treated as continuous variables. Restrictive cubic spline functions of tumour size and number in training and validation sets showed that both prognostic factors presented a linear profile (non-linearity *p* values were 0.08 and 0.15 for tumour number and size in training set, respectively; and 0.39 and 0.43 in validation set, respectively, Fig. 3A–D). When validated in patients with HBV, other aetiologies and in patients who underwent CT and MRI evaluation, the results were similar: non-linearity *p* values were 0.06 and 0.06 for tumour size and number in patients with HBV, respectively; 0.41 and 0.78 in patients with other aetiology, respectively (Fig. S2); 0.44 and 0.91 in MRI evaluation group, respectively; and 0.06 and 0.44 in CT evaluation group, respectively (Fig. S3). These results suggested that tumour size and number could be modelled as continuous linear variables. The relationship between tumour size, tumour number and 3-year survival probability were depicted in a contour plot (Fig. 3E). Interaction tests between variables suggested no interaction terms. Two possible models were thus developed: model 1, only tumour burden profiles were included (tumour size and number); model 2, tumour burden and AFP were included.

The coefficients of variables derived from Cox regression analyses in each model were multiplied by 10 and rounded to one decimal for clinical use, resulting in the following 2 models:

Model 1: linear predictor (LP) = tumour size + tumour number

Model 2: LP = 1.1 × tumour size + 1.1 × tumour number + 3.4 × AFP

(Where "AFP" is a binary variable with a value of 0 for AFP ≤400 ng/ml and 1 for AFP >400 ng/ml, and tumour size and number are continuous variables).

The time-dependent AUROC value, likelihood ratio (LR) χ^2, C-index and AIC of the possible models are shown in Table 2. Compared with model 1, model 2 (with the addition of AFP levels) showed only slight improvement regarding time-dependent AUROC values and C-indices, with insignificant statistical differences.

Since different aetiology indicates different outcome, and MRI can be more accurate in detecting viable HCC nodules under 2 cm and in treatment response evaluation, the entire cohort was further divided into subgroups according to aetiology (HBV and other aetiology) and image assessment modality

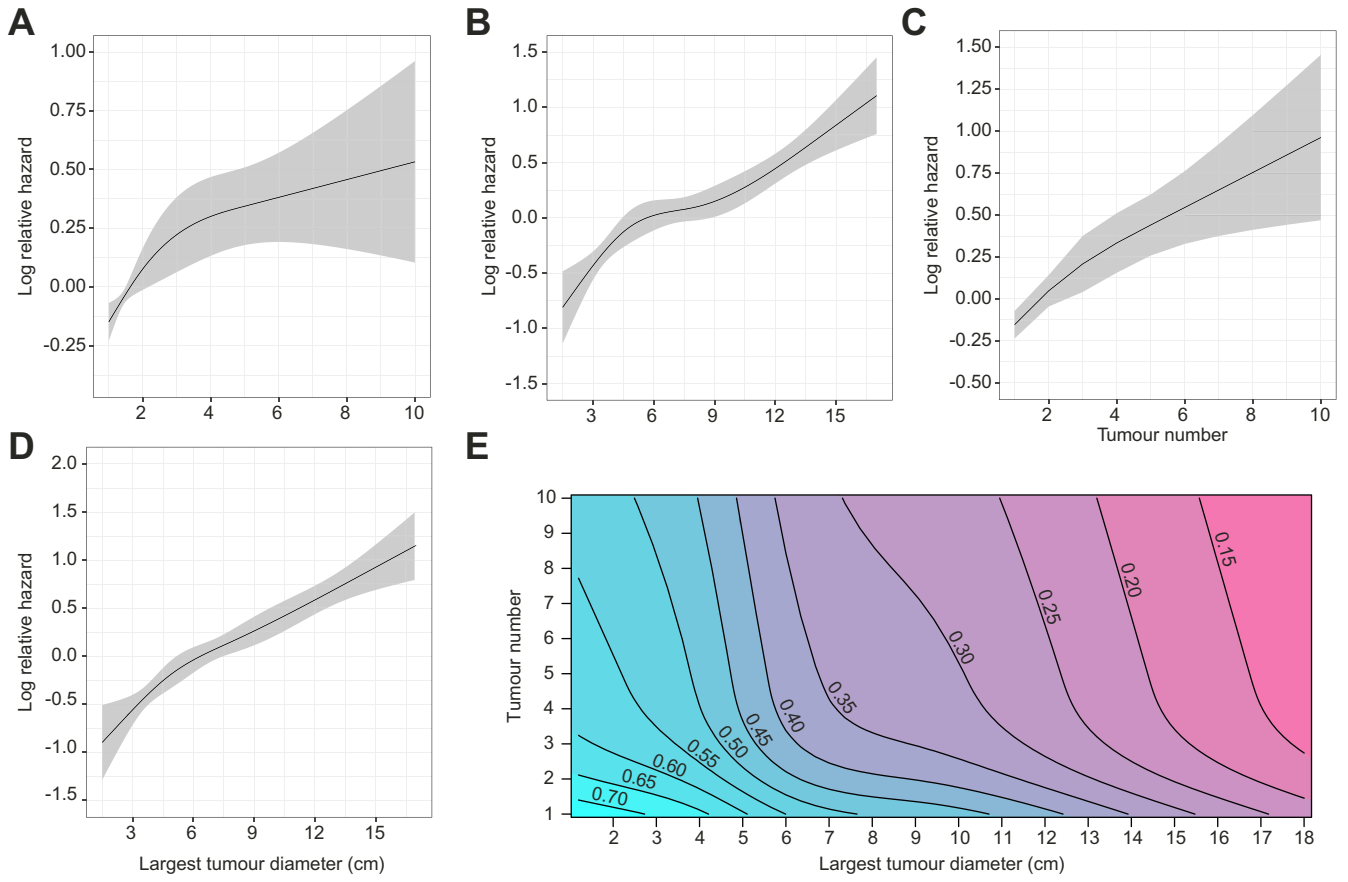


Fig. 3. Relation between largest tumour diameter, tumour number and relative hazard. (A) Restricted cubic spline of largest tumour diameter in training set; (B) Restricted cubic spline of tumour number in training set; (C) Restricted cubic spline of largest tumour diameter in validation set; (D) Restricted cubic spline of tumour number in validation set; (E) Contour plot of 3-year survival probability according to largest tumour diameter and tumour number. (This figure appears in colour on the web.)

Table 2. Comparison of the performance and discriminative ability of 2 possible models.

Cohort	Models	1-yr AUROC (95% CI)	2-yr AUROC (95% CI)	3-yr AUROC (95% CI)	LR χ^2	df	C-index (95% CI)	AIC
Training*	Model 1 LP = T_S + T_N	0.73 (0.69–0.77)	0.69 (0.65–0.73)	0.65 (0.61–0.70)	77.58	2	0.66 (0.63–0.69)	4,972.50
	Model 2 LP = 1.1 × T_S + 1.1 × T_N + 3.4 × AFP	0.75 (0.70–0.79)	0.69 (0.65–0.73)	0.65 (0.61–0.70)	87.88	3	0.66 (0.63–0.69)	4,805.24
Validation**	Model 1 LP = T_S + T_N	0.72 (0.68–0.77)	0.71 (0.67–0.75)	0.69 (0.65–0.74)	105.34	2	0.67 (0.64–0.70)	4,554.98
	Model 2 LP = 1.1 × T_S + 1.1 × T_N + 3.4 × AFP	0.73 (0.68–0.77)	0.72 (0.67–0.76)	0.70 (0.65–0.75)	110.81	3	0.68 (0.64–0.71)	4,399.97

AFP, alpha-fetoprotein; AIC, Akaike Information Criterion; AUROC, area under receiver operating characteristic curve; LP, linear predictor; LR, likelihood ratio; PS, performance status; TBIL, total bilirubin; T_S, tumour size; T_N, tumour number.
* In training cohort, *p* values for LR Chi-square test and C-index comparison between models 1 and 2 were 0.001 and 0.276, respectively.
** In validation cohort, *p* values for LR Chi-square test and C-index comparison between models 1 and 2 were 0.004 and 0.398, respectively.

(MRI and CT). In these subgroups, the difference in performance and discrimination between the 2 candidate models was also insignificant (Table S3).

Given this situation and the purpose of developing an easy-to-use bedside stratification tool, model 1 was selected as the final model. Optimism-corrected calibration slopes were 0.97 (standard error: 0.02) in the training set and 0.98 (standard error: 0.01) in the validation set (Fig. S4). Calibration curves were also performed in patients with HBV and other aetiologies (Fig. S5), as well as in patients who underwent MRI and CT assessment (Fig. S6), with consistent findings.

Based on these findings, a nomogram for individual patient risk stratification was created (Fig. 4). The 1-year, 2-year, and 3-year survival probability and estimated median survival of individual patients could be predicted before the TACE procedure with the sum of tumour size and number.

Assessment and comparison of the performance and discrimination of the current model and other models and prognostic systems

The performance and discrimination of the selected model and other tumour-burden-based models (up-to-seven criteria and

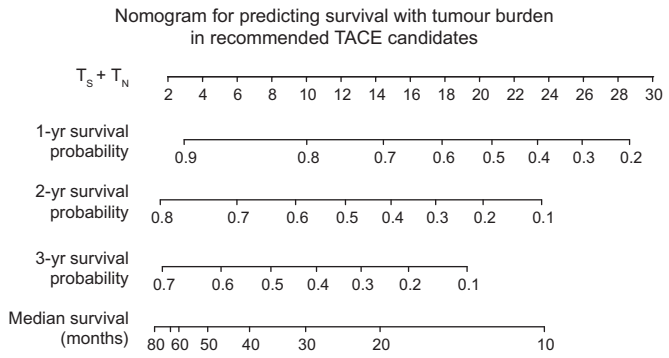


Fig. 4. Nomogram of the current model for individual survival prediction. TACE, transarterial chemoembolization; T_S: largest tumour diameter in cm; T_N: tumour number.

four-and-seven criteria), as well as models involving other parameters (HAP score, mHAP score, mHAP-II score, mHAP-III score, BCLC-B sub-classification and ALBI score) were compared (Table 3). The 1-year, 2-year, and 3-year AUROC values and C-indices of current model were higher than other models (Fig. 5), suggesting a favourable performance and discrimination, which remained consistent in different subgroups with different age, gender, AFP levels, liver function, aetiology and imaging assessment modality (Table S4–7). In the comparison with currently available prognostic systems, the current score remained favourable for risk stratification in the current target population (Table S8).

Cut-off value for risk stratification

Apart from individual level outcome prediction, quick evaluation of prognosis levels with cut-off values is also relevant to routine clinical practice. By examining the C-indices and AUROC values for different percentiles and quartiles of the selected LP, cut-off values at the thirtieth percentile and the third quartile (6.1 and 12.0) were identified for the sum of tumour size and tumour number, and were then rounded as 6 and 12, respectively. The current model was named as the six-and-twelve score accordingly. Consequently, patients were divided into 3 strata: stratum 1, the sum of tumour size and number ≤6; stratum 2, the sum >6 but ≤12; stratum 3, the sum >12.

In the entire cohort, the median OS of the 3 strata from the lowest tumour burden stratum to the highest were 49.1 (95% CI 43.7–59.4), 32.0 (95% CI 29.9–37.5), and 15.8 (95% CI 14.1–17.7) months, respectively. With stratum 1 as reference, the hazard ratios (HRs) for strata 2 and 3 were 1.76 (95% CI 1.47–2.11; p <0.001) and 3.69 (95% CI 3.03–4.50; p <0.001), respectively. The median OS of strata 1, 2, and 3 in the training cohort were 44.3 (95% CI 41.1–57.3), 32.6 (95% CI 27.8–38.5), and 16.8 (95% CI 15.1–19.7) months, respectively. With stratum 1 as reference, the HRs for strata 2 and 3 were 1.64 (95% CI 1.28–2.10; p <0.001) and 3.20 (95% CI 2.46–4.17; p <0.001), respectively. In the validation cohort, the median OS of these 3 strata were 51.9 (95% CI 46.3–NA), 31.8 (95% CI 29.8–39.6), and 14.8 (95% CI 12.3–17.9) months, respectively. With stratum 1 being the reference, the HRs for strata 2 and 3 were 1.92 (95% CI 1.46–2.51; p <0.001) and 4.37 (95% CI 3.25–5.90; p <0.001), respectively. Survival curves were significantly different among the 3 strata in training and validation sets (both with log-rank p <0.001, Fig. 2C-D). Rates of BCLC stage A and B within each stratum were listed in Table S9.

Subgroup analysis

The current model could stratify patients into the aforementioned 3 strata across subgroups including in patients with different AFP levels (≤400 ng/ml and >400 ng/ml), ALBI grades (1 and 2), and aetiologies (HBV and other aetiologies), and image modalities for evaluation (MRI and CT), suggesting a consistent performance in these populations (Fig. S7). The median survival and HRs with 95% CIs of the 3 strata in different subgroups are listed (Table S10).

Discussion

Based on a multicentre, nationwide cohort comprising 1,604 recommended TACE candidates (BCLC-A/B), we developed the six-and-twelve score that can predict individual outcome with

Table 3. Comparison of the performance and discriminative ability between the current model and other models.

Cohort	Models	1-yr AUROC (95% CI)	2-yr AUROC (95% CI)	3-yr AUROC (95% CI)	C-index (95% CI)
Training	The current model	0.73 (0.69–0.77)	0.69 (0.65–0.73)	0.65 (0.61–0.70)	0.66 (0.63–0.69)
	Up-to-seven	0.66 (0.62–0.70)	0.63 (0.60–0.67)	0.60 (0.56–0.64)	0.61 (0.58–0.63)
	Four-and-seven	0.66 (0.61–0.70)	0.63 (0.59–0.67)	0.59 (0.55–0.64)	0.61 (0.58–0.63)
	HAP	0.67 (0.63–0.72)	0.64 (0.60–0.68)	0.58 (0.54–0.63)	0.61 (0.58–0.64)
	mHAP	0.68 (0.64–0.73)	0.65 (0.61–0.69)	0.63 (0.59–0.68)	0.63 (0.60–0.66)
	mHAP II	0.67 (0.63–0.72)	0.65 (0.61–0.69)	0.58 (0.54–0.63)	0.61 (0.58–0.64)
	mHAP III	0.69 (0.64–0.74)	0.65 (0.61–0.70)	0.63 (0.57–0.68)	0.63 (0.59–0.66)
	BCLC-B sub-classification	0.61 (0.58–0.64)	0.61 (0.57–0.64)	0.60 (0.55–0.64)	0.58 (0.56–0.61)
	ALBI	0.52 (0.47–0.56)	0.53 (0.49–0.57)	0.51 (0.47–0.56)	0.53 (0.50–0.55)
Validation	The current model	0.72 (0.68–0.77)	0.71 (0.67–0.75)	0.69 (0.65–0.74)	0.67 (0.64–0.70)
	Up-to-seven	0.65 (0.62–0.69)	0.65 (0.61–0.68)	0.64 (0.59–0.68)	0.61 (0.59–0.64)
	Four-and-seven	0.67 (0.62–0.71)	0.63 (0.60–0.67)	0.63 (0.59–0.68)	0.62 (0.59–0.64)
	HAP	0.66 (0.61–0.71)	0.63 (0.59–0.67)	0.61 (0.56–0.66)	0.61 (0.58–0.64)
	mHAP	0.66 (0.61–0.71)	0.64 (0.69–0.68)	0.60 (0.55–0.65)	0.61 (0.58–0.64)
	mHAP II	0.65 (0.61–0.70)	0.64 (0.60–0.68)	0.61 (0.56–0.66)	0.61 (0.58–0.64)
	mHAP III	0.70 (0.64–0.75)	0.70 (0.66–0.75)	0.66 (0.61–0.71)	0.65 (0.62–0.69)
	BCLC-B sub-classification	0.61 (0.57–0.64)	0.61 (0.57–0.64)	0.62 (0.58–0.66)	0.59 (0.56–0.61)
	ALBI	0.51 (0.46–0.56)	0.53 (0.49–0.57)	0.56 (0.52–0.61)	0.53 (0.50–0.55)

ALBI, albumin-bilirubin, AUROC, area under receiver operating characteristic curve; BCLC, Barcelona Clinic Liver Cancer; HAP, hepatoma arterial-embolization prognostic; mHAP, modified HAP.

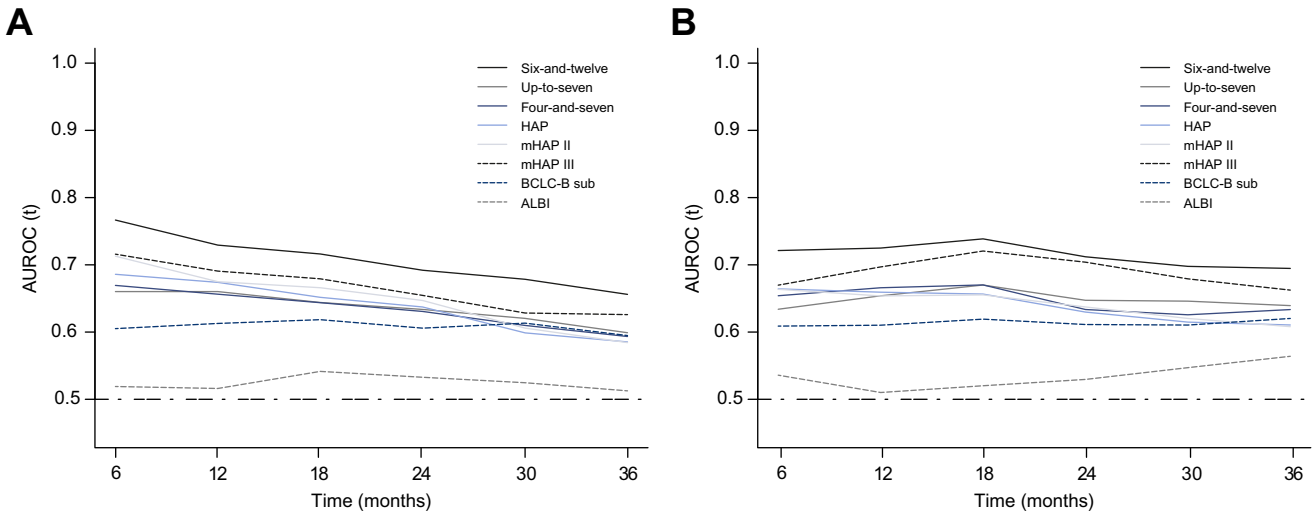


Fig. 5. Time-dependent AUROC values of the current model and other available models. (A) Time-dependent AUROC values in training set; (B) Time-dependent AUROC values in validation set. ALBI, albumin-bilirubin, AUROC, area under receiver operating characteristic curve; BCLC, Barcelona Clinic Liver Cancer; HAP, hepatoma arterial-embolization prognostic; mHAP, modified HAP.

favourable performance and discrimination, and with the sum of tumour size and number ≤6, >6 but ≤12, and >12, the score identified 3 prognostic strata presenting significantly different median survival of 49.1 months, 32.0 months, and 15.8 months, respectively. The strength and novelty of the current study lies in: i) conducting the study based on a consecutive multicentre cohort with a large sample size; ii) developing the first prognostic model specifically for ideal TACE candidates rather than all patients undergoing TACE; iii) finding that tumour burden alone can be used for outcome prediction in recommended TACE candidates; iv) adopting a continuous instead of categorized model presentation to provide individualized and stratified survival estimates for clinical practice and clinical trials.

The median OS of 32.9 months in the current cohort was obviously longer than the 19.8 month median OS reported by Lencioni *et al.* in a recent systematic review on TACE,[43] probably because the current study identified ideal TACE candidates as the target population. On the contrary, the 2018 EASL guidelines mentioned a median survival of 40 months in well-selected candidates,[14] which was much more similar to the current study. Indeed, the studies used to draw this conclusion in the guidelines[17,44,45] shared similar study population staging characteristics as the current study. In these studies, 22–41% patients were at stage A under the previous version of BCLC staging criteria according to the information provided by the authors, and in the current study, about 25% patients belong to stage A under the previous BCLC criteria (Table 1). The median survival of the 3 strata identified by the six-and-twelve score varies approximately within a range of 13–43 months, consistent with previous reports.[13,43,46] Patients within the best outcome strata had similar median survival to untreated patients with early stage HCC, whereas the stratum with the highest tumour burden had a median survival close to the 16 months reported for untreated patients with intermediate stage HCC.[47-49] This further demonstrates that this population includes patients with different prognoses even when liver function and performance status are preserved, supporting the importance of risk stratification of these patients in clinical practice and trials.

The target population is the core issue of prognostic models, since the characteristics and heterogeneity of the population determine the parameters, presentation, and performance of the final model. The best target population for TACE has been identified as the study population in pivotal randomized controlled trials and cohort studies of TACE, yet investigations regarding its internal homogeneity are lacking. This may introduce unrecognized confounding which could influence the results of these studies, thus a risk stratification score for identifying and reducing heterogeneity is needed. Particularly, as is shown in Table S8, the constituent ratio of BCLC stage A and B within strata 2 and 3 identified by current prognostic model were similar, yet patient outcomes were significantly different, suggesting that the current score can further stratify patients on the basis of the BCLC staging system. Previous scores such as four-and-seven score and HAP score targeted all patients treated by TACE, with the inclusion of cases beyond recommendation potentially influencing patient characteristics and heterogeneity profiles. The current study, however, fixed on this population for the first time, which is characterized by unresectable HCC, preserved liver function (mainly Child-Pugh A grade) and acceptable performance status due to the inclusion criteria set by these trials and studies, with no exclusion criteria on tumour size or number as in previous pivotal studies (Table S1A–1C).

In this population, tumour burden is a fundamental prognostic factor of HCC among all candidate predictors regardless of systemic or loco-regional treatment.[17,27,50,51] Intriguingly, although the predictive value of tumour burden seems obvious, the cut-off values for risk stratification have long been controversial, and most of the models featuring tumour burden adopted arbitrary or empirical cut-offs,[16,21,28,29,52–57] except for the "up-to-seven" and "four-and-seven" criteria. However, the former was originally derived from liver transplantation candidates, rendering its application in recommended TACE candidates an open issue,[27] whereas the latter criteria stemmed from a cohort with limited sample size, calling for further investigations in large cohorts.[56] In light of this situation, the six-and-twelve score was advantageous in being developed specifically

for ideal TACE candidates and adopting evidence-based cut-off values on the basis of a large multicentre cohort, thus providing more precise criteria for stratification.

During the model development process, performance status was good in all patients, therefore was not included as a variable. However, liver function and biological features still might be crucial determinants of prognosis[23] and were primarily considered for model development. Interestingly, unlike in previous scores where bilirubin and albumin are major components, these liver function parameters were not identified as independent prognostic factors according to Cox regression analysis, whereas the addition of the biomarker AFP on the basis of tumour burden (model 2) only provided marginal improvement compared to the tumour-burden-based model, without statistically significant difference (Table 2) regarding model performance. This is probably due to the difference in the degree of heterogeneity regarding these parameters within the current population. While tumour burden varies in a relatively wide range, liver function parameters were at a relatively well-preserved level, therefore, tumour burden profiles may play a major role in prediction, while other parameters contribute less to model performance. This disparity of homogeneity in different parameters may also explain the findings that the six-and-twelve score exhibited better performance and discriminative capability than models including more non-tumour-burden parameters but fewer tumour characteristics (HAP score, BCLC-B sub-classification and ALBI score). Understandably, mHAP-II score performed better than HAP score but the advantage was marginal, probably because the former added tumour number but encoded it as a categorized variable. However, it should be acknowledged that models such as HAP score and ALBI score have been validated in other settings, and the absence of their superiority over the six-and-twelve score in this population does not necessarily devalue these scores, especially in populations other than ideal candidates of TACE, due to the increased heterogeneity regarding liver function and performance status. For the same reason, the six-and-twelve score should be further validated in other patients, but still, it appears to be a favourable solution for risk stratification of the best target population for TACE.

An ideal prognostic model for risk stratification should not only be developed with appropriate methods, but also be concise for clinical application.[23] A continuous prognostic tool might better meet the demand. For one thing, loss of information can be reduced with the maintenance of continuity, as was supported in comparison between different models (Table 3). As the only continuous prognostic tool among all previous models, mHAP-III score presented the highest AUROC values among them, approaching that of the six-and-twelve score. Nevertheless, the six-and-twelve score is simpler for quick bedside assessment of patient outcome regardless of AFP level, liver function, or aetiology according to subgroup analysis. Additionally, individualized prediction can be achieved with a continuous prognostic score. Using a nomogram, the current score can be more informative by providing survival probability prediction at different time points for ideal candidates awaiting the TACE procedure, compared to models and scores with categorized presentation. Last but not least, another strength of this study is the clinical usefulness of the model. The current prognostic score, not including radiological response despite its predictive value, reserved the feature of being applied at baseline. When the TACE procedure is considered for a candidate with unresectable HCC, the six-and-twelve model can return an estimated survival probability and median survival as a referential for comparison with expected survival under other treatments.[15] For instance, the current score identifies that patients with a sum of tumour size and number exceeding 12 have a median survival of about 15 months, shorter than that of intermediate stage patients who receive sorafenib for being unfit for or unresponsive to loco-regional treatment.[58,59] This suggests that although all the best target population is supposed to benefit from TACE, the survival improvement for patients in the highest tumour burden stratum may be moderate, and the decision of performing exclusive TACE should be scrutinized, calling for further investigations on strategy refinement in this stratum. In a broader sense, the current model may provide risk stratification criteria to control heterogeneity in future clinical trials that continue to target ideal candidates of TACE.

There are some limitations to our study. Firstly, the risk of selection bias is unavoidable in observational studies. However, this risk has been minimized by the inclusion of all consecutive patients and a large cohort of unresectable HCC candidates of TACE to date. Secondly, HCC in Chinese patients is mainly HBV-related, and might present different tumour characteristics from those in patients from other regions where HCC is characterized by less tumour burden and by other aetiologies such as HCV or alcohol use.[18] Moreover, although liver function parameters were not identified as significant predictors in the current cohort, their prognostic value may still be crucial in patients at a later stage, especially in patients with other aetiologies who might be more prone to liver function impairment after TACE.[6] However, the subgroup analysis regarding aetiology with bootstrapping suggested that our model could be effectively applied in patients with aetiologies other than HBV. Still, it should be admitted that further large-scale external validation in different regions is necessary. Thirdly, treatment response was not considered a candidate parameter. Nevertheless, models including this parameter can only be applied when response can be evaluated after the procedure, leading to a delay in the timing of stratification. Given this limitation, as well as other aforementioned points, the application of our model as a vital prognostic tool in other settings should be further validated.

In conclusion, stemming from a large multicentre cohort of 1,604 patients, the six-and-twelve score is the first prognostic model for stratifying recommended TACE candidates. With an easy-to-use presentation consisting of tumour burden characters, the model exhibited adequate performance with individualized prediction and can stratify patients into 3 strata with significantly different median survival. Therefore, the six-and-twelve score may be helpful for assessing outcomes in clinical practice and for designing clinical trials with comparable criteria and stratified risk. Further validations in patients with different aetiologies from both Eastern and Western cohorts remain highly warranted.

Financial support

This study was supported by the National Natural Science Foundation of China (81172145 and 81420108020) and National Key Technology R&D Program (2015BAI13B07).

Conflict of interest

All authors have nothing to disclose.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors contributions

Qiuhe Wang: acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, and administrative, technical, or material support. Dongdong Xia: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, and administrative, technical, or material support. Wei Bai: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, and administrative, technical, or material support. Enxin Wang: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, and administrative, technical, or material support. Ming Huang, Junhui Sun, Hailiang Li, Jing Li, Hui Zhao, Xingnan Pan, Guowen Yin, Weidong Gong, Xiaoli Zhu, Guoliang Shao, Wei Mu, Zhengyu Lin, Jianbing Wu, Jueshi Liu, Jiaping Li, Chunqing Zhang, Haibin Shi, Zixiang Li, Jinlong Song, Shufa Yang, Wenhui Wang, Yanbo Zheng, Jian Xu, Rong Ding, Yuelin Zhang, Lin Zheng, Nan You, Weiwei Gu, Tao Xu, Hui Yu, Peng Zhang, Shuai Zhang, Hui Zeng, Hui Zhang, Jian Chen, Long Feng, Yongjin Zhang, Lin Liu, Guangchuan Wang, Weizhong Zhou, Xueda Li, Wenbo Shao, Wukui Huang, Lei Li, Wenjin Jiang, Wen Zhang, Jing Niu, Jie Yuan, Xiaomei Li, Hui Chen, Yong Lv, Kai Li, Zhanxin Yin, and Daiming Fan: acquisition of data, critical revision of the manuscript for important intellectual content, and administrative, technical, or material support. Jielai Xia: analysis and interpretation of data, critical revision of the manuscript for important intellectual content, statistical analysis, and administrative, technical, or material support. Guohong Han: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, and administrative, technical, or material support.

Acknowledgements

This study was partially presented as an oral presentation in the International Liver Congress 2017 for the European Association for the Study of the Liver in Amsterdam, Netherlands, on April 2017. We gratefully acknowledge Drs. Wengang Guo, Chuangye He, Xulong Yuan, Tianlei Yu, Zhengyu Wang, Bohan Luo, Na Han and Ying Zhu from Xijing Hospital of Digestive Disease, Dr. Shuangxi Li from the First Affiliated Hospital of Lanzhou University, Dr. Junfeng Li from Tumour Hospital of Yunnan province, Dr. Li Jiang from Henan Cancer Hospital, and Dr. An Li from the Second Affiliated Hospital of Nanchang University for their dedication in performing TACE procedures and follow-up of patients; as well as Drs. Jiahao Fan, Wenjun Wang, Lei Liu, and Haibo Liu from Xijing Hospital of Digestive Disease for their assistance in revising the manuscript.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2019.01.013.

References

Author names in bold designate shared co-first authorship

[1] Bruix J, Reig M, Sherman M. Evidence-based diagnosis, staging, and treatment of patients with hepatocellular carcinoma. Gastroenterology 2016;150:835–853.
[2] Venkatesh SK, Chandan V, Roberts LR. Liver masses: a clinical, radiologic, and pathologic perspective. Clin Gastroenterol Hepatol 2014;12:1414–1429.
[3] Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. Nat Rev Gastroenterol Hepatol 2010;7:448–458.
[4] Bruix J, Sherman M. American Association for the Study of Liver Disease. Management of hepatocellular carcinoma: an update. Hepatology 2011;53:1020–1022.
[5] European Association for the Study of the Liver, European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J. Hepatol. 2012;56:908–943.
[6] Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LR, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology 2018;67:358–380.
[7] Meyer T, Fox R, Ma YT, Ross PJ, James MW, Sturgess R, et al. Sorafenib in combination with transarterial chemoembolisation in patients with unresectable hepatocellular carcinoma (TACE 2): a randomised placebo-controlled, double-blind, phase 3 trial. Lancet Gastroenterol Hepatol 2017;2:565–575.
[8] **Lencioni R**, **Llovet JM**, Han G, Tak WY, Yang J, Guglielmi A, et al. Sorafenib or placebo plus TACE with doxorubicin-eluting beads for intermediate stage HCC: The SPACE trial. J Hepatol 2016;64:1090–1098.
[9] Chao Y, Chung YH, Han G, Yoon JH, Yang J, Wang J, et al. The combination of transcatheter arterial chemoembolization and sorafenib is well tolerated and effective in Asian patients with hepatocellular carcinoma: final results of the START trial. Int J Cancer 2015;136:1458–1467.
[10] Kudo M, Han G, Finn RS, Poon RT, Blanc JF, Yan L, et al. Brivanib as adjuvant therapy to transarterial chemoembolization in patients with hepatocellular carcinoma: a randomized phase III trial. Hepatology 2014;60:1697–1707.
[11] **Yin L**, **Li H**, **Li AJ**, Lau WY, Pan ZY, Lai EC, et al. Partial hepatectomy vs. transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan Criteria: a RCT. J Hepatol 2014;61:82–88.
[12] Lammer J, Malagari K, Vogl T, Pilleul F, Denys A, Watkinson A, et al. Prospective randomized study of doxorubicin-eluting-bead embolization in the treatment of hepatocellular carcinoma: results of the PRECISION V study. Cardiovasc Intervent Radiol 2010;33:41–52.
[13] Sangro B, Salem R. Transarterial chemoembolization and radioembolization. Semin Liver Dis 2014;34:435–443.
[14] European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
[15] Cappelli A, Cucchetti A, Cabibbo G, Mosconi C, Maida M, Attardo S, et al. Refining prognosis after trans-arterial chemo-embolization for hepatocellular carcinoma. Liver Int 2016;36:729–736.
[16] Kadalayil L, Benini R, Pallan L, O'Beirne J, Marelli L, Yu D, et al. A simple prognostic scoring system for patients receiving transarterial embolisation for hepatocellular cancer. Ann Oncol 2013;24:2565–2570.
[17] Takayasu K, Arii S, Kudo M, Ichida T, Matsui O, Izumi N, et al. Superselective transarterial chemoembolization for hepatocellular carcinoma. Validation of treatment algorithm proposed by Japanese guidelines. J Hepatol 2012;56:886–892.
[18] Kim DY, Ryu HJ, Choi JY, Park JY, Lee DY, Kim BK, et al. Radiological response predicts survival following transarterial chemoembolisation in patients with unresectable hepatocellular carcinoma. Aliment Pharmacol Ther 2012;35:1343–1350.
[19] Kim BK, Kim SU, Kim KA, Chung YE, Kim MJ, Park MS, et al. Complete response at first chemoembolization is still the most robust predictor for favorable outcome in hepatocellular carcinoma. J Hepatol 2015;62:1304–1310.
[20] **Wang W**, **Zhao Y**, Bai W, Han G. Response assessment for HCC patients treated with repeated TACE: the optimal time-point is still an open issue. J Hepatol 2015;63:1530–1531.
[21] Bolondi L, Burroughs A, Dufour JF, Galle PR, Mazzaferro V, Piscaglia F, et al. Heterogeneity of patients with intermediate (BCLC B) Hepatocellular Carcinoma: proposal for a subclassification to facilitate treatment decisions. Semin Liver Dis 2012;32:348–359.

[22] Sieghart W, Hucke F, Peck-Radosavljevic M. Transarterial chemoembolization: modalities, indication, and patient selection. J Hepatol 2015;62:1187–1195.
[23] Pinato DJ, Howell J, Ramaswami R, Sharma R. Review article: delivering precision oncology in intermediate-stage liver cancer. Aliment Pharmacol Ther 2017;45:1514–1523.
[24] Hsu CY, Lee YH, Hsia CY, Huang YH, Su CW, Lin HC, et al. Performance status in patients with hepatocellular carcinoma: determinants, prognostic impact, and ability to improve the Barcelona Clinic Liver Cancer system. Hepatology 2013;57:112–119.
[25] Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet 2018;391:1301–1314.
[26] Matsui O, Kadoya M, Yoshikawa J, Gabata T, Arai K, Demachi H, et al. Small hepatocellular carcinoma: treatment with subsegmental transcatheter arterial embolization. Radiology 1993;188:79–83.
[27] Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. Lancet Oncol 2009;10:35–43.
[28] Yamakado K, Miyayama S, Hirota S, Mizunuma K, Nakamura K, Inaba Y, et al. Subgrouping of intermediate-stage (BCLC stage B) hepatocellular carcinoma based on tumor number and size and Child-Pugh grade correlated with prognosis after transarterial chemoembolization. Jpn J Radiol 2014;32:260–265.
[29] Park Y, Kim SU, Kim BK, Park JY, Kim DY, Ahn SH, et al. Addition of tumor multiplicity improves the prognostic performance of the hepatoma arterial-embolization prognostic score. Liver Int 2016;36:100–107.
[30] **Pinato DJ**, **Sharma R**, Allara E, Yen C, Arizumi T, Kubota K, et al. The ALBI grade provides objective hepatic reserve estimation across each BCLC stage of hepatocellular carcinoma. J Hepatol 2017;66:338–346.
[31] **Yau T**, **Tang VY**, **Yao TJ**, Fan ST, Lo CM, Poon RT. Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma. Gastroenterology 2014;146, 1691–1700.e1693.
[32] Kudo M, Chung H, Osaki Y. Prognostic staging system for hepatocellular carcinoma (CLIP score): its value and limitations, and a proposal for a new staging system, the Japan Integrated Staging Score (JIS score). J Gastroenterol 2003;38:207–215.
[33] Chun YS, Pawlik TM, Vauthey JN. 8th edition of the AJCC Cancer staging manual: pancreas and hepatobiliary cancers. Ann Surg Oncol 2018;25:845–847.
[34] Tateishi R, Yoshida H, Shiina S, Imamura H, Hasegawa K, Teratani T, et al. Proposal of a new prognostic model for hepatocellular carcinoma: an analysis of 403 patients. Gut 2005;54:419–425.
[35] **Farinati F**, **Vitale A**, Spolverato G, Pawlik TM, Huo TL, Lee YH, et al. Development and validation of a new prognostic system for patients with hepatocellular carcinoma. PLoS Med 2016;13 e1002006.
[36] Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995;48:1503–1510.
[37] Buuren Sv, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Softw 2010:1–68.
[38] Harrell F. Regression modeling strategies. BIOS 2014;330.
[39] Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med 2013;32:5381–5397.
[40] Davison AC, Hinkley DV. Bootstrap methods and their application. Cambridge University Press; 1997.
[41] European Association for the Study of the Liver. EASL 2017 clinical practice guidelines on the management of hepatitis B virus infection. J Hepatol 2017;67:370–398.
[42] European Association for the Study of the Liver. EASL recommendations on treatment of hepatitis C 2018. J Hepatol 2018;69:461–511.
[43] Lencioni R, de Baere T, Soulen MC, Rilling WS, Geschwind JF. Lipiodol transarterial chemoembolization for hepatocellular carcinoma: a systematic review of efficacy and safety data. Hepatology 2016;64:106–116.
[44] Malagari K, Pomoni M, Moschouris H, Bouma E, Koskinas J, Stefaniotou A, et al. Chemoembolization with doxorubicin-eluting beads for unresectable hepatocellular carcinoma: five-year survival analysis. Cardiovasc Int Radiol 2012;35:1119–1128.
[45] **Burrel M**, **Reig M**, Forner A, Barrufet M, de Lope CR, Tremosini S, et al. Survival of patients with hepatocellular carcinoma treated by transarterial chemoembolisation (TACE) using drug eluting beads. Implications for clinical practice and trial design. J Hepatol 2012;56:1330–1335.
[46] Galle PR, Tovoli F, Foerster F, Worns MA, Cucchetti A, Bolondi L. The treatment of intermediate stage tumours beyond TACE: from surgery to systemic therapy. J Hepatol 2017;67:173–183.
[47] Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 1999;19:329–338.
[48] Giannini EG, Farinati F, Ciccarese F, Pecorelli A, Rapaccini GL, Di Marco M, et al. Prognosis of untreated hepatocellular carcinoma. Hepatology 2015;61:184–190.
[49] Llovet JM, Bruix J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: chemoembolization improves survival. Hepatology 2003;37:429–442.
[50] Sala M, Llovet JM, Vilana R, Bianchi L, Sole M, Ayuso C, et al. Initial response to percutaneous ablation predicts survival in patients with hepatocellular carcinoma. Hepatology 2004;40:1352–1360.
[51] Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med 1996;334:693–699.
[52] Giannini EG, Moscatelli A, Pellegatta G, Vitale A, Farinati F, Ciccarese F, et al. Application of the intermediate-stage subclassification to patients with untreated hepatocellular carcinoma. Am J Gastroenterol 2016;111:70–77.
[53] Hucke F, Pinter M, Graziadei I, Bota S, Vogel W, Muller C, et al. How to STATE suitability and START transarterial chemoembolization in patients with intermediate stage hepatocellular carcinoma. J Hepatol 2014;61:1287–1296.
[54] Ogasawara S, Chiba T, Ooka Y, Kanogawa N, Motoyama T, Suzuki E, et al. A prognostic score for patients with intermediate-stage hepatocellular carcinoma treated with transarterial chemoembolization. PLoS ONE 2015;10 e0125244.
[55] Xu L, Peng ZW, Chen MS, Shi M, Zhang YJ, Guo RP, et al. Prognostic nomogram for patients with unresectable hepatocellular carcinoma after transcatheter arterial chemoembolization. J Hepatol 2015;63:122–130.
[56] **Kim BK**, **Shim JH**, Kim SU, Park JY, Kim DY, Ahn SH, et al. Risk prediction for patients with hepatocellular carcinoma undergoing chemoembolization: development of a prediction model. Liver Int 2016;36:92–99.
[57] Kim JH, Shim JH, Lee HC, Sung KB, Ko HK, Ko GY, et al. New intermediate-stage subclassification for patients with hepatocellular carcinoma treated with transarterial chemoembolization. Liver Int 2017;37:1861–1868.
[58] Bruix J, Raoul JL, Sherman M, Mazzaferro V, Bolondi L, Craxi A, et al. Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma: subanalyses of a phase III trial. J Hepatol 2012;57:821–829.
[59] Iavarone M, Cabibbo G, Piscaglia F, Zavaglia C, Grieco A, Villa E, et al. Field-practice study of sorafenib therapy for hepatocellular carcinoma: a prospective multicenter study in Italy. Hepatology 2011;54:2055–2063.