Recommendation of mHAP and ABCR scoring systems for the decision-making of the first and subsequent TACE session in HCC patients

Kefeng Jia[a,*], Weili Yin[b,*], Zhongsong Gao[a], Wen Shen[c], Fang Wang[b], Shuangshuang Xie[c], Mingge Li[a] and Rong Lv[a]

Background Due to the high heterogeneity among hepatocellular carcinoma (HCC) patients receiving transarterial chemoembolization (TACE), the prognosis of patients varies significantly. Various predictive scoring systems have been developed to identify the patients who could benefit from TACE. However, there is no consensus on which is better. This study aims to validate and compare the predictive capabilities of scoring systems for first and subsequent TACE.

Materials A total of 524 HCC patients were treated with TACE, and 222 patients who met the inclusion criteria were included. Log-rank test was used to verify the predictive value of six scoring systems for the first TACE and four TACE retreatment scoring systems. Harrell's concordance (C)-index, likelihood ratio and integrated Brier score (IBS) were used to compare the predictive performance.

Results For the scoring systems of TACE, the overall survival (OS) of candidates screened by Hepatoma Arterial-embolization Prognostic (HAP), modified HAP (mHAP), mHAP3, alpha-fetoprotein, Barcelona Clinic Liver Cancer, Child-Pugh and Response (ABCR), albumin-bilirubin grade (ALBI), tumor size, alpha-fetoprotein, first TACE response and pre-/post-TACE was significantly longer than that of the noncandidates (all $P < 0.05$), whereas the mHAP2 and assessment for retreatment with TACE did not distinguish the candidates from noncandidates ($P = 0.206$, 0.115, respectively). The predictive and calibration performances of mHAP and ABCR were the highest for the first TACE and TACE retreatment, respectively.

Conclusion mHAP identifies the patients who could benefit from the first TACE, whereas ABCR distinguishes patients who could benefit from subsequent TACE sessions Eur J Gastroenterol Hepatol 35: 461–470 Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc.

Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignant tumor and the second leading cause of cancer-related deaths worldwide. About 750 000 patients are diagnosed with HCC every year, among which about 55% are from China [1,2]. To date, transarterial chemoembolization (TACE) has been widely endorsed by several guidelines, including the Barcelona Clinic Liver Cancer (BCLC) staging system [3], the American Association for the Study of Liver Diseases (AASLD) [4] and the European

Association for the Study of the Liver (EASL) [5] and is the recommended first-line treatment for asymptomatic patients with intermediate stage, large size/multifocal HCC without major vascular invasion or extrahepatic metastasis. However, the recommended patients for TACE treatment constitute a highly heterogeneous subpopulation characterized by different tumor burdens and hepatic functional reserve [6]. The difference in the median survival ranged from 13–43 months in this population [7,8], which indicated that some of the patients might not be suitable for TACE treatment, and hence, alternative therapies should be administered. Therefore, several TACE prognostic scoring systems have been developed to address the heterogeneity of patients. Among these, the influencing prognostic models include the Selection for Transarterial chemoembolization Treatment (STATE) scoring system [9], hepatoma arterial-embolization prognostic (HAP) scoring system [10], modified versions of HAP [modified HAP (mHAP) [11], mHAP2 [12] and mHAP3 [13]] that predict the prognosis after first TACE session, the assessment for retreatment with TACE (ART) [14] and alpha-fetoprotein (AFP), BCLC, Child-Pugh, and response (ABCR) scoring systems [15] that evaluate whether patients could benefit from TACE retreatment. In addition, other scoring systems were reported in 2020, namely the albumin-bilirubin grade (ALBI), tumor size, AFP, first TACE response (ASAR) [16] and pre-/post-TACE scoring system [17]. Although these TACE scoring systems have been modified and developed, the predictive values in clinical practices

are yet debatable due to differences in HCC patients from different regions, TACE practices, and the professional levels of different settings. In fact, TACE scoring systems have been gradually recommended by multiple guidelines or consensus, but there is no consensus on which one is better [5,18,19]. In addition, the development and clinical verification of most of the scoring systems did not include Chinese HCC patients worldwide. Thus, it is necessary to verify these scoring systems in Chinese patients, mainly suffering from HBV infection. The present study aimed to verify the scoring systems for the first and subsequent TACE and compare the predictive performance of the scoring systems.

Patients and methods

Patients

From January 2008 to December 2013, 524 HCC patients were treated with the first TACE in our hospital. HCC was diagnosed according to the AASLD/EASL [4,5] guidelines and then further proven by histological or imaging examination. The inclusion criteria were as follows: (1) all patients age >18-year-old; (2) received ≥2 TACE sessions; (3) liver function of Child-Pugh class A or B; (4) BCLC stage A, B or C HCC, but the portal vein tumor thrombus (PVTT) involved only at branches; (5) Eastern Cooperative Oncology Group activity score of 0 or 1 point. The exclusion criteria were as follows: (1) any previous antitumor therapies or received TACE as a bridge for any other therapy before the end of follow-up (including surgical resection, liver transplantation, radiofrequency ablation, sorafenib or other systemic therapies); (2) patients with PVTT at the main trunk; (3) imaging evaluation after first TACE showed complete response (CR); (4) data were incomplete. Consequently, 302 patients were excluded, and finally, 222 patients

were included in this study (Fig. 1). Baseline and follow-up characteristics, including the laboratory examination results within 3 days before the first and subsequent TACE and the imaging data, were recorded. This study was performed in accordance with the Ethics of the Declaration of Helsinki of the World Medical Association and approved by the Ethics Committee of the Tianjin Third Central Hospital (IRB2020-025-01). In this retrospective study, informed consent of the patients was waived. However, all the data were obtained after the patients or their families signed informed consent.

TACE session, treatment effectiveness evaluation and follow-up

TACE was performed by experienced interventional radiologists with over 10-year experience (K.J. and C.Y.). The modified Seldinger method was adopted to puncture the femoral artery and insert a 5 F catheter sheath in patients under local anesthesia. The 4 F RH catheter was used for selective catheterization into the celiac trunk artery and superior mesenteric artery for angiography to verify the anatomical structure of the variant hepatic artery. Computed tomography during hepatic angiography and arterial portography was performed to evaluate the locations, sizes, numbers and patency of tumors supplying arteries and portal veins. Then, the microcatheter was superselective intubated into the third or fourth hepatic artery branch supplying blood to the target tumor. The infusion of iodized oil (2–20 ml) and adriamycin (10–60 mg) was injected, followed by embolization with gel foam particles (150–350 μm or 350–560 μm). The endpoint of embolization was that the tumor intravascular blood flow was significantly reduced compared with the initial blood flow, and the contrast medium was not emptied in 3–5 cardiac cycles. The doses of the chemotherapeutic drugs and embolization materials were decided according to the tumor burden, tumor characteristics and hepatic functional reserve [20]. None of the patients in this study underwent drug-eluting beads-TACE (DEB-TACE).

Adverse events (AEs) were evaluated in accordance with the National Cancer Institute Common termination criteria for adverse events version 5.0, and the AEs were divided into five levels according to the severity. Grade 1: asymptomatic or mild, only clinical or diagnostic findings, without treatment; Grade 2: requires minor, local or noninvasive treatment; Grade 3: serious or medically significant but not immediately life-threatening, resulting in hospitalization or prolonged hospitalization; Grade 4: life-threatening; emergency treatment is required; Grade 5: AE related death.

CT or MRI examination was performed at 6–8 weeks after TACE. For patients with residual activity of tumor or new lesions, 'on-demand' TACE was performed according to the tumor response and hepatic functional reserve. The modified Response Evaluation Criteria in Solid Tumors (mRECIST) criteria [21] were used to evaluate the tumor response. CR and partial response (PR) were considered as the presence of radiological responses, whereas stable disease (SD) and progressive disease (PD) were considered as the absence of radiological response. The time of all-cause death or last follow-up of the patients was recorded by reviewing the medical records or telephone follow-up. The date of last follow-up in this study was 30 November 2020.
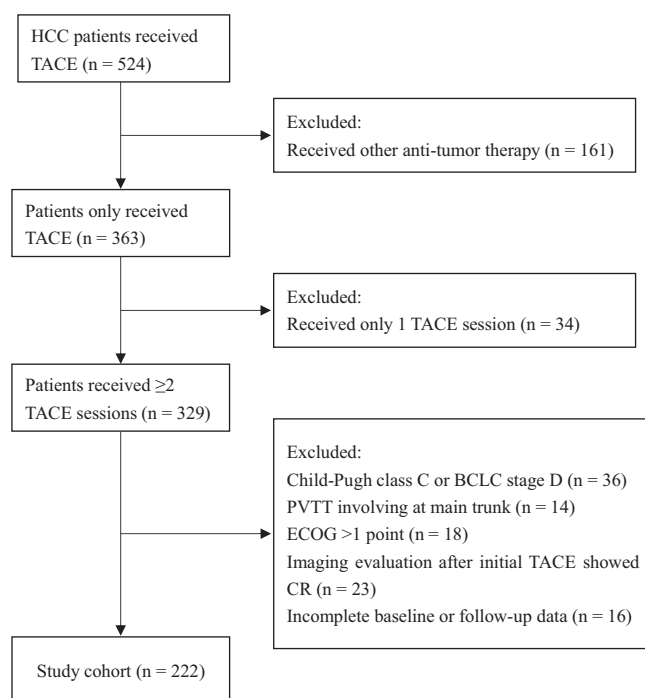

Fig. 1. Schematic of the screening of patients.

Risk stratification and candidate identification by transarterial chemoembolization scoring systems

According to previous studies, the baseline and follow-up characteristics required for the six TACE scoring systems for the first TACE session and four TACE scoring systems for subsequent TACE sessions, as well as the risk stratification and candidate identification, are shown in Table 1. Among these scoring systems, ASA(R) [16] was the modified version of the ASAR scoring system, which removed the characteristic of 'radiological response' from the original model and was used for the validation of the first TACE session. Moreover, the STATE scoring system was not verified in this study, as the C-reactive protein (CRP), which is not a routinely measured index in our center, was required.

Statistical analysis

Quantitative data in normal distribution were presented as mean ± SD. Quantitative data not in normal distribution were described with median and interquartile range (IQR), and qualitative data were described with frequencies and percentages. The overall survival (OS) was defined as the day before TACE initiation until death or last follow-up. Kaplan–Meier survival curve was plotted, and the log-rank test was used to compare the differences in the survival of candidates and noncandidates identified by each scoring system. Harrell's concordance (C)-index and likelihood ratio (LR) were calculated to evaluate and compare the prognostic performance of each model in the patients included in this study. The higher the C-index and LR χ^2 values, the better the performance of the scoring system in distinguishing the survival differences between candidates and noncandidates. Brier score-based predictive error curve was used to quantitatively and qualitatively evaluate and compare the calibration performances of different scoring systems. The integrated Brier score (IBS) within 60 months was calculated as the summary measure of predictive error. The lower the Brier score, the lower the predictive error and the higher the calibration performance of the corresponding scoring system. SPSS software version 20.0 (SPSS Inc., Chicago, Illinois, USA) and R version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria) were used for all statistical analyses. A two-sided P-value < 0.05 indicated statistical significance.

Results

General characteristics of the patients

The baseline and follow-up characteristics of the patients are shown in Table 2. The mean age of the cohort was 59.7 ± 8.8 years old and consisted of 80.2% (178/222) males. The majority of the patients were at BCLC stage B (65.8%, 146/222). In our population, 214 patients (96.4%) had liver cirrhosis. Most of them were associated with HBV infection (79.0%, 169/214). Among them, 116 patients had a Child-Pugh class of A (54.2%) and 98 patients had Child-Pugh class B (45.8%). Grade 1/2 AEs occurred in 78 patients (35.1%) after the first TACE, mainly including abdominal pain, fever and nausea with or without vomiting. The increase of alanine aminotransferase

(ALT) or aspartate aminotransferase (AST) exceeded the upper limit of the normal value by three times but not more than five times, the increase of total bilirubin (T-Bil) exceeded the upper limit of the normal value by 1.5 times but not more than three times; Grade 3 AEs occurred in 16 (7.2%) patients, including severe abdominal pain, fever or femoral artery pseudoaneurysm which prolonged hospitalization, ALT or AST rising more than five times the upper limit of normal value and T-Bil rising more than three times the upper limit of normal value. No grade 4/5 patients were found in this study cohort. The median time interval between the first and second TACE sessions was 54.5 (46.0, 68.3) days. The mean follow-up time of the patients was 32.4 ± 26.6 (2.0–144.8) months. At the end of the follow-up, 13 patients (5.9%) survived and 209 patients (94.1%) died.

Verification of using transarterial chemoembolization scoring system to predict the survival differences between candidates and noncandidates

The median OS of the 222 patients was 24.2 (95% CI, 28.8–35.9) months. The numbers and corresponding OS of the candidates and noncandidates screened by the risk stratification of TACE scoring systems are shown in Table 3. Among these 222 patients, eight with chronic hepatitis but no cirrhosis were not included in the Child-Pugh class and increase in the score. Therefore, only 214 patients were included in the analysis of ART and ABCR scoring systems. The log-rank test showed that among the scoring systems, mHAP2 and ART could not predict the survival difference between the candidates and noncandidates ($P = 0.206$, $P = 0.115$, respectively). The Kaplan–Meier survival curves of the scoring systems are shown in Figs. 2 and 3.

In total 214 patients with liver cirrhosis were divided into the Child A group (116 cases) and Child B group (98 cases). The OS was 27.0 (95% CI, 21.0–31.5) months in the Child A group and 24.0 (95% CI, 21.6–29.0) months in the Child B group. The difference was statistically significant ($P = 0.017$). Kaplan–Meier survival curve of the Child A group and Child B group was shown in Supplementary Figure S1, Supplemental digital content 1, *http://links.lww.com/EJGH/A828*. The 1-, 3- and 5-year survival rates of the patients in the Child A group were 81.8, 36.2 and 20.8%, respectively, and those in the Child B group were 76.5, 28.5 and 6.1%, respectively.

Comparison of performance of transarterial chemoembolization scoring systems in predicting the prognosis of the patient cohort in this study

For the first TACE, the mHAP scoring system had the highest distinguishing capability at all time points during the follow-up (Harrell's C-index 0.575; 95% CI, 0.540–0.610) and the optimal homogeneity within the classification (LR $\chi^2 = 13.91$) (Table 4 and Fig. 4a), whereas for the TACE retreatment scoring systems, ABCR had the highest Harrell's C-index (0.604; 95% CI, 0.571–0.637) and highest LR χ^2 (27.01) (Table 4 and Fig. 4b). When evaluating the calibration performances of the prognostic models within 60 months based on Kaplan–Meier analysis, mHAP and ABCR scoring systems had the lowest IBS

Table 1. Characteristics for the predictive scoring systems, risk stratification, and candidate identification

Scoring system	Characteristics (score)		Risk stratification (prognosis score)		Candidates
For first TACE					
HAP	Albumin <36 g/L	1	HAP A	0	Yes
	Bilirubin >17 µmol/L	1	HAP B	1	Yes
	AFP >400 ng/mL	1	HAP C	2	No
	Tumor diameter >7 cm	1	HAP D	3–4	No
mHAP	Albumin <36 g/L	1	mHAP A	0	Yes
	AFP >400 ng/mL	1	mHAP B	1	Yes
	Tumor diameter >7 cm	1	mHAP C	2	No
			mHAP D	3	No
mHAP2	Albumin <36 g/L	1	mHAP2 A	0	Yes
	Bilirubin >0.9 mg/dL	1	mHAP2 B	1	Yes
	AFP >400 ng/mL	1	mHAP2 C	2	No
	Tumor diameter >7 cm	1	mHAP2 D	3–5	No
	Tumor number ≥2	1			
mHAP3	The prognostic index (PI) formula = $(0.104 \times$ size in cm$) + [0.3089 \times$ number (single nodule = 1; 2-3 nodules = 2; more than three nodules = 3$)] + (0.2185 \times \log_{10}$AFP in ng/mL$) - (0.4049 \times$ albumin in g/dL$) + (0.1506 \times$ bilirubin in mg/dL$)$		<Median PI		Yes
			≥Median PI		No
ASA(R)	ALBI grade 1	0	ASA(R) <4		Yes
	2	1	ASA(R) ≥4		No
	3	2			
	Maximal tumor diameter <3 cm	0			
	3–5 cm	1			
	≥5 cm	2			
	AFP <400 ng/mL	0			
	≥400 ng/mL	1			
Pre-TACE	Linear predictor = $0.313 \times$ tumor number $(0 =$ solitary, $1 =$ multifocal$) + 1.252 \times \log_{10}$tumor size (cm$) + 0.230 \times$ baseline $\log_{10}$AFP (ng/mL$) - 0.0176 \times$ baseline albumin (g/L$) + 0.458 \times$ baseline $\log_{10}$bilirubin (µmol/L$) + 0.437 \times$ VI $(0 =$ no, $1 =$ yes$) + 0.149 \times$ HBV $(0 =$ no, $1 =$ yes$) + 0.333 \times$ alcohol $(0 =$ no, $1 =$ yes$) + 0.211 \times$ other cause if not HCV/HBV/alcohol $(0 =$ no, $1 =$ yes$)$		Risk category 1	≤0.94	Yes
			Risk category 2	>0.94 to ≤ 1.47	Yes
			Risk category 3	>1.47 to ≤ 2.10	No
			Risk category 4	>2.10	No
For re-TACE					
ART	AST increase >25 %	4	ART 0–1.5		Yes
	Child–Pugh increase 1 point	15	ART ≥2.5		No
	≥2 points	3			
	Absence of radiological response	1			
ABCR	AFP ≥200 ng/mL	1	ABCR ≤2		Yes
	BCLC A	0	ABCR ≥3		No
	B	2			
	C	3			
	Child–Pugh increase ≥2 points	2			
	Presence of radiologic response	−3			
ASAR	ALBI grade 1	0	ASAR <4		Yes
	2	1	ASAR ≥4		No
	3	2			
	Maximal tumor diameter <3 cm	0			
	3–5 cm	1			
	≥5 cm	2			
	AFP <400 ng/mL	0			
	≥400 ng/mL	1			
	Tumor response CR, PR	0			
	SD, PD	1			
Post-TACE	Linear predictor = $0.207 \times$ tumor number $(0 =$ solitary, $1 =$ multifocal$) + 1.129 \times \log_{10}$tumor size (cm$) + 0.147 \times$ baseline $\log_{10}$AFP (ng/mL$) + 0.750 \times$ baseline $\log_{10}$bilirubin (µmol/L$) + 0.447 \times$ VI $(0 =$ no, $1 =$ yes$) + 0.469 \times$ PR $(0 =$ no, $1 =$ yes$) + 1.143 \times$ SD $(0 =$ no, $1 =$ yes$) + 1.354 \times$ PD $(0 =$ no, $1 =$ yes$)$		Risk category 1	≤1.82	Yes
			Risk category 2	>1.82 to ≤ 2.49	Yes
			Risk category 3	>2.49 to ≤ 3.37	No
			Risk category 4	> 3.37	No

AFP, alpha-fetoprotein; ALBI, albumin-bilirubin grade; AST, aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancer; CR, complete response; HBV, hepatitis B virus; HCV, hepatitis C virus; PD, progressive disease; PR, partial response; re-TACE, repeated transarterial chemoembolization; SD, stable disease; TACE, transarterial chemoembolization; VI, vascular invasion.

(0.165 and 0.159, respectively) in the prognosis models for first and subsequent TACE, respectively. The predictive error curves are shown in Fig. 5. Subsequently, the prediction performance of some scoring systems was compared again in the patients of the Child A group and Child B group, respectively. The results showed that the mHAP and ABCR models still had relatively good performance (results were shown in Supplementary Table S1 and S2, Supplemental digital content 1, *http://links.lww.com/EJGH/A828*).

Table 2. Baseline and follow-up characteristics and tumor responses of the patients

Characteristic	N (%), or mean ± SD, or median (IQRs)
Before first TACE	
Age (years)	59.7 ± 8.8
Gender	
Male/female	178 (80.2%)/44 (19.8%)
Cirrhosis	
Yes/no	214 (96.4%)/8 (3.6%)
Cause of cirrhosis	
HBV/HCV/Alcohol/other	169 (79.0%)/9 (4.2%)/10 (4.7%)/26 (12.1%)
Child-Pugh class[a]	
A/B	116 (54.2%)/98 (45.8%)
BCLC stage	
A/B/C	25 (11.3%)/146 (65.8%)/51 (23.0%)
ALBI grade	
1/2/3	49 (22.1%)/163 (73.4%)/10 (4.5%)
Maximal tumor diameter (cm)	5.0 (3.5, 7.5)
Maximal tumor diameter	
≤7 cm/>7 cm	156 (70.3%)/66 (29.7%)
Maximal tumor diameter	
<3 cm/ 3–5 cm/≥5 cm	42 (18.9%)/63 (28.4%)/117 (52.7%)
Tumor number	
1/2–3/>3	77 (34.7%)/114 (51.4%)/31 (14.0%)
VI	
Yes/no	51 (23.0%)/171 (77.0%)
Albumin (g/L)	35.6 ± 5.1
Albumin	
<36 g/L/≥36 g/L	106 (47.7%)/116 (52.3%)
Bilirubin (μmol/L)	19.5 (14.2, 28.7)
Bilirubin	
≤17 μmol/L/>17 μmol/L	90 (40.5%)/132 (59.5%)
Bilirubin	
≤0.9 mg/dL/>0.9 mg/dL	73 (32.9%)/149 (67.1%)
AFP (ng/mL)	77.7 (10.3, 816.5)
AFP	
<200 ng/mL/≥200 ng/mL	128 (57.7%)/94 (42.3%)
AFP	
<400 ng/mL/≥400 ng/mL	142 (64.0%)/80 (36.0%)
AST (U/L)	33.0 (20.0, 52.0)
Before re-TACE	
AST (U/L)	33.0 (20.0, 54.0)
AST increase	
>25%/≤25%	77 (34.7%)/145 (65.3%)
Child-Pugh increase[a]	
<1 point/1 point/≥2 points	166 (77.6%)/38 (17.8%)/10 (4.6%)
Tumor response	
PR/SD/PD	122 (55.0%)/67 (30.2%)/33 (14.9%)

AFP, alpha-fetoprotein; ALBI, albumin-bilirubin grade; AST, aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancer; HBV, hepatitis B virus; HCV, hepatitis C virus; PD, progressive disease; PR, partial response; re-TACE, repeated transarterial chemoembolization; SD, stable disease; TACE, transarterial chemoembolization; VI, vascular invasion.
[a]Child-Pugh class was only applied to patients with liver cirrhosis, and thus 8 patients with chronic hepatitis but no cirrhosis were not included in the Child-Pugh class and increase in the score.

Table 3. Log-rank test for the transarterial chemoembolization scoring system-based candidates and noncandidates

Scoring system	Candidates/noncandidates			
	Number of patients	Median OS, mo (95% CI)	χ^2	*P* value
For first TACE				
HAP	85/137	31.0 (27.0–40.0)/22.6 (19.0–24.5)	11.590	<0.001
mHAP	153/69	28.5 (25.0–32.3)/17.5 (14.0–23.3)	15.371	<0.001
mHAP2	37/185	27.0 (22.5–39.5)/24.0 (21.6–28.3)	1.601	0.206
mHAP3	111/111	30.5 (25.0–39.6)/22.0 (17.5–25.0)	10.572	0.001
ASA(R)	172/50	26.5 (24.0–31.0)/16.5 (11.5–24.5)	8.140	0.005
Pre-TACE	70/152	31.5 (25.3–41.0)/23.0 (18.5–25.0)	6.223	0.013
For re-TACE				
ART	147/67	26.0 (24.0–30.6)/22.0 (16.0–30.0)	2.542	0.115
ABCR	161/53	29.3 (25.0–35.0)/11.5 (9.0–17.5)	33.621	<0.001
ASAR	142/80	29.3 (25.0–34.0)/16.5 (11.5–24.0)	19.136	<0.001
Post-TACE	40/182	32.3 (27.0–46.9)/23.3 (20.0–26.0)	3.975	0.046

ABCR, alpha-fetoprotein, Barcelona clinic liver cancer, Child-Pugh, and response; ART, assessment for retreatment with TACE; ASAR, albumin-bilirubin grade, tumor size, alpha-fetoprotein, first TACE response; HAP, hepatoma arterial-embolization prognostic; IBS, integrated Brier score; mHAP, modified hepatoma arterial-embolization prognostic; tumor size, alpha-fetoprotein, first TACE response; OS: overall survival; re-TACE, repeated transarterial chemoembolization; TACE, transarterial chemoembolization.

to predictive and calibration performances; (3) TACE was the only treatment received by the included patients. The prediction of the prognosis of patients could be reflected by the scoring models after any other antitumor treatments were excluded.

Since not only the prognostic scores predicted by the first TACE treatment should be verified and compared but also the prognostic model of the second TACE treatment should be evaluated. Therefore, patients who received two or more TACE treatments were included in this study cohort, while patients who received only one treatment were excluded. Patients with tumor response assessed as CR after the first TACE treatment are also excluded because these patients will directly enter the follow-up process rather than TACE treatment again, that is, implement 'on-demand' TACE. To be closer to the actual clinical situation, in addition to patients with BCLC stage B, this study also included a part of HCC patients with BCLC stage A ($n = 25$, 11.3%) and stage C ($n = 51$, 23.0%), which was similar to some original study cohorts that developed TACE scores [10–13,15,17].

The findings of this study showed that the mHAP2 scoring system could not predict the survival difference between the candidates and noncandidates ($P = 0.206$), and the comparisons between predictive and calibration performances of the various models also showed that mHAP2 had dismal performance (Harrell's C-index = 0.517; LR $\chi^2 = 1.69$; IBS = 0.171) (Table 4). The disadvantages of the mHAP2 score were (1) compared with HAP and mHAP scores, the mHAP2 score includes the characteristic of tumor number. Tumor number seems to be an indicator of TACE prognosis but it has not been verified in this study cohort. This may be related to the segmental or subsegmental embolization of each lesion during TACE, which can usually enable multiple HCC with similar cell differentiation types or blood supply to obtain similar curative effects, which may weaken the prognostic role of the characteristic of tumor number to a certain extent. For instance, ART, ABCR and ASAR scoring systems did not
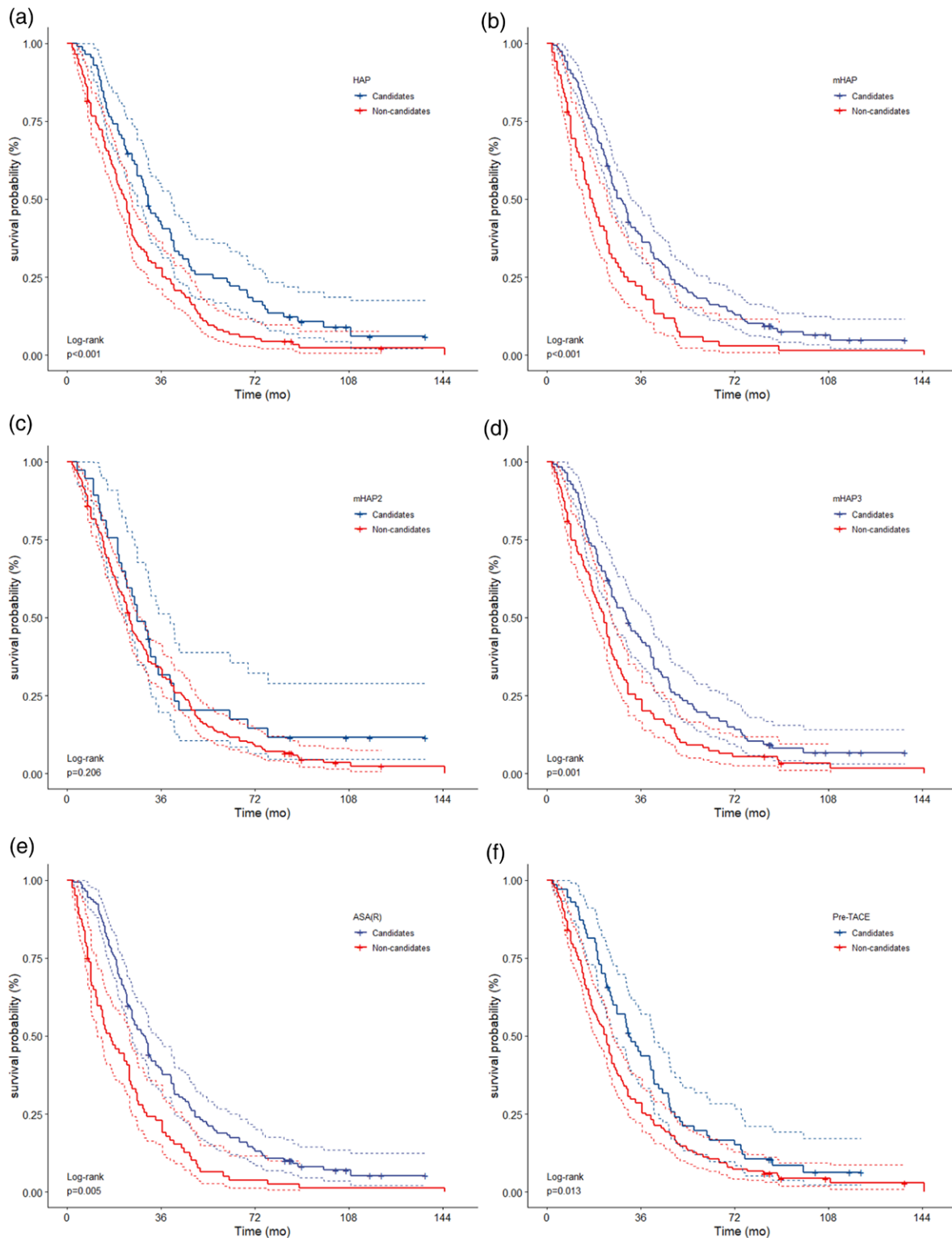
Discussion

The heterogeneity of HCC patients leads to variances in the effects of TACE treatment. The HCC management guidelines of EASL [5] and Chinese [18], as well as TACE clinical practice guideline [19] indicate that TACE scoring systems can be used as the individualized prognostic evaluation and risk stratification model applied before TACE and recalibrated after TACE, whereas the previously developed TACE scoring systems did not show consistent results in consequent studies. This study evaluated these scoring systems and had several advantages as follows compared to previous studies: (1) first verified the feasibility of 10 prognostic models, including the ASAR/ASA(R), pre-TACE and post-TACE scoring systems; (2) comprehensive comparison of the scoring systems with respect

Fig. 2. Kaplan–Meier survival curves between candidates and noncandidates according to the scoring systems of HAP, mHAP, mHAP2, mHAP3, ASA(R) and Pre-TACE (a–f).

include tumor numbers as a characteristic [14–16]. (2) mHAP2 score included one additional characteristic than the HAP score, but did not increase the risk stratification correspondingly; instead, patients with prognosis scores of 3–5 were categorized into high-risk (grade D) stratification, which comprises more patients as the noncandidates for TACE session. In this study, the great differences in

the number of candidates (16.7%; 37/222) and non-candidates (83.3%; 185/222), as per the discrimination by the mHAP2 scoring system, which might influence the prediction performance of mHAP2 scoring system. Interestingly, the mHAP3 model divided the patients into two groups after calculating the continuous prediction scoring systems of each patient; also, it accurately predicted the
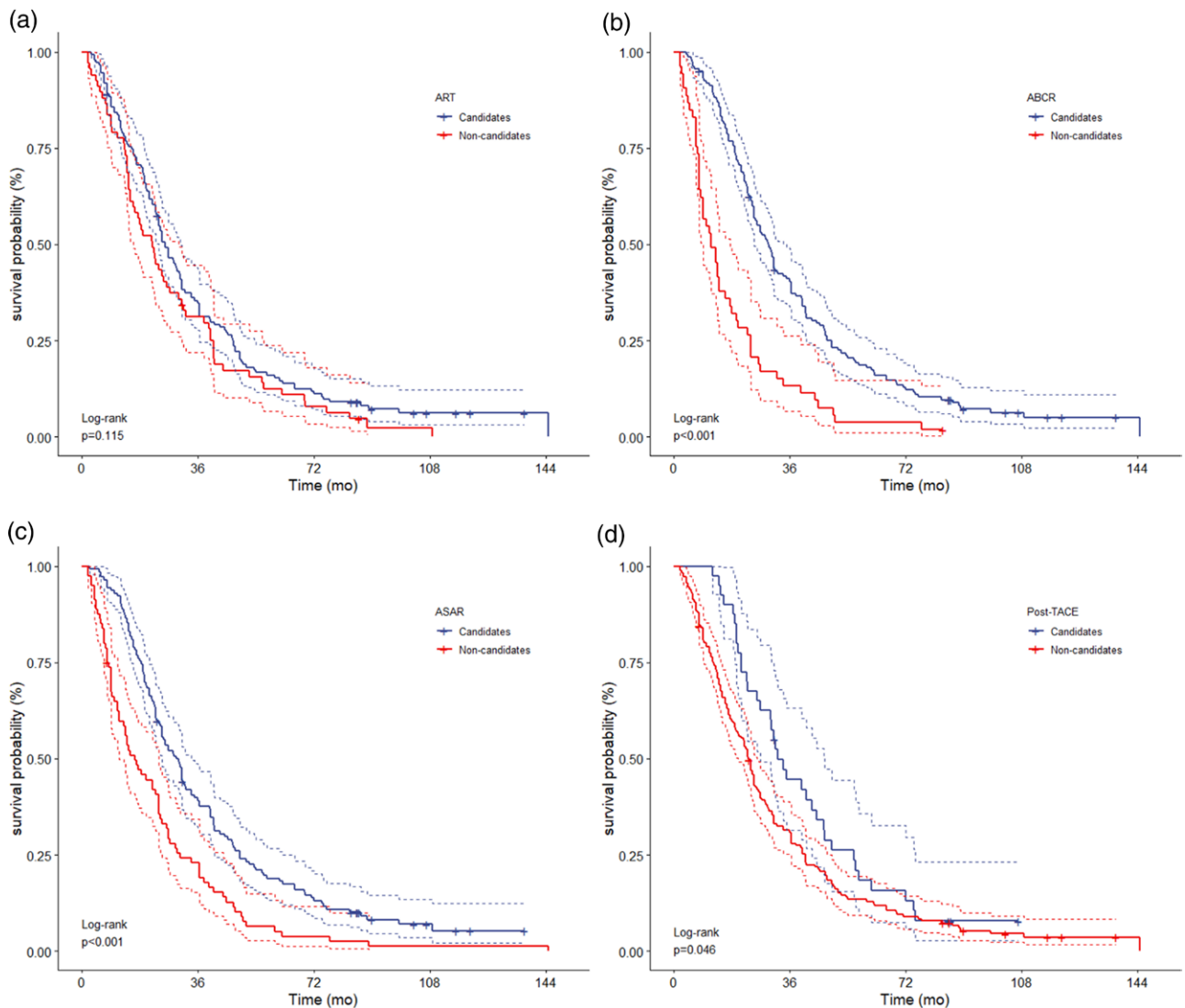
Fig. 3. Kaplan–Meier survival curves between candidates and noncandidates according to the scoring system of ART, ABCR, ASAR and post-TACE (a–d).

Table 4. Comparison of the predictive and calibration performances of transarterial chemoembolization scoring systems

Scoring system	Harrell's C-index (95% CI)	LR χ^2	1-year IBS	3-year IBS	5-year IBS
For first TACE					
HAP	0.569 (0.534–0.604)	11.78	0.067	0.170	0.167
mHAP	0.575 (0.540–0.610)	13.91	0.066	0.168	0.165
mHAP2	0.517 (0.490–0.544)	1.69	0.068	0.175	0.171
mHAP3	0.570 (0.533–0.606)	10.43	0.067	0.170	0.166
ASA(R)	0.554 (0.521–0.587)	7.29	0.067	0.171	0.168
Pre-TACE	0.559 (0.526–0.592)	6.47	0.067	0.171	0.168
For re-TACE					
ART	0.530 (0.493–0.567)	2.41	0.068	0.174	0.171
ABCR	0.604 (0.571–0.637)	27.01	0.061	0.157	0.159
ASAR	0.590 (0.555–0.625)	17.41	0.064	0.166	0.164
Post-TACE	0.548 (0.523–0.573)	4.30	0.067	0.172	0.169

ABCR, alpha-fetoprotein, Barcelona clinic liver cancer, Child-Pugh, and response; ART, Assessment for Retreatment with TACE; ASAR, albumin-bilirubin grade; HAP, hepatoma arterial-embolization prognostic; IBS, integrated Brier score; LR, likelihood ratio; mHAP, modified hepatoma arterial-embolization prognostic; tumor size, alpha-fetoprotein, first TACE response; re-TACE, repeated transarterial chemoembolization; TACE, transarterial chemoembolization.

survival difference between candidates and noncandidates ($P = 0.001$) despite the fact that mHAP3 presented characteristics same as mHAP2.

The mHAP scoring system showed the best predictive performance (Harrell's C-index = 0.575; LR $\chi^2 = 13.91$) and calibration performance (IBS = 0.165) among the six models for the first TACE, the discrimination ability (C-index) of the model in this study was similar to the relevant research results [22]. The advantages of the mHAP score applied to this study cohort were (1) the review of previous original studies showed that bilirubin is the most unimportant variable for the HAP scoring system (multivariate analysis $P = 0.047$) [10]. During the establishment of the mHAP scoring system by Pinato etc., bilirubin was removed as it did not show any predictive value for survival difference and was considered as a weak variable in a robust algorithm, which in turn reduced the accuracy of the HAP model [11]. The various TACE scoring systems [12–17] modified the cutoff/weight of bilirubin or involved other factors such as Child-Pugh class or ALBI grade which include bilirubin to evade such influence. (2)
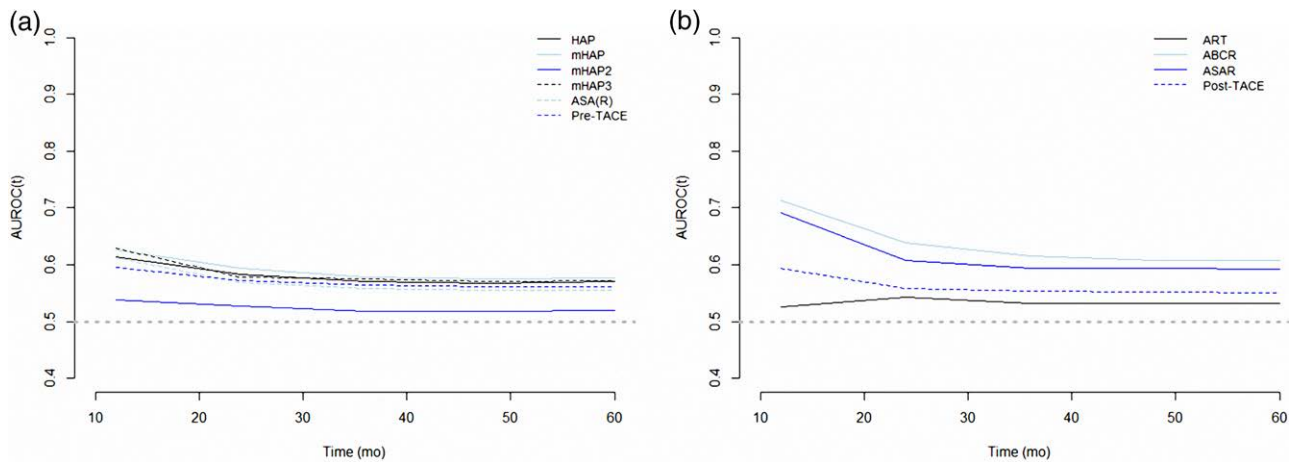
Fig. 4. (a,b) Time-dependent receiver operating characteristic (ROC) curves for comparisons among (a) clinical scoring system in the first TACE and (b) clinical scoring system in repeated TACE. AUROC: area under the receiver operating characteristic curve.
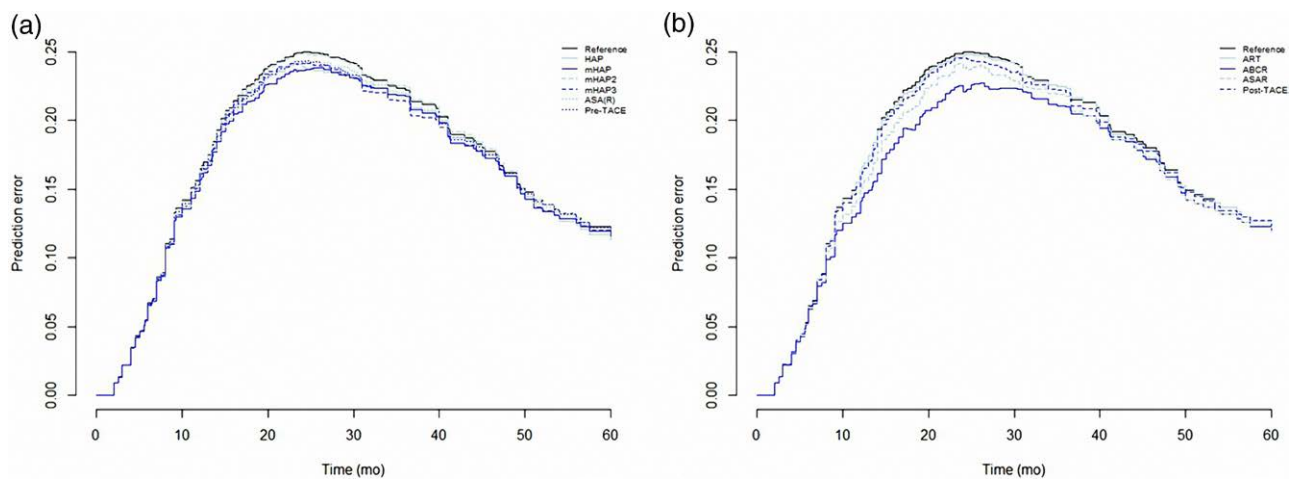

Fig. 5. Prediction error curve of (a) the first TACE clinical scoring system and (b) the repeated TACE clinical scoring system compared to the reference (Kaplan–Meier analysis).

The inclusion criteria of patients in the original study on mHAP were similar to this study [11], and 723 patients in the verification cohort were from Asia, which could also underlie the good performance of mHAP in the cohort of Asian HCC patients in HBV endemic areas.

Among the four scoring systems for subsequent TACE, only the ART scoring system was not verified ($P = 0.115$), which was in agreement with the results reported previously [23–27]. The disadvantages of the ART score applied to this study cohort were: (1) its inclusion standard required that the interval between the first and second TACE <90 days, which was not adopted as a criterion in this study. In clinical practice, the second session is usually performed 'on-demand' according to the tumor response and hepatic functional reserve after the first TACE [28,29], which prevents a subgroup of patients from receiving secondary TACE within 90 days; [23,24] (2) the diseases in most of the patients in this study were caused by HBV infection, and the AST level was mainly influenced by the activity of virus and responses to antiviral drugs, which were differed from the AST change pattern in ART cohorts characterized by alcoholic liver disease. Moreover, AST was not an important prognostic factor of liver

function. Therefore, it was not a constituent parameter of the Child-Pugh grade and model for end-stage liver disease. However, the characteristic 'AST increase >25%' has a high weight (4 points) in the ART scoring system, which influences the predictive performance in the cohort of this study. (3) ART scoring system utilized the radiological response criteria in the EASL guideline [30] but not the mRECIST criteria [21], whereas the mRECIST was accurate in evaluating the prognosis of patients with PR and SD [31]. The findings of this study also showed that the ART scoring system had no advantage in the consequent comparisons of performances (Harrell's C-index = 0.530; LR $\chi^2 = 2.41$; IBS = 0.171).

The ABCR model showed the best predictive performance (Harrell's C-index = 0.604; LR $\chi^2 = 27.01$) and calibration performance (IBS = 0.159) in the cohort of this study, the discrimination ability (C-index) of the model in this study was similar to the relevant research results [22,26]. The advantages of the ABCR score were (1) similar to ART, ABCR also included the characteristics of radiological response and altered Child-Pugh scores. Unlike the ART scoring system, ABCR incorporates AFP changes and BCLC stages instead of AST

increase. Several previous studies have reported that AFP changes following TACE sessions are closely associated with treatment efficiency; [32,33] (2) the consideration of inclusion criteria was the same as that of this study, the BCLC A, B and C stage patients in the original ABCR scoring system cohort were assigned the score of 0, 2 and 3, respectively, in the subsequently established prognosis model, which had critical effects in guiding the risk stratification in the cohort of this study. (3) The original study establishing the ABCR system also pointed out that the cohort consisted mainly of patients with viral liver diseases, and the model aimed to provide a convenient scoring system for HCC patients with higher percentages of viral liver diseases [15]. This also explicated why the ABCR scoring system is optimal for HCC patients with HBV as the primary cause.

Nevertheless, the present study has several limitations. First, as a single-center retrospective study, there could be severe selection bias and information bias in this study. Second, since TACE was the only treatment for the patients included, whether it is applicable to the current combination therapy needs to be explored further. Third, this study consisted of Chinese HCC patients, of which HBV was the primary cause, and thus, the findings need to be generalized and extrapolated cautiously. Finally, patients receiving only conventional TACE therapies were included in this study, and patients receiving DEB-TACE were not evaluated.

In summary, the predictive values of mHAP2 and ART scoring systems were limited and not applicable for the first and subsequent TACE selection in HCC patients, respectively. mHAP and ABCR scoring systems had the best predictive capabilities for the prognosis of HCC patients receiving TACE and acted as effective tools for clinical decisions. mHAP scoring system identified candidates who could benefit from the first TACE session, whereas the ABCR scoring system is suitable for prognosticating the patients who could benefit from the TACE retreatment.

Conflicts of interest

There are no conflicts of interest.

References

1 Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, *et al*. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; 136:E359–E586.

2 Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, *et al*. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; 66:115–132.

3 Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet* 2018; 391:1301–1314.

4 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.

5 Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J-L, *et al*. EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018; 69:182–236.

6 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.

7 Sangro B, Salem R. Transarterial chemoembolization and radioembolization. *Semin Liver Dis* 2014; 34:435–443.

8 Lencioni R, de Baere T, Soulen MC, Rilling WS, Geschwind JH. Lipiodol transarterial chemoembolization for hepatocellular carcinoma: a systematic review of efficacy and safety data. *Hepatology* 2016; 64:106–116.

9 Hucke F, Pinter M, Graziadei I, Bota S, Vogel W, Müller C, *et al*. How to STATE suitability and START transarterial chemoembolization in patients with intermediate stage hepatocellular carcinoma. *J Hepatol* 2014; 61:1287–1296.

10 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.

11 Pinato DJ, Arizumi T, Allara E, Jang JW, Smirne C, Kim YW, *et al*. Validation of the hepatoma arterial embolization prognostic score in European and Asian populations and proposed modification. *Clin Gastroenterol Hepatol* 2015; 13:1204–1208.e2.

12 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.

13 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.

14 Sieghart W, Hucke F, Pinter M, Graziadei I, Vogel W, Müller C, *et al*. The ART of decision making: retreatment with transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology* 2013; 57:2261–2273.

15 Adhoute X, Penaranda G, Naude S, Raoul JL, Perrier H, Bayle O, *et al*. Retreatment with TACE: the ABCR SCORE, an aid to the decision-making process. *J Hepatol* 2015; 62:855–862.

16 Nam JY, Choe AR, Sinn DH, Lee JH, Kim HY, Yu SJ, *et al*. A differential risk assessment and decision model for transarterial chemoembolization in hepatocellular carcinoma based on hepatic function. *BMC Cancer* 2020; 20:504.

17 Han G, Berhane S, Toyoda H, Bettinger D, Elshaarawy O, Chan AWH, *et al*. Prediction of survival among patients receiving transarterial chemoembolization for hepatocellular carcinoma: a response-based approach. *Hepatology* 2020; 72:198–212.

18 Bureau of Medical Administration, National Health Commission of the People's Republic of China. [Standardization for diagnosis and treatment of hepatocellular carcinoma (2022 edition)]. *Zhonghua Gan Zang Bing Za Zhi* 2022; 30:367–388. Chinese.

19 Clinical Guidelines Committee of Chinese Interventionalists College. [Chinese clinical practice guidelines for transarterial chemoembolization of hepatocellular carcinoma]. *Zhonghua Nei Ke Za Zhi* 2021; 60:599–614. Chinese.

20 Kim HY, Park JW, Joo J, Jung SJ, An S, Woo SM, *et al*. Severity and timing of progression predict refractoriness to transarterial chemoembolization in hepatocellular carcinoma. *J Gastroenterol Hepatol* 2012; 27:1051–1056.

21 Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis* 2010; 30:05252–05060.

22 Wang ZX, Wang EX, Bai W, Xia DD, Mu W, Li J, *et al*. Validation and evaluation of clinical prediction systems for first and repeated transarterial chemoembolization in unresectable hepatocellular carcinoma: a Chinese multicenter retrospective study *World J Gastroenterol* 2020; 26:657–669.

23 Kudo M, Arizumi T, Ueshima K. Assessment for retreatment (ART) score for repeated transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology* 2014; 59:2424–2425.

24 Arizumi T, Ueshima K, Iwanishi M, Minami T, China H, Kono M, *et al*. Evaluation of ART scores for repeated transarterial chemoembolization in Japanese patients with hepatocellular carcinoma. *Oncology (Huntingt)* 2015; 89:4–10.

25 Terzi E, Terenzi L, Venerandi L, Croci L, Renzulli M, Mosconi C, *et al*. The ART score is not effective to select patients for transarterial chemoembolization retreatment in an Italian series. *Dig Dis* 2014; 32:711–716.

26 Kloeckner R, Pitton MB, Dueber C, Schmidtmann I, Galle PR, Koch S, *et al*. Validation of clinical scoring systems ART and ABCR after transarterial chemoembolization of hepatocellular carcinoma. *J Vasc Interv Radiol* 2017; 28:94–102.

27   Tseng CL, Lai WJ, Huang CJ, Huang YH, Su CW, Lee IC, *et al*. The effectiveness of ART score in selecting patients for transarterial chemoembolization retreatment: a cohort study in Taiwan. *Medicine (Baltim)* 2015; 94:e1659.
28   Antoch G, Roelle G, Ladd SC, Kuehl H, Heusner TA, Sotiropoulos GC, *et al*. Selective and sequential transarterial chemoembolization: survival in patients with hepatocellular carcinoma. *Eur J Radiol* 2012; 81:2290–2297.
29   Terzi E, Golfieri R, Piscaglia F, Galassi M, Dazzi A, Leoni S, *et al*. Response rate and clinical outcome of HCC after first and repeated cTACE performed 'on demand'. *J Hepatol* 2012; 57:1258–1267.
30   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.
31   Shim JH, Lee HC, Kim SO, Shin YM, Kim KM, Lim Y-S, *et al*. Which response criteria best help predict survival of patients with hepatocellular carcinoma following chemoembolization? A validation study of old and new models. *Radiology* 2012; 262:708–718.
32   Lee YK, Kim SU, Kim DY, Ahn SH, Lee KH, Lee DY, *et al*. Prognostic value of α-fetoprotein and des-γ-carboxy prothrombin responses in patients with hepatocellular carcinoma treated with transarterial chemoembolization. *BMC Cancer* 2013; 13:5.
33   Memon K, Kulik L, Lewandowski RJ, Wang E, Ryu RK, Riaz A, *et al*. Alpha-fetoprotein response correlates with EASL response and survival in solitary hepatocellular carcinoma treated with transarterial therapies: a subgroup analysis. *J Hepatol* 2012; 56:1112–1120.