Transarterial chemoembolization for hepatocellular carcinoma: development and external validation of the Munich-TACE score

Mark op den Winkel[a], Dorothea Nagel[b], Philip op den Winkel[a], Jörg Trojan[e], Philipp M. Paprottka[c], Christian J. Steib[a], Laura Schmidt[a], Markus Göller[f], Petra Stieber[b], Peter Göhring[b], Andreas Herbst[b], Markus Rentsch[d], Enrico N. De Toni[a], Burkhard Göke[g], Alexander L. Gerbes[a] and Frank T. Kolligs[a,h]

Background Allocation of patients with hepatocellular carcinoma (HCC) to the adequate therapy is determined by both tumor burden and liver function. The Barcelona Clinic Liver Cancer (BCLC) staging system and therapeutic algorithm recommends transarterial chemoembolization (TACE) based on the best evidence available to patients with intermediate-stage HCC (BCLC-B). However, many centers also treat subgroups of patients outside these recommendations and with more advanced disease by TACE. The purpose of this study was to identify prognostic factors in a TACE cohort, including BCLC-B patients, as well as patients treated outside of BCLC-B, to test the prognostic capabilities of published staging systems and to optimize prognostication for TACE patients.

Patients and methods A cohort of 186 first-line TACE patients was analyzed. Independent prognostic factors were identified and used to construct the Munich-TACE score (M-TACE). M-TACE was tested against established staging systems (including BCLC and two recently published TACE-specific scores) and a ranking using concordance index and Akaike Information Criterion was performed. Finally, an external validation in an independent TACE cohort ($n = 71$) was conducted.

Results Bilirubin, Quick/international normalized ratio, C-reactive protein, creatinine, α-feto protein, and tumor extension were identified as independent prognostic factors and used to construct M-TACE. M-TACE identifies three distinct subgroups ($P < 0.0001$) with median survival times of 35.2, 16.9, and 8.6 months, respectively. Compared with established staging systems, M-TACE showed the best prognostic capabilities in both cohorts of patients (cohort 1: *c*-index, 0.71; Akaike Information Criterion: 1276; cohort 2: *c*-index, 0.754).

Conclusion We identified independent risk factors for patients treated with TACE. The newly constructed M-TACE score is superior to established staging systems and might prove helpful to identify patients who are most suitable for TACE.

Eur J Gastroenterol Hepatol 30:44–53

Introduction

Hepatocellular carcinoma (HCC) is the sixth most frequent cancer entity worldwide. The course of the disease often is complicated by underlying liver cirrhosis [1]. Transarterial chemoembolization (TACE) is an established treatment method in the palliative setting with solid evidence for improvement of survival [2]. However, conflicting data on the efficacy of TACE indicate that selecting the right patients is a critical issue. HCC staging systems taking into account both liver function and tumor parameters can be a valuable tool for treatment decisions. The Barcelona Clinic Liver Cancer (BCLC) [3], the staging system endorsed by the major liver associations [4,5], only allocates patients in the intermediate stage B (asymptomatic large or multinodular tumors without portal invasion or extrahepatic spread, Child A-B) to TACE. Conversely, in real-world clinical practice, this widely used modality is not limited to stage B patients [6]. Recent studies have demonstrated inferior prognostic abilities of BCLC for TACE patients when compared with other staging systems [7]. As all established staging systems have been developed in inhomogeneous HCC populations, it remains unclear whether they consider the relevant prognostic factors for TACE patients at all. In line with the call for treatment-specific staging systems [8,9], two pretherapeutic TACE-specific scores have been proposed recently [10,11]. A major aim of the current study was to apply the available staging systems to a homogeneous

European Journal of Gastroenterology & Hepatology 2018, 30:44–53

Keywords: Barcelona Clinic Liver Cancer, Cancer of the Liver Italian Program, Hepatoma Arterial embolization Prognostic, hepatocellular carcinoma staging systems, Selection for Transarterial Chemoembolization Treatment

[a]Department of Medicine II, [b]Institute of Laboratory Medicine, [c]Department of Clinical Radiology, [d]Department of General, Visceral, Transplantation, Vascular and Thoracic Surgery, University Hospital, Ludwig Maximilians University, Munich, [e]Department of Internal Medicine I, Johann-Wolfgang-Goethe-University, Frankfurt am Main, [f]Department of Internal Medicine II, University of Erlangen, Erlangen, [g]University Medical Center Hamburg-Eppendorf, Hamburg and [h]Department of Internal Medicine and Gastroenterology, Helios-Klinikum Berlin-Buch, Berlin, Germany

Correspondence to Mark op den Winkel, MD, Department of Medicine II, Liver Center Munich, University Hospital, Ludwig Maximilians University, Marchioninistrasse 15, 81377 Munich, Germany

Tel: +49 89 44000; fax: +49 89 4400 78829;
e-mail: mark.op.den.winkel@med.uni-muenchen.de

Received 29 June 2017 **Accepted** 24 August 2017

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, *www.eurojgh.com*.

TACE cohort and compare their prognostic quality. The identification of prognostic factors should then be used to create a new, potentially more accurate prognostic score for patients evaluated for TACE; providing a tool for identifying those patients most likely to benefit from this treatment.

Patients and methods

Patients

In this retrospective study, we identified patients with HCC between December 1999 and March 2011 (including patients from a previous study [12]), who were treated with conventional TACE at the University Hospital of Munich ($n = 234$). The ethics committees of Munich and Frankfurt (external cohort) had approved the study. Patients who could not be classified in all staging systems owing to missing clinical or radiological parameters were excluded ($n = 21$). Non-first-line TACE patients and those receiving TACE as a mode of 'bridging' before liver transplantation were excluded ($n = 27$). The 186 patients remaining formed the study population.

Data collection

Aiming at creating clinically practical tumor extension criteria for TACE patients, the tumor extension criteria defined by BCLC were modified: the term 'portal invasion' was replaced by 'vascular involvement' (defined by radiological signs of direct tumor invasion of any hepatic vessel or by portal vein thrombosis with or without signs of direct tumor invasion); in addition, enlarged local lymph nodes as a sign of extrahepatic disease were ignored, as this parameter is considered an unspecific radiological finding in the context of chronic liver disease. The external cohort consisted of 71 TACE patients of the University Hospital Frankfurt (2005–2012). Similar to the University Hospital of Munich, this liver center has a longstanding experience in HCC treatment including a liver transplant program.

Tested staging systems

Five multidimensional [BCLC [3], Cancer-of-the-Liver-Italian-Program (CLIP) [13], Groupe-d'Etude-et-de-Traitementdu-Carcinome-Hepatocellulaire (GETCH) [14], Japan Integrated Staging (JIS) [15], and Okuda [16]] and two unidimensional scores (TNM [17] and Child–Pugh [18]) were tested. In addition, two recently published TACE-specific scores were applied: Selection for Transarterial Chemoembolization Treatment (STATE) score [10] and Hepatoma Arterial embolization Prognostic (HAP) score [11].

Transarterial chemoembolization

Indication for TACE was based upon interdisciplinary discussion. In general, the TACE procedure was conducted as follows: before each TACE (< 4 weeks), a contrast-enhanced computed tomography for intervention planning purposes was performed. The TACE was carried out as selectively as possible aiming at segmental or subsegmental feeder arteries and dosage of the TACE components (epirubicin, lipiodol, gelatin sponge) was adjusted to tumor size, vascularization, and liver function. Sequential TACE therapies were performed in patients without contraindications, if signs of tumor viability on follow-up imaging persisted.

Statistical analysis

For statistical analysis, SAS software (SAS, version 9.2; SAS Institute Inc., Cary, North Carolina, USA) was used. *P* values less than 0.05 indicated statistical significance and *P* values less than 0.0001 indicated high statistical significance.

Univariate analysis

Overall survival (OS) was defined as time from the date of first TACE until death or final follow-up. In univariate analysis, OS was estimated by using the Kaplan–Meier method. The log-rank test was used to compare the survival curves. In addition, the *P*-value medians of survival time and 95% confidence intervals (95% CI) for the different strata are given. Both single parameters and the whole scores were analyzed concerning their prognostic significance. For Kaplan–Meier analysis of continuous variables, one or more cutoff values are necessary; therefore, laboratory values were divided into quartiles.

Multivariate analysis and Munich-TACE construction

Those parameters with significance in univariate analysis underwent multivariate testing by using the Cox proportional hazards regression model. Creatinine and Quick were categorized because Kaplan–Meier analysis of their quartiles did not indicate linearity. All other laboratory values were taken as base two logarithms and used as continuous variables. Clinical and tumor parameters were modified by joining together those categories that lacked significant distinction in univariate analysis. As an example, the tumor extension criteria were reduced from four to two categories (category A: singular nodule < 2 cm, three nodules ≤ 3 cm, one nodule ≤ 5 cm; category B: large or multilocular, vascular involvement, M1). The significant variables and interactions between them were tested simultaneously and a model was created using backward selection. Parameters with significant prognostic meaning in multivariate analysis should serve as the components of the new prognostic score. In order to keep the score as simple as possible, a second, less complex Cox model was calculated on the basis of the first one. For this, final model interactions were ignored and continuous variables were categorized. Variables no longer significant in this setting were excluded. Internal validation of the model was performed by calculating the same model for 200 bootstrap samples and determining the concordance index (*c*-index) [19] for each sample. The mean value of these 200 *c*-indices was used as an estimation, how good the new model would be applicable to a new collective. In addition, for graphic illustration, the cohort was split into tertiles by applying the new model, and in the next step Kaplan–Meier curves for the tertiles and the curves estimated from the Cox model were plotted. For construction of the score, the specific number of points for each category was estimated on the basis of the coefficient of the Cox model. The final arrangement of the Munich-TACE score (M-TACE) into its stages was achieved by formation of tertiles.

Ranking

Ranking of staging systems was conducted by using the *c*-index [19] and the Akaike Information Criterion (AIC) [20] derived from the Cox model. *C*-indices together with 95% CI were calculated using the SAS macro (SAS Institute Inc.). As only the *c*-index is capable of directly comparing the performance of a prognostic model in different populations, only the *c*-index was used in external validation.

Results

Baseline patient characteristics

Baseline patient characteristics of the internal cohort are shown in Table 1. The leading etiological factor was alcohol abuse (45.7%) (Table S1, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*). The majority of patients were male (80.6%), and the median age was 65.5 (range: 31.5–88.8) years. Liver cirrhosis was present in 83.9%. Partial or total portal vein thrombosis was seen in 11.8 and 1.6%. Liver function was compensated (no cirrhosis or Child A) in 62.3%; only 9.1% had end-stage liver disease (Child C). Almost all of the patients (97.9%) were in a good or fairly good general condition [Eastern Cooperative Oncology Group (ECOG) 0–1]. In all, 39.7% showed more than three nodules. Vascular invasion and distant metastasis were present in the minority of cases (15.1 and 9.2%). The characteristics of the most relevant baseline laboratory parameters are summarized in Table S2 (Supplemental digital content 1, *http://links.lww.com/EJGH/A236*). The median number of TACE interventions was 3. The majority of patients received up to five (72.6%) TACE interventions; 21.5% were treated only once. In all, 9.2% of the patients underwent the TACE procedure more than 10 (total range: 1–25) times (Table S3, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*). The most frequent reason for cessation of TACE therapy was untreatable tumor progression (29%) (Table S4, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*).

Survival analysis and prognostic factors

By the end of follow-up, 155/186 (83.3%) of all patients had died. Median follow-up duration for patients alive was 34.9 (range: 3.5–80.6) months, and median follow-up for all patients as estimated by the reverse Kaplan–Meier method was 53.7 months. Overall median survival was 16.9 (95% CI: 14.4–20.6) months (Fig. 1a). The 1-, 3-, and 5-year OS rates were 65.1, 22.5, and 9.7%, respectively.

Univariate analysis

A total of 17 parameters showed a significant influence on survival in univariate analysis: tumor-related (Table 1) – tumor extension ($P = 0.026$), number of tumor nodes ($P = 0.002$), up-to-seven criteria ($P = 0.016$), distant metastasis ($P = 0.035$), and tumor burden of at least 50% ($P = 0.011$); clinical (Table 1) – ascites ($P = 0.031$) and age ($P = 0.025$); laboratory (Table S5, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*) – α-feto protein (AFP) ($P < 0.0001$), bilirubin ($P = 0.003$), alkaline phosphatase ($P < 0.0001$), glutamic oxaloacetic transaminase ($P < 0.0001$), glutamic pyruvic transaminase ($P = 0.006$), Quick ($P = 0.0117$), albumin ($P = 0.028$), C-reactive protein (CRP) ($P < 0.0001$), creatinine ($P = 0.0154$), and lactate dehydrogenase ($P = 0.005$).

Table 1. Baseline parameter, univariate analysis, Munich-TACE cohort

Parameters	*n* (%)	Median survival (95% CI) (months)	*P* value
Sex			0.632
Female	36 (19.4)	15.2 (12.7–28.3)	
Male	150 (80.6)	16.9 (14.3–21.0)	
Age (median: 65.5) (years)			0.025
< 60	46 (24.7)	13.8 (7.2–16.9)	
60–65	47 (25.3)	16.5 (11.5–25.3)	
66–71	47 (25.3)	26.8 (18.3–35.0)	
≥ 72	46 (24.7)	14.4 (9.2–20.6)	
Etiology (three most frequent)			0.826
Alcohol	85 (45.7)	20.3 (15.2–22.5)	
HCV	36 (19.4)	14.6 (10.8–27.6)	
Cryptogenic	23 (12.4)	16.4 (9.4–26.8)	
Liver cirrhosis			0.125
No	30 (16.1)	28.3 (12.7–35.2)	
Yes	156 (83.9)	15.8 (13.8–18.3)	
Ascites			0.031
No	134 (72.0)	20.1 (15.8–25.3)	
Moderate	36 (19.4)	10.6 (6.9–20.3)	
Severe	16 (8.6)	5.8 (2.7–21.7)	
Hepatic encephalopathy			0.071
No	153 (82.3)	18.6 (15.6–22.5)	
Moderate	33 (17.7)	7.3 (5.2–13.7)	
ECOG			0.261
0	117 (62.8)	18.6 (14.6–25.3)	
1	65 (35.0)	15.2 (9.2–20.3)	
2	4 (2.2)	15.1 (2.7–26.0)	
Portal hypertension			0.107
No	65 (36.1)	23.8 (14.4–29.7)	
Yes	115 (63.9)	16.4 (14.3–20.3)	
Portal vein thrombosis			0.695
No	161 (86.6)	17.1 (14.6–21.3)	
Yes	25 (13.4)	14.3 (6.0–30.1)	
Number of tumor nodes			0.002
1	52 (27.9)	23.8 (14.4–32.5)	
2	36 (19.4)	22.9 (12.8–35.0)	
3	24 (12.9)	18.2 (3.9–25.6)	
> 3	74 (39.8)	14.3 (11.5–16.8)	
Tumor extension[a] (cm)			0.026
Singular <2	5 (2.7)	14.6 (2.2–)	
Three nod ≤ 3, one nod ≤ 5	22 (11.8)	34.6 (10.8–85.5)	
Large or multilocular	115 (61.8)	17.1 (14.4–21.0)	
Vascular involvement, M1	44 (23.7)	13.8 (6.9–17.4)	
Tumor burden			0.011
≤ 50%	170 (91.3)	17.1 (14.6–21.3)	
> 50%	16 (8.7)	13.7 (3.4–22.5)	
Up-to-seven criteria[b]			0.016
In	56 (30.3)	28.0 (16.0–34.6)	
Out	129 (69.7)	15.6 (13.7–18.6)	
Tumor size > 7 cm			0.947
No	127 (69.0)	17.1 (14.4–21.3)	
Yes	57 (31.0)	16.7 (11.4–23.8)	
Lymph nodes ≥ 1 cm			0.678
No	132 (70.9)	16.1 (14.2–20.6)	
Yes	54 (29.1)	20.3 (13.7–24.0)	
Vascular invasion			0.908
No	158 (84.9)	16.9 (14.4–21.3)	
Yes	28 (15.1)	15.2 (6.9–28.3)	
Distant metastasis			0.035
No	169 (90.8)	18.1 (15.2–21.7)	
Yes	17 (9.2)	11.6 (3.9–16.5)	

HCV, hepatitis C virus; CI, confidence interval; EOCG, Eastern Cooperative Oncology Group; nod, nodule.
[a]Modified from BCLC (Forner *et al.* [3]). [b]Sum of the size of the largest tumor and the total number of tumors.

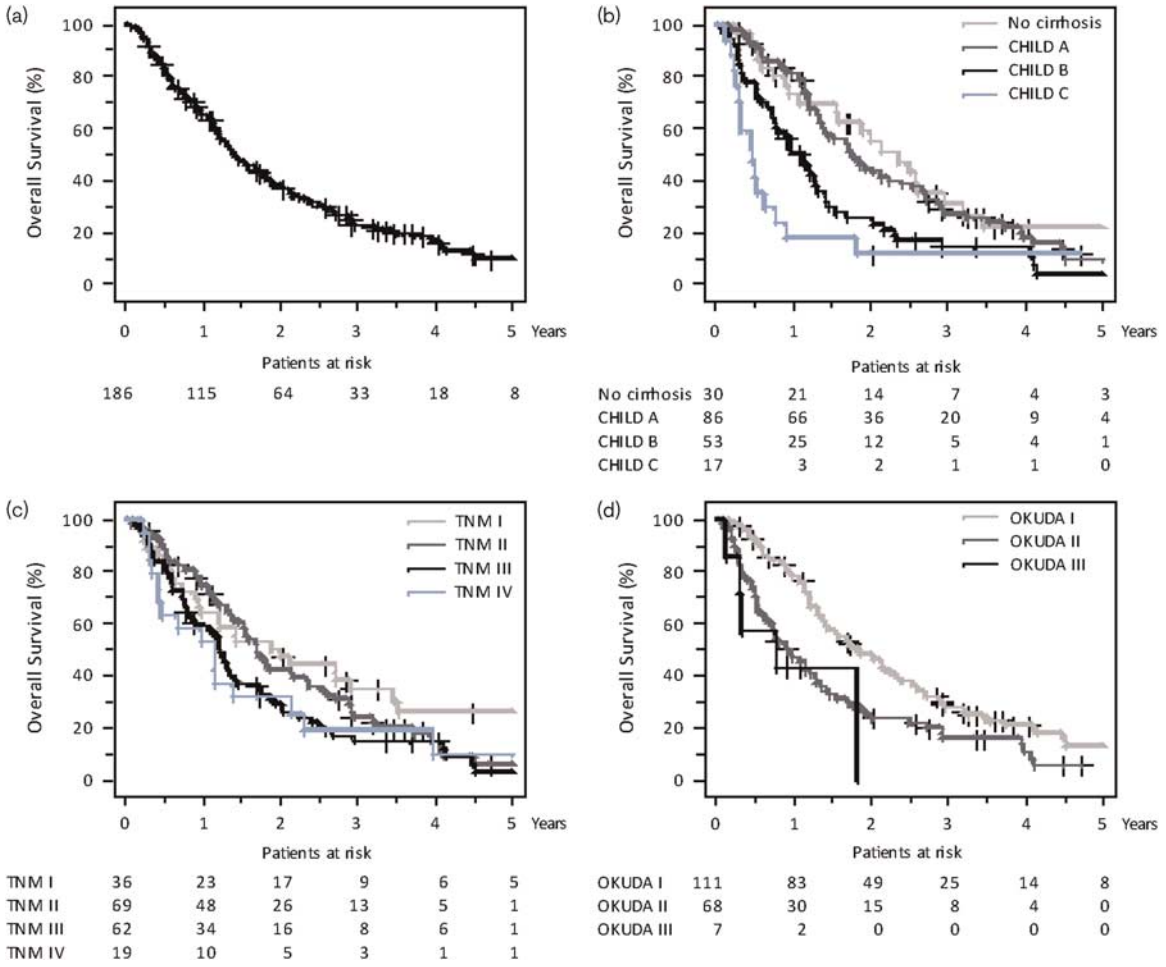

Fig. 1. Overall survival internal transarterial chemoembolization (TACE) cohort (a); Kaplan–Meier survival curves for Child–Pugh (b), TNM (c), Okuda (d).

Staging systems

Stratification of patients and estimated median survival time according to the staging systems are depicted in Table 2. Stratification according to BCLC revealed that 109 of the 186 patients received TACE outside of stage B, especially stage C was represented strongly (41.9%). Apart from TNM ($P = 0.066$), all staging systems possessed a significant association with prognosis. Figures 1 and 2 show the Kaplan–Meier survival analysis stratified according to the staging systems. In the next step, the discriminatory quality of the staging systems was examined. All of the strata in GETCH and STATE characterized distinct survival groups. All other scores had at least two adjacent stages lacking distinct survival times. Stages B and C of BCLC did not have a distinct survival ($P = 0.95$) (Table 2). Thirty-five BCLC-C patients were assigned to this more advanced stage and not to stage B solely on the basis of their ECOG 1 status (Table S6, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*). When including these patients in BCLC-B, there was a slight tendency toward separation of the BCLC-B and BCLC-C survival times (Fig. S1, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*); however, this remained nonsignificant ($P = 0.657$).

Multivariate analysis

In the final prognostic model, five laboratory parameters and one radiological parameter [tumor extension; $P = 0.008$; hazard ratio (HR): 2.077] remained significant predictors of survival (Table S7, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*). Of the laboratory parameters, one was tumor related (AFP: $P < 0.0001$; HR: 3.618), two were liver function related (Quick: $P = 0.0124$; HR: 1.617; bilirubin: $P = 0.0001$; HR: 3.144), and two were neither directly tumor related nor liver related (creatinine: $P = 0.0207$; HR: 1.657; CRP: $P < 0.0001$; HR: 3.349). The nearly congruent course of the Kaplan–Meier and Cox regressions curves in each of the three stages of the prognostic model based on these parameters (Fig. S2, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*), as well as the high *c*-index of 0.736, demonstrated an excellent applicability on the internal TACE collective (internal validation).

Construction of Munich-TACE

On the basis of the six parameters of the final model (Table S7, Supplemental digital content 1, *http://links.lww.com/EJGH/A236*), the M-TACE score was constructed. Table 3 shows the allocation of score points subject to a patient's exact laboratory values and tumor extension, respectively.

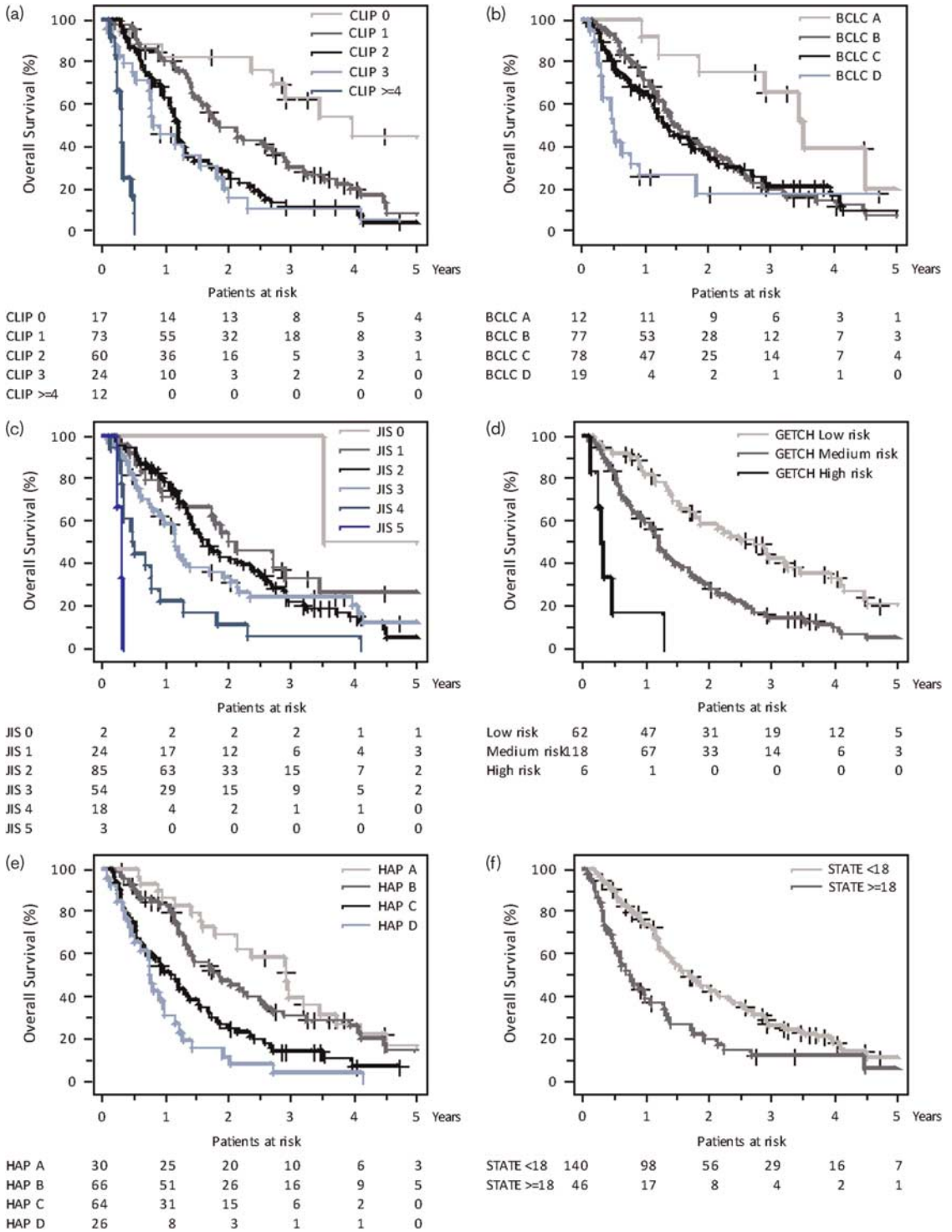

Fig. 2. Kaplan1Meier survival curves for CLIP (a), BCLC (b), JIS (c), GETCH (d), HAP (e), STATE (f). BCLC, Barcelona-Clinic-Liver-Cancer; CLIP, Cancer-of-the-Liver-Italian-Program; GETCH, Groupe-d'Etude-et-de-Traitement-du-Carcinome-Hepatocellulaire; HAP, Hepatoma-Arterial-embolization-Prognostic; JIS, Japan Integrated Staging; STATE, Selection-for-Transarterial-Chemoembolisation-Treatment.

Patients in the range of 0–9 points are allocated to stage I (low mortality risk group); stage II (intermediate mortality risk group) is defined by the range of 10–13; and stage III patients (high mortality risk group) have 14–26 points. In particular, high CRP (≥2.0 mg/dl), AFP (≥1000 ng/ml), and bilirubin values (≥3.1 mg/dl) have a high impact on mortality risk, each resulting in the allocation of 6 points. Three points are given, whenever these parameters do not reach the above-mentioned cutoff value, but still are in the pathological range. An advanced tumor extension (category B) is accounted for with 4 points, whereas deranged values of creatinine and Quick have a comparatively low weighting (2 points). Of note, substituting Quick with international normalized ratio (INR) [similar HR of 1.617 (95% CI: 1.109–2.357) vs. 1.59 (1.053–2.279)] did not result in changes of M-TACE stratification. When applied

Table 2. Univariate analysis of the staging systems

Scores	*n* (%)	Median survival months (95% CI)	*P* value
Child–Pugh	186		< 0.0001
No cirrhosis	30 (16.1)	28.3 (12.7–35.2)	–
A	86 (46.2)	21.0 (16.4–28.8)	nc vs. A: 0.62
B	53 (28.5)	13.5 (8.8–15.6)	A vs. B: 0.003
C	17 (9.1)	5.5 (3.0–9.2)	B vs. C: 0.092
TNM	186		0.066
I	36 (19.4)	23.1 (11.2–41.3)	–
II	69 (37.1)	20.3 (16.4–28.0)	I vs. II: 0.255
III	62 (33.3)	14.4 (9.8–17.1)	II vs. III: 0.101
IV	19 (10.2)	13.7 (4.9–25.6)	III vs. IV: 0.875
Okuda	186		0.001
I	111 (59.7)	22.2 (16.8–27.6)	–
II	68 (36.6)	10.8 (7.3–15.2)	I vs. II: 0.0006
III	7 (3.8)	9.2 (1.3–21.7)	II vs. III: 0.539
BCLC	186		0.001
A	12 (6.5)	42.0 (14.6–85.5)	–
B	77 (41.4)	18.1 (14.4–24.2)	A vs. B: 0.005
C	78 (41.9)	16.0 (13.7–21.3)	B vs. C: 0.945
D	19 (10.2)	5.9 (3.4–10.8)	C vs. D: 0.029
CLIP	186		< 0.0001
0	17 (9.1)	47.5 (28.3–63.8)	–
1	73 (39.2)	22.3 (18.1–32.2)	0 vs. 1: 0.016
2	60 (32.3)	14.3 (11.6–15.6)	1 vs. 2: 0.001
3	24 (12.9)	10.1 (6.1–18.3)	2 vs. 3: 0.406
≥ 4	12 (6.5)	3.4 (2.3–5.2)	3 vs. ≥ 4: < 0.0001
JIS	186		< 0.0001
0	2 (1.0)	– (42.0–)	–
1	24 (12.9)	24.5 (11.2–41.3)	0 vs. 1: 0.299
2	85 (45.7)	20.1 (16.8–26.8)	1 vs. 2: 0.189
3	54 (29.0)	13.7 (10.5–20.3)	2 vs. 3: 0.357
4	18 (9.7)	5.7 (3.6–9.2)	3 vs. 4: 0.004
5	3 (1.6)	3.4 (2.7–3.8)	4 vs. 5: 0.011
GETCH	186		< 0.0001
Low risk	62 (33.3)	32.5 (21.3–41.3)	–
Intermediate risk	118 (63.4)	14.4 (11.6–16.8)	L vs. I: < 0.0001
High risk	6 (3.2)	3.4 (1.3–15.2)	I vs. H: < 0.0001
HAP	186		< 0.0001
A	30 (35.0)	35.0 (21.3–41.3)	–
B	66 (22.2)	22.2 (16.0–30.1)	A vs. B: 0.208
C	64 (12.8)	12.8 (7.9–16.5)	B vs. C: 0.002
D	26 (9.0)	9.0 (5.2–11.6)	C vs. D: 0.063
STATE	186		< 0.0001
≥ 18	140 (75.3)	20.3 (16.4–25.3)	
< 18	46 (24.7)	9.0 (5.5–15.2)	
M-TACE	186		< 0.0001
I	55 (29.6)	35.2 (32.4–53.9)	–
II	70 (37.6)	16.9 (13.7–21.0)	I vs. II: < 0.0001
III	61 (32.8)	8.6 (5.9–11.6)	II vs. III: < 0.0001
M-TACE (external)	71		< 0.0001
I	38 (53.5)	68.1 (30.9–68.1)	
II	22 (31.0)	12.9 (9.3–16.9)	I vs. II: < 0.0001
III	11 (15.5)	5.6 (3.0–13.2)	II vs. III: < 0.015
Child–Pugh (external)	71		0.003
A + no cirrhosis	45 (63.4)	40.0 (16.9–68.1)	
B	24 (33.8)	9.4 (5.0–26.8) (B + C)	
C	2 (2.8)		
CLIP (external)	71		< 0.0001
0	6 (8.5)	– (30.9–)	–
1	34 (47.9)	68.1 (13.5–68.1)	0 vs. 1: 0.206
2	21 (29.6)	16.9 (9.3–40.0)	1 vs. 2: 0.118
3	6 (8.5)	6.4 (1.9–)	2 vs. 3: 0.136
≥ 4	4 (5.6)	4.0 (1.5–13.2)	3 vs. ≥ 4: 0.397

Staging systems were applied to the internal cohort. In addition, M-TACE, CLIP, and Child–Pugh were applied to the external cohort as well.
BCLC, Barcelona-Clinic-Liver-Cancer; CI, confidence interval; CLIP, Cancer-of-the-Liver-Italian-Program; HAP, Hepatoma-Arterial-embolization-Prognostic; JIS, Japan Integrated Staging; M-TACE, Munich-TACE; nc, no cirrhosis; STATE, Selection-for-Transarterial-Chemoembolisation-Treatment; TACE, transarterial chemoembolization.

Table 3. M-TACE score

	Points				
	0	2	3	4	6
AFP (ng/ml)	< 35	–	35–999	–	≥ 1000
Bilirubin (mg/dl)	< 1.1	–	1.1–3.0	–	≥ 3.1
CRP (mg/dl)	< 0.5	–	0.5–1.9	–	≥ 2
Tumor extension[a]	Category A	–	–	Category B	–
Creatinine (mg/dl)	< 1.3	≥ 1.3	–	–	–
Quick (%)	≥ 75	< 75	–	–	–
Stage I (low mortality risk): 0–9 points					
Stage II (intermediate mortality risk): 10–13 points					
Stage III (high mortality risk): 14–26 points					

AFP, α-feto protein; CRP, C-reactive protein; INR, international normalized ratio.
[a]Tumor extension category B: positive if one of the following criteria is met: large (one nodule > 5 cm) or multilocular (exceeding the limits of three nodules ≤ 3 cm) or vascular involvement or M1. Otherwise category A. For further external use, the parameter Quick can be replaced by INR (>1.2 vs. ≤ 1.2).

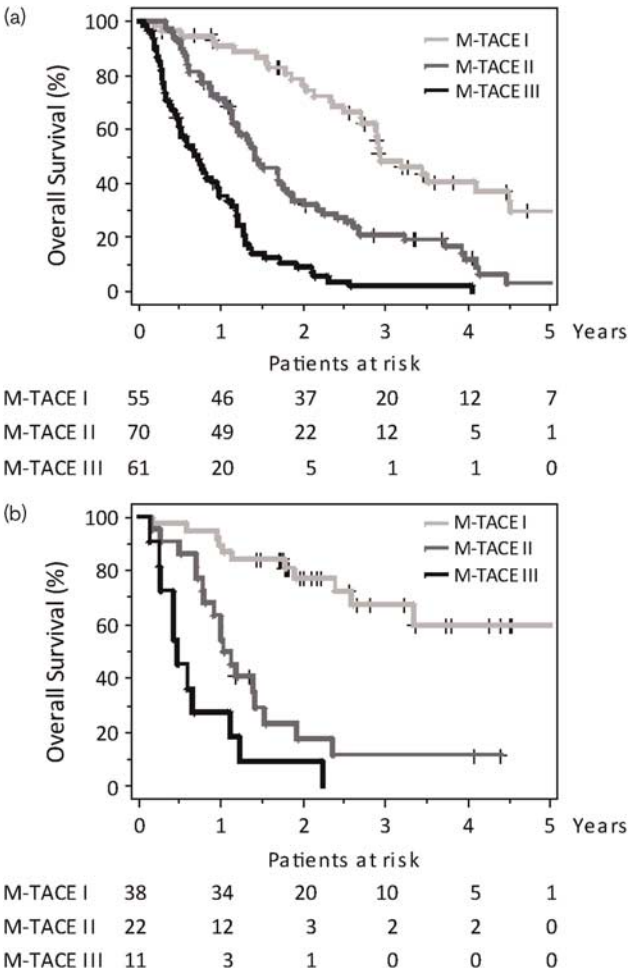

Fig. 3. M-TACE applied to internal TACE cohort. Stage I versus II: $P < 0.0001$; stage II versus III: $P < 0.0001$ (a). M-TACE applied to external TACE cohort (external validation); stage I versus II: $P < 0.0001$; stage II versus III: $P = 0.015$ (b). M-TACE, Munich-TACE; TACE, transarterial chemoembolization.

to the 186 TACE patients, M-TACE yielded three Kaplan–Meier survival curves with distinct survival (each $P < 0.0001$) (Fig. 3a). Patients in stage I ($n = 55$) had a median survival of 35.2 (95% CI: 32.4–53.9) months, stage II ($n = 70$) patients of 16.9 (95% CI: 13.7–21) months, and stage III patients ($n = 61$) had a low median survival of 8.6 (95% CI: 5.9–11.6) months.

Ranking

To distinguish the staging systems' prognostic ability, AIC and *c*-index were calculated. Table 4 depicts the ranking results on the basis of these two tests. Regarding the established staging systems, CLIP (AIC: 1305; *c*-index: 0.677) was identified as the best score, followed by the TACE-specific HAP score (AIC: 1322; *c*-index: 0.654). Of note, BCLC had inferior prognostic qualities (AIC: 1345; *c*-index: 0.597). M-TACE outperformed all staging systems in terms of AIC (1247) and *c*-index (0.71). As a sign of consistency, the top three scores took the same rank in either of the tests, as did the bottom-ranked TNM.

External validation

The major characteristics of the external TACE patient population ($n = 71$) are shown in Tables S8 and S9 (Supplemental digital content 1, *http://links.lww.com/EJGH/A236*). TACE components used were mitomycin C, lipiodol, and degradable starch microspheres. Age and sex distribution was similar as was the ECOG status. However, viral hepatitis (43.7%) and not alcohol (28.2%) was the leading etiological factor. There was a clear tendency toward less advanced tumors: only 63.4% (vs. 85.4%) had the higher tumor extension category B (large or multilocular tumors, vascular infiltration, or distant metastases). Only 22.5% had a tumor size of more than 7 cm compared with 31% in the internal cohort. Moreover, less Child C patients were documented (2.8 vs. 9.2%). As opposed to the internal cohort, the majority of patients were assigned to M-TACE stage I (53.3 vs. 29.6%). Consequentially, the median OS was longer (22.9 vs. 16.9 months). M-TACE stages I, II, and III had a median survival of 68.1, 12.9, and 5.6 months, respectively (Fig. 3b and Table 2). The good discriminatory ability of M-TACE was underlined by the significantly different survival times between the three stages (stage I vs. II, $P < 0.0001$; stage II vs. III, $P = 0.015$). As a whole, M-TACE revealed a highly significant meaning for survival ($P < 0.0001$). CLIP, the established system with the best performance in the internal cohort, and Child–Pugh were tested in the external patient population for comparative reasons. Although both scores showed an overall good performance ($P < 0.0001$ and $P = 0.002$, respectively), their discriminatory ability was not as good as that of M-TACE, with nonsignificant survival differences between all CLIP stages and between Child B and C (Table 2). Ranking according to *c*-index confirmed these results: 0.754 (M-TACE) versus 0.689 (CLIP) and 0.637 (Child–Pugh).

Table 4. Internal TACE cohort: ranking of seven established HCC staging systems and three new TACE-specific scores according to *c*-index and AIC

Ranks	Score	*c*-Index (95% CI)	AIC (AIC rank)
1	M-TACE[a]	0.710 (0.673–0.748)	1276 (1)
2	CLIP	0.677 (0.628–0.725)	1305 (2)
3	HAP[a]	0.652 (0.612–0.693)	1323 (3)
4	CHILD	0.625 (0.577–0.673)	1338 (6)
5	JIS	0.624 (0.574–0.674)	1334 (5)
6	GETCH	0.616 (0.574–0.658)	1326 (4)
7	OKUDA	0.600 (0.558–0.642)	1341 (7)
8	BCLC	0.597 (0.548–0.645)	1345 (9)
9	STATE[a]	0.584 (0.545–0.623)	1341 (8)
10	TNM	0.564 (0.512–0.616)	1348 (10)

In cases of discrepancy between *c*-index and AIC, the *c*-index was favored.
AIC, Akaike Information Criterion; BCLC, Barcelona-Clinic-Liver-Cancer; CI, confidence interval; CLIP, Cancer-of-the-Liver-Italian-Program; HAP, Hepatoma-Arterial-embolization-Prognostic; HCC, hepatocellular carcinoma; JIS, Japan Integrated Staging; M-TACE, Munich-TACE; STATE, Selection-for-Transarterial-Chemoembolisation-Treatment; TACE, transarterial chemoembolization; TNM, tumor, node and metastasis.
[a]TACE-specific scores.

Discussion

Patient cohort and transarterial chemoembolization characteristics

The study identified M-TACE as the superior prognostic score for TACE patients when compared with all relevant staging systems. The major advantage of the study was the homogeneous treatment collective M-TACE was developed in, exclusively considering first-line conventional TACE patients. Even the recently published TACE-specific scores were developed in less homogeneous (and smaller) cohorts, including different techniques [51% bland embolization (HAP) and 43% DEB-TACE (STATE)] and 26% non-first-line TACE patients (STATE), respectively [10,11]. Patient selection in the STATE study was on the basis of BCLC, including only stage A and B patients (+ECOG 1) and excluding 61 BCLC-C patients, although the prognostic accurateness of BCLC in nonsurgical patients has repeatedly been challenged [7,21] and our data show that BCLC-B and BCLC-C TACE patients do not show distinct survival. In terms of baseline parameters, the main results of our study are comparable to those of other western TACE studies [22]. OS was 16.9 months, compared with 15.0 in the HAP score cohort and 14.7 months in the STATE cohort. The rather long OS in our external cohort (22.9 months) could be explained by the high percentage of patients with comparatively good liver function and less advanced tumors. Another TACE study from Japan reported a median OS of 34 months in patients with even better liver function [23], an observation highlighting the impact of patient selection on survival. The TACE patients of the external cohort (2005–2012) were mainly treated in the era of Sorafenib, offering an option for systemic treatment after TACE failure [1]. Whether this factor contributed to the survival difference compared with the internal cohort (1999–2011) could not be assessed in this study.

The median number of TACE interventions in the internal cohort was 3 (range: 1–25); these data are well in line with other TACE studies, especially the one from Hucke and colleagues (identical median of three TACE interventions in both Austrian cohorts). The percentage of patients receiving only a single TACE intervention was 21.5 in the current study, compared with 14 and 18% in the Austrian cohorts [10]. These data reflect the general recommendation to repeat TACE if possible, to achieve the maximum benefit [22].

Prognostic parameters

Six independent prognostic parameters were identified: AFP is known to correlate with tumor extent and has frequently been acknowledged as a risk factor for TACE patients [23]. The observation in our study that values

greater than or equal to 1000 ng/ml are especially associated with a poor prognosis in TACE patients, is supported by earlier studies [24]. With a comparable HR, no difference in M-TACE patient stratification was observed when using INR or Quick, respectively. For future external and prospective validation, the use of INR to avoid interlaboratory variation is justified. Bilirubin is incorporated in many HCC staging systems [3,14,16], and studies focusing on TACE confirmed its prognostic significance in this subgroup [25]. A large study on prognostic factors in nonsurgical HCC patients even found bilirubin to be the most important prognostic factor. The authors concluded that liver function is more important for prognosis than tumor characteristics [26]; this is further supported by the fact that in our study the unidimensional liver function score Child–Pugh worked much better than did the unidimensional tumor classification TNM. Nevertheless, all multidimensional HCC staging systems consider some sort of radiological tumor characteristics, and the inclusion of the tumor extension criteria within M-TACE is in line with this observation. Of the six significant factors, CRP and creatinine stand out, because neither of them has been incorporated in any of the established HCC staging systems. However, CRP appeared within the recently introduced TACE-specific STATE score [10]. Creatinine, although obviously not a liver parameter, has repeatedly been linked to prognosis in end-stage liver disease patients as documented by the MELD score for liver transplant allocation. An elevated creatinine was associated with a higher risk of TACE-related mortality and reduced OS [25]. The observation that in total two liver function, two tumor, and two other parameters comprise the M-TACE score underlines the need for a multidimensional HCC staging system owing to the complex interplay of the malignant tumor with a dysfunctional organ.

Performance of staging systems

It has been hypothesized that consequent utilization of staging systems can actually improve survival in HCC patients [1]. However, it has become increasingly clear that because of the wide range of clinical presentations of HCC, different staging systems should be applied to different subgroups [9]. TNM, which lacks any parameter reflecting a patients liver function, is endorsed for surgical patients [9] and its parameters in part rely on precise pathological results; TNMs' inferior performance in a cohort of nonsurgical patients therefore comes as no surprise. CLIP was identified as the established staging system with the best performance. This confirms the results of an earlier TACE study, which identified CLIP as the staging system with the best AIC [7]. The HAP score authors ranked their score highest compared with all other tested scores, but constrictively stated an 'at least as good performance as CLIP' [11]; thus a new staging system proposed for TACE patients at least has to prove its superior performance against the 'benchmark' score CLIP.

Lately, two TACE-specific pretherapeutic staging systems have been published [10,11]. Both the HAP and STATE score are well designed and easy to use. The only prognostic factor shared by all three relevant TACE scores (HAP, STATE, and M-TACE) is tumor extension, although characterized by different criteria: whereas M-TACE includes number and size of tumor nodes, vascular involvement, and extrahepatic spread, HAP considers a maximum diameter of more than 7 cm and STATE a tumor exceeding the 'up-to-seven criteria' [10]. In addition, overlapping criteria in two of the three scores are as follows: bilirubin [>1 mg/dl (HAP) vs. >1 and >3 mg/dl (M-TACE)], AFP [>400 ng/ml (HAP) vs. 35–999 and ≥ 1000 ng/ml (M-TACE)], CRP [>1 mg/dl (STATE) and 0.5–1.9 and ≥ 2 mg/dl (M-TACE)], and albumin (HAP and STATE). Although an ideal staging system should be as simple as possible, the higher number of prognostic factors in M-TACE and the finely graduation of some of its parameters could be an explanation for its superior prognostic performance.

European Association for the Study of the Liver and American Association for the Study of Liver Diseases endorse the Barcelona classification [4,5]. BCLC selectively recommends TACE for stage B patients [3]. In contrast, HCC patients outside of stage B do receive TACE therapy in clinical practice [1,6,27]; in fact, TACE is the most frequent first-line treatment in Europe, North America, and China across all BCLC stages [28]. This observation was underlined by the results of our retrospective study. Here, less than half of the examined patients were in stage B at the time of first TACE treatment. The two largest subgroups BCLC-B ($n=77$) and BCLC-C patients ($n=78$) did not show distinct survival ($P=0.945$). Another study suggested that a subset of BCLC-C patients (HAP stages A or B) may benefit from TACE [29]. In addition, an ECOG of more than 0 automatically excludes patients from BCLC-B and therefore from TACE treatment, although a recent study indicated that ECOG 1 could be included in BCLC-B [30] and our own data show that ECOG 0 versus 1 is not a prognostic factor in TACE patients ($P=0.261$). However, even after inclusion of ECOG 1 patients in the BCLC-B stage, survival times of BCLC-B and BCLC-C patients in our TACE cohort did not separate significantly ($P=0.657$). In the light of these results, the allocation to TACE through the BCLC treatment algorithm should be reassessed. Prognostic parameters, partly not included in BCLC, determine the prognosis of TACE patients.

Clinical implications

It is reassuring to find that M-TACE worked even better in the external than in the internal population (*c*-index: 0.754 vs. 0.710), although the external cohort showed different patient characteristics. Further study including a prospective validation is needed before a routine use of M-TACE for TACE prognosis in patient and study management can be endorsed. For a meaningful staging system, linkage to treatment indication is mandatory [3]. A possible M-TACE treatment algorithm is outlined in Fig. 4. Patients in stage II are 'typical' TACE candidates, whereas patients in stage I can receive TACE with presumably longer survival time. Patients in stage III, however, are characterized by a poor prognosis. Therefore, alternate, noninvasive palliative treatment strategies such as sorafenib or best supportive care should be favored. Although this treatment algorithm can be useful for conceptualizing treatment options, a too dogmatic approach holds the danger of withholding patients from more aggressive therapy. Therefore, M-TACE should not be a

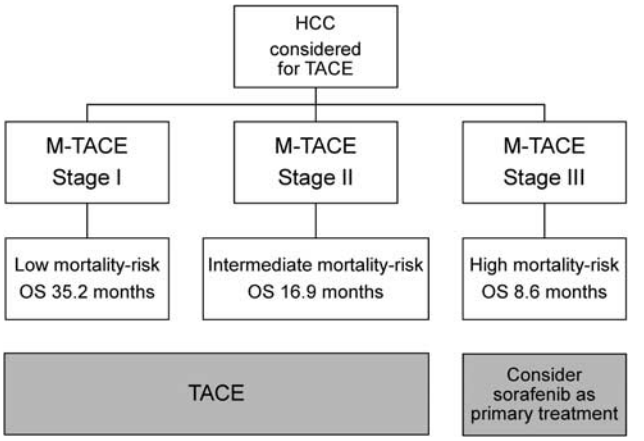

Fig. 4. M-TACE treatment algorithm. HCC, hepatocellular carcinoma; Munich-TACE. Munich-transarterial chemoembolization.

substitute but rather a supportive tool for individual multidisciplinary treatment decisions. The putative advantage of the two other TACE scores HAP and STATE is the lower score complexity. However, a 'bedside calculation of the score' as indicated by Hucke *et al.* [10] is not of great clinical importance, because every HCC treatment should be decided upon in a tumor board of an experienced HCC center [1] and our score still is easy enough to be applied in this setting: M-TACE demands the decision of an experienced radiologist whether the tumor is large, multifocal, or shows vascular or extrahepatic involvement, with one positive factor being sufficient to assign the four score points; the remaining M-TACE points are solely dependent on objective laboratory parameters. The M-TACE score is focussed on pretherapeutic prognosis estimation, it does not provide advice to the question whether to perform sequential TACE sessions. Scores considering post-first-TACE prognostic factors do exist [31] but follow a different approach and are not useful for the process of deciding whether to TACE a HCC patient at all. Therefore, clinicians need two different TACE scores as sketched by the START strategy [10]. On the basis of the performance of M-TACE in the present study, it can be regarded as a promising staging system for the initial c-TACE treatment decision in treatment-naive HCC patients.

Acknowledgements

The authors thank Julia von Baross for technical assistance, and Stephanie Hempe (S.H.) and Dr. Julia Sappl for assistance in data collection.

This research was presented in part to the Faculty of Medicine, Ludwig Maximilians University, as part of the doctoral thesis of P.O.D.W.

Conflicts of interest

There are no conflicts of interest.

References

1 El-Serag HB. Hepatocellular carcinoma. *N Engl J Med* 2011; 365:1118–1127.
2 Llovet JM, Bruix J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: Chemoembolization improves survival. *Hepatology* 2003; 37:429–442.
3 Forner A, Reig ME, de Lope CR, Bruix J. Current strategy for staging and treatment: the BCLC update and future prospects. *Semin Liver Dis* 2010; 30:61–74.
4 Bruix J, Sherman M. American Association for the Study of Liver Diseases. 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 Kim BK, Kim SU, Park JY, Kim DY, Ahn SH, Park MS, *et al*. Applicability of BCLC stage for prognostic stratification in comparison with other staging systems: single centre experience from long-term clinical outcomes of 1717 treatment-naive patients with hepatocellular carcinoma. *Liver Int* 2012; 32:1120–1127.
7 Cho YK, Chung JW, Kim JK, Ahn YS, Kim MY, Park YO, *et al*. Comparison of 7 staging systems for patients with HCC undergoing transarterial chemoembolization. *Cancer* 2008; 112:352–361.
8 Chen C-H, Hu F-C, Huang G-T, Lee P-H, Tsang Y-M, Cheng A-L, *et al*. Applicability of staging systems for patients with hepatocellular carcinoma is dependent on treatment method-analysis of 2010 Taiwanese patients. *Eur J Cancer* 2009; 45:1630–1639.
9 Vauthey J-N, Dixon E, Abdalla EK, Helton WS, Pawlik TM, Taouli B, *et al*. Pretreatment assessment of hepatocellular carcinoma: expert consensus statement. *HPB (Oxford)* 2010; 12:289–299.
10 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.
11 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.
12 op den Winkel M, Nagel D, Sappl J, op den Winkel P, Lamerz R, Zech CJ, *et al*. Prognosis of patients with hepatocellular carcinoma. Validation and ranking of established staging-systems in a large western HCC-cohort. *Plos One*, 7:e45066.
13 [No authors listed]. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: the Cancer of the Liver Italian Program (CLIP) investigators. *Hepatology* 1998; 28:751–755.
14 Chevret S, Trinchet JC, Mathieu D, Rached AA, Beaugrand M, Chastang C. A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire. *J Hepatol* 1999; 31:133–141.
15 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.
16 Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, *et al*. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* 1985; 56:918–928.
17 Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* 2010; 17:1471–1474.
18 Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973; 60:646–649.
19 Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984; 3:143–152.
20 Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control,* 1974; 19:716–723.
21 Huitzil-Melendez F-D, Capanu M, O'Reilly EM, Duffy A, Gansukh B, Saltz LL, *et al*. Advanced hepatocellular carcinoma: which staging systems best predict prognosis? *J Clin Oncol* 2010; 28:2889–2895.
22 Raoul J-L, Sangro B, Forner A, Mazzaferro V, Piscaglia F, Bolondi L, *et al*. Evolving strategies for the management of intermediate-stage hepatocellular carcinoma: available evidence and expert opinion on the use of transarterial chemoembolization. *Cancer Treat Rev* 2011; 37:212–220.
23 Takayasu K, Arii S, Ikai I, Omata M, Okita K, Ichida T, *et al*. Prospective cohort study of transarterial chemoembolization for unresectable hepatocellular carcinoma in 8510 patients. *Gastroenterology* 2006; 131:461–469.
24 O'Suilleabhain CB, Poon RTP, Yong JL, Ooi GC, Tso WK, Fan ST. Factors predictive of 5-year survival after transarterial chemoembolization for inoperable hepatocellular carcinoma. *Br J Surg* 2003; 90:325–331.

25 Garwood ER, Fidelman N, Hoch SE, Kerlan RK, Yao FY. Morbidity and mortality following transarterial liver chemoembolization in patients with hepatocellular carcinoma and synthetic hepatic dysfunction. *Liver Transpl* 2013; 19:164–173.
26 Carr BI, Pancoska P, Branch RA. Tumor and liver determinants of prognosis in unresectable HCC: a large case cohort study. *Hepatol Int* 2010; 4:396–405.
27 Sangiovanni A, Colombo M. Treatment of hepatocellular carcinoma: beyond international guidelines. *Liver Int* 2016; 36 (Suppl 1):124–129.
28 Park JW, Chen M, Colombo M, Roberts LR, Schwartz M, Chen PJ, *et al*. Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. *Liver Int* 2015; 35:2155–2166.
29 Pinato DJ, Arizumi T, Jang JW, Allara E, Suppiah P, Smirne C, *et al*. Combined sequential use of HAP and ART scores to predict survival outcome and treatment failure following chemoembolization in hepatocellular carcinoma: a multi-center comparative study. *Oncotarget* 2016; 7:44705–44718.
30 Hsu C-Y, Lee Y-H, Hsia C-Y, Huang Y-H, Su C-W, Lin H-C, *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.
31 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.