



# Novel Albumin–Bilirubin Grade-Based Risk Prediction Model for Patients with Hepatocellular Carcinoma Undergoing Chemoembolization

Jung Hee Kim<sup>1</sup> · Dong Hyun Sinn<sup>2</sup> · Jeong-Hoon Lee<sup>3</sup> · Dongho Hyun<sup>4</sup> · Sung Ki Cho<sup>4</sup> · Sung Wook Shin<sup>4</sup> · Young Chang<sup>3</sup> · Yoon Jun Kim<sup>3</sup> · Jung-Hwan Yoon<sup>3</sup> · Wonseok Kang<sup>2</sup> · Geum-Youn Gwak<sup>2</sup> · Yong-Han Paik<sup>2</sup> · Joon Hyeok Lee<sup>2</sup> · Kwang Cheol Koh<sup>2</sup> · Seung Woon Paik<sup>2</sup> · Moon Seok Choi<sup>2</sup>

Received: 20 July 2017 / Accepted: 16 January 2018 / Published online: 13 February 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

**Background** Recently, albumin–bilirubin (ALBI) grade has been suggested as a better surrogate for hepatic functional reserve for patients with hepatocellular carcinoma (HCC).

**Aims** We developed and validated a novel prediction model to predict outcome for HCC patients who underwent transcatheter arterial chemoembolization (TACE) as a first-line therapy.

**Methods** From a multivariate Cox regression model for overall survival, five objective variables (ALBI grade), the Barcelona clinic liver cancer (BCLC) stage, response after the first TACE session, Alpha-fetoprotein level, and sex were chosen and the ABRAS score was developed from the derivation cohort ( $n = 476$ ) and scored to generate an 8-point risk prediction model. The model's prognostic performance was assessed in the randomly assigned internal validation set ( $n = 475$ ) and external validation set ( $n = 243$ ).

**Results** The ALBI grade was able to stratify patient survival within the same Child–Pugh class. The time-dependent area under receiver operating characteristics curves (AUROCs) for overall survival at 1 and 3 years were 0.78 and 0.73 in the training set, 0.78 and 0.71 in the internal validation set, and 0.70 and 0.65 in the external validation set, respectively. When stratified by BCLC stage, ABRAS score at a cutoff point of more than 3, 4, and 5 for BCLC stage 0/A, B, and C could identify subset of patients with dismal prognosis.

**Conclusion** ABRAS score was useful in estimating prognosis for patients who underwent TACE as a first-line therapy. This score can be useful in planning and guiding treatment strategies with TACE, which warrants prospective validation.

**Keywords** Hepatocellular carcinoma · Transcatheter arterial chemoembolization · Albumin–bilirubin grade

---

Jung Hee Kim, Dong Hyun Sinn and Jeong-Hoon Lee have contributed equally to this study.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10620-018-4934-6>) contains supplementary material, which is available to authorized users.

---

✉ Moon Seok Choi  
drmschoi@gmail.com

<sup>1</sup> Department of Internal Medicine, Hallym University Hangang Sacred Heart Hospital of Hallym University Medical Center, Chuncheon, Korea

<sup>2</sup> Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Korea

## Abbreviations

ALBI	Albumin–bilirubin
HCC	Hepatocellular carcinoma
TACE	Transcatheter arterial chemoembolization
BCLC	Barcelona clinic liver cancer
AUROC	Area under receiver operating characteristics curves

<sup>3</sup> Department of Internal Medicine and Liver Research Institute, Seoul National University College of Medicine, Seoul, Korea

<sup>4</sup> Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

AST	Aspartate aminotransferase
ALT	Alanine aminotransferase
AFP	Alpha-fetoprotein
PIVKA-II	Protein induced by vitamin K antagonist
mRECIST	Modified response evaluation criteria in solid tumors
CR	Complete response
PR	Partial response
SD	Stable disease
PD	Progressive disease

## Introduction

Transcatheter arterial chemoembolization (TACE) is the most commonly used nonsurgical treatment modality for patients with hepatocellular carcinoma (HCC) [1]. TACE has shown a survival benefit over best supportive care for patients with unresectable HCC in two randomized controlled trials [2, 3] and is the first-line treatment recommended for Barcelona clinic liver cancer (BCLC) intermediate stage (multinodular, preserved liver function, and good performance status) HCC [4, 5]. TACE has also shown efficacy for early-stage and advanced-stage HCC [6–8] and is widely practiced across the BCLC stages [9].

Proper assessment of risk and benefit is a fundamental step in planning TACE. To date, there are several tools to guide further treatment strategies. ART score has been proposed to guide the re-treatment strategy for patients who underwent TACE [10]; it is composed of aspartate aminotransferase (AST) increase, radiologic tumor response, and Child–Pugh score increase after the first TACE. ABCR score has been proposed, which is composed of serum alpha-fetoprotein (AFP) level, BCLC stage, Child–Pugh score increase, and radiologic tumor response [11]. SNACOR score has also been proposed, which is composed of tumor size, tumor number, AFP level, Child–Pugh class, and radiologic response [12]. These models can help predict outcome for patients who underwent TACE and can be used to guide further treatment strategies. Notably, Child–Pugh score is included in all of these models [10–12].

Recently, the albumin–bilirubin (ALBI) grade was suggested and validated as a better method of assessing liver function than Child–Pugh score in HCC patients [13, 14]. Therefore, we tested whether ALBI is better than Child–Pugh score at predicting overall survival in HCC patients who underwent TACE. Then, we developed a risk prediction model that integrates the ALBI grade as a stratifying biomarker of liver reserves. Finally, we validated the newly developed risk prediction model by investigating whether it can estimate prognosis and guide further treatment strategies.

## Methods

### Study Population

This is a retrospective cohort study based on a prospective HCC registry from Samsung Medical Center, Seoul, South Korea. The HCC registry for the period of January 2007 to December 2012 was reviewed for potential inclusion in this study. A detailed description of the Samsung Medical Center HCC registry is described in our previous paper [15]. Briefly, all newly diagnosed HCC patients who received care at our institution were prospectively registered. HCC was diagnosed either histologically or clinically, according to regional guidelines [16]. Well-trained abstractors collected data, including age at diagnosis, gender, date of diagnosis, etiology, liver function, tumor characteristics, tumor stage, and initial treatment modality. From this registry, we included 1254 patients who received conventional iodized oil-based TACE as an initial treatment. Of those, we excluded 303 patients who met any of the following exclusion criteria: (1) major vascular invasion ( $n = 126$ ), defined by a tumor invading main portal vein, (2) bile duct invasion ( $n = 21$ ), (3) extrahepatic spread ( $n = 4$ ), (4) combined treatment (radiofrequency ablation, radiation, or sorafenib) ( $n = 86$ ), (5) received a liver transplantation during the follow-up period ( $n = 58$ ), and (6) Child–Pugh class C ( $n = 8$ ). Finally, a total of 951 treatment-naïve patients with BCLC stage 0 to C HCC who received TACE as the first-line treatment were analyzed. The entire study population was randomly divided into a derivation set (50%,  $n = 476$ ) and an internal validation set (50%,  $n = 475$ ). We also validated the risk score consisted of patients who received care at Seoul National University Hospital, Seoul, Korea, using the same inclusion and exclusion criteria and consisted of 243 patients (external validation cohort). The study was reviewed and approved by the Institutional Review Board at Samsung Medical Center and by the Institutional Review Board at Seoul National University Hospital. Because the study is based on a retrospective analysis of existing administrative and clinical data, the requirement for obtaining informed patient consent was waived by the Institutional Review Board.

### TACE Procedure and Follow-Up

After selective arteriography of the superior mesenteric, celiac, and common hepatic arteries using a 5-French catheter, a microcatheter was positioned into or as close as possible to the tumor feeding branch. A mixture of doxorubicin hydrochloride (adriamycin, Ildong, Seoul,

Korea) and iodized oil (Lipiodol; Guerbet, Aulnay-sous-Bois, France) was slowly infused through that catheter. The doses of iodized oil and doxorubicin were determined based on the size and vascularity of the tumor; the maximum doses of iodized oil and doxorubicin administered in a single session were 25 ml mg and 70 mg, respectively. Thereafter, embolization was performed with 1- to 2-mm-diameter gelatin sponge pledgets (Cutanplast, Mascia Brunelli, Milan, Italy) until blood flow ceased. Patients were followed up in 1 month with contrast-enhanced CT or MRI to evaluate their response to treatment. Then, patients were monitored at 3–6-month intervals. When recurrence was identified, subsequent treatment was performed if clinically feasible.

## Variables

The primary end point was overall survival, which was defined as the time from the initial TACE procedure to death or last follow-up, whichever comes first. We used the following variables for the HCC registry: age at diagnosis, sex, etiology of liver disease, Child–Pugh score, albumin, bilirubin, alanine aminotransferase (ALT), AST, maximum tumor size, number of tumors, presence of vascular invasion, AFP, protein induced by vitamin K antagonist (PIVKA-II) level, and BCLC stage. For this study, we additionally collected tumor response and Child–Pugh score, albumin, bilirubin, AST, AFP, and PIVKA-II level at the time of response evaluation after the initial TACE session. Radiologic response was evaluated with mRECIST criteria, as described in a prior report [17], with four response categories: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD).

## Statistical Analyses

We tested whether ALBI grade can further stratify overall survival within patients with the same Child–Pugh score or vice versa using the Kaplan–Meier method, and differences were assessed using the log-rank test. To generate the risk prediction model, the entire study population was randomly divided into a derivation set (50%,  $n = 476$ ) and a validation set (50%,  $n = 475$ ). We tested baseline variables (age, sex, etiology of liver disease, ALBI grade, AST, ALT, alkaline phosphatase, prothrombin time, platelet count, albumin, creatinine, AFP, PIVKA-II, number of tumors, maximum tumor size, presence of segmental vascular invasion, and BCLC stage), post-TACE variables at the first response valuation (week 4) (radiologic response, platelet, prothrombin time, albumin, AST, ALT, alkaline phosphatase, creatinine, AFP, PIVKA, and ALBI grade), and delta variables (change in AST, change in ALBI grade, change in AFP, and change in PIVKA-II) for possible associations with overall survival

using the Cox proportional hazards model in the derivation set. The performance of the prediction model was tested using the C-index in the validation set. Numeric variables (e.g., AFP) were categorized based on cutoffs. When choosing the cutoffs for numeric variables, different cutoffs were tested and the cutoff point was determined as the point that produced the highest C-index (data not shown). For multiple category variables (e.g., ALBI grade, BCLC stage, and radiologic response), patients were simplified into two or three categories based on the C-index of the final model. The regression coefficient of each risk predictor from the multivariable Cox proportional hazards model was divided by the regression coefficient of the lowest value, and the resulting number was rounded to an integer value to generate each score. Area under the receiver operating characteristics (AUROC) curve was calculated to measure the performance of the risk prediction model, and cutoff point was chosen to group patients into low and high risk scores. The developed risk score was validated internally using randomly assigned validation cohort ( $n = 457$ ) and using independent cohort from another institution ( $n = 243$ ). A two-sided  $p$  value  $< 0.05$  was considered significant.

## Results

### Baseline Characteristics

The baseline characteristics of the study subjects are summarized in Table 1. Vascular invasion limited to a segmental or subsegmental branch of the portal vein was observed in 113 patients (11.9%). The internal validation cohort had similar baseline characteristics as the derivation cohort (Table 1). However, there were significant differences between derivation and external validation cohort, in terms of age, Child–Pugh class, ALBI grade, incidence of vascular invasion, baseline AFP levels, and initial TACE response (Table 1).

### Comparison Between Child–Pugh Score and ALBI Grade

The time-dependent AUROC for overall survival showed the ALBI grade as a better prognostic indicator than Child–Pugh class (Table 2). Among 814 patients with Child–Pugh class A HCC, the ALBI grade was 1, 2, and 3 for 429 (52.7%), 384 (47.2%), and 1 patient (0.1%), and overall survival was significantly different according to the ALBI grade among patients with Child–Pugh class A (Fig. 1a). Of the 137 patients with Child–Pugh class B, the ALBI grade was 1, 2, and 3 for 1 (0.7%), 112 (81.8%), and 24 patients (17.5%), and overall survival was significantly different according to the ALBI grade among patients with Child–Pugh class B

**Table 1** Comparison of baseline characteristics

Characteristics	Derivation ( <i>n</i> = 476)	Internal validation ( <i>n</i> = 475)	External validation ( <i>n</i> = 243)	<i>p</i> value	<i>p</i> value
Age (years)	62.3 ± 10.3	63.1 ± 10.8	65.8 ± 10.1	0.24	< 0.001
Male ( <i>n</i> , %)	381 (80.0)	376 (79.2)	189 (77.8)	0.73	0.77
Etiology, HBV ( <i>n</i> , %)	346 (72.7)	328 (69.1)	173 (71.2)	0.21	0.67
Child–Pugh class				0.14	< 0.001
5	290 (60.9)	309 (65.1)	110 (45.3)		
6	107 (22.5)	108 (22.7)	68 (28.0)		
7–9	79 (16.6)	58 (12.2)	65 (26.7)		
ALBI grade				0.11	0.001
1	200 (42.0)	230 (48.4)	70 (28.8)		
2	261 (54.8)	235 (49.5)	159 (65.4)		
3	15 (3.2)	10 (2.1)	14 (5.8)		
Tumor number				0.59	0.35
Single	225 (47.3)	238 (50.1)	110 (45.3)		
2–3	179 (37.6)	164 (34.5)	86 (35.4)		
≥ 4	72 (15.1)	73 (15.4)	47 (19.3)		
Tumor size (cm)				0.88	0.077
≤ 2.0	129 (27.1)	134 (28.2)	70 (28.8)		
2.1–5.0	181 (38.0)	182 (38.3)	108 (44.4)		
> 5.0	166 (34.9)	159 (33.5)	65 (26.7)		
Vascular invasion ( <i>n</i> , %) <sup>a</sup>	50 (10.5)	63 (13.3)	13 (5.3)	0.18	0.025
BCLC stage				0.35	0.057
O	34 (7.1)	43 (9.1)	26 (10.7)		
A	258 (54.2)	242 (50.9)	131 (53.9)		
B	134 (28.2)	127 (26.7)	73 (30.0)		
C	50 (10.5)	63 (13.3)	13 (5.3)		
AFP (log ng/ml)	47 (11–479)	52 (10–523)	17 (6–172)	0.70	< 0.001
PIVKA-II (log mIU/ml)	143 (24–500)	120 (28–500)	50 (22–949)	0.85	0.69
Treatment response				0.80	< 0.001
Complete response	124 (26.1)	125 (26.3)	116 (47.7)		
Partial response	240 (50.4)	243 (51.2)	63 (25.9)		
Stable disease	22 (4.6)	16 (3.4)	30 (12.3)		
Progressive disease	90 (18.9)	91 (19.2)	34 (14.0)		

Values are expressed as mean ± standard deviation or number (%)

HBV hepatitis B virus, ALBI albumin–bilirubin grade, BCLC stage Barcelona clinic liver cancer stage, AFP alpha-fetoprotein, PIVKA-II prothrombin induced by vitamin K absence-II

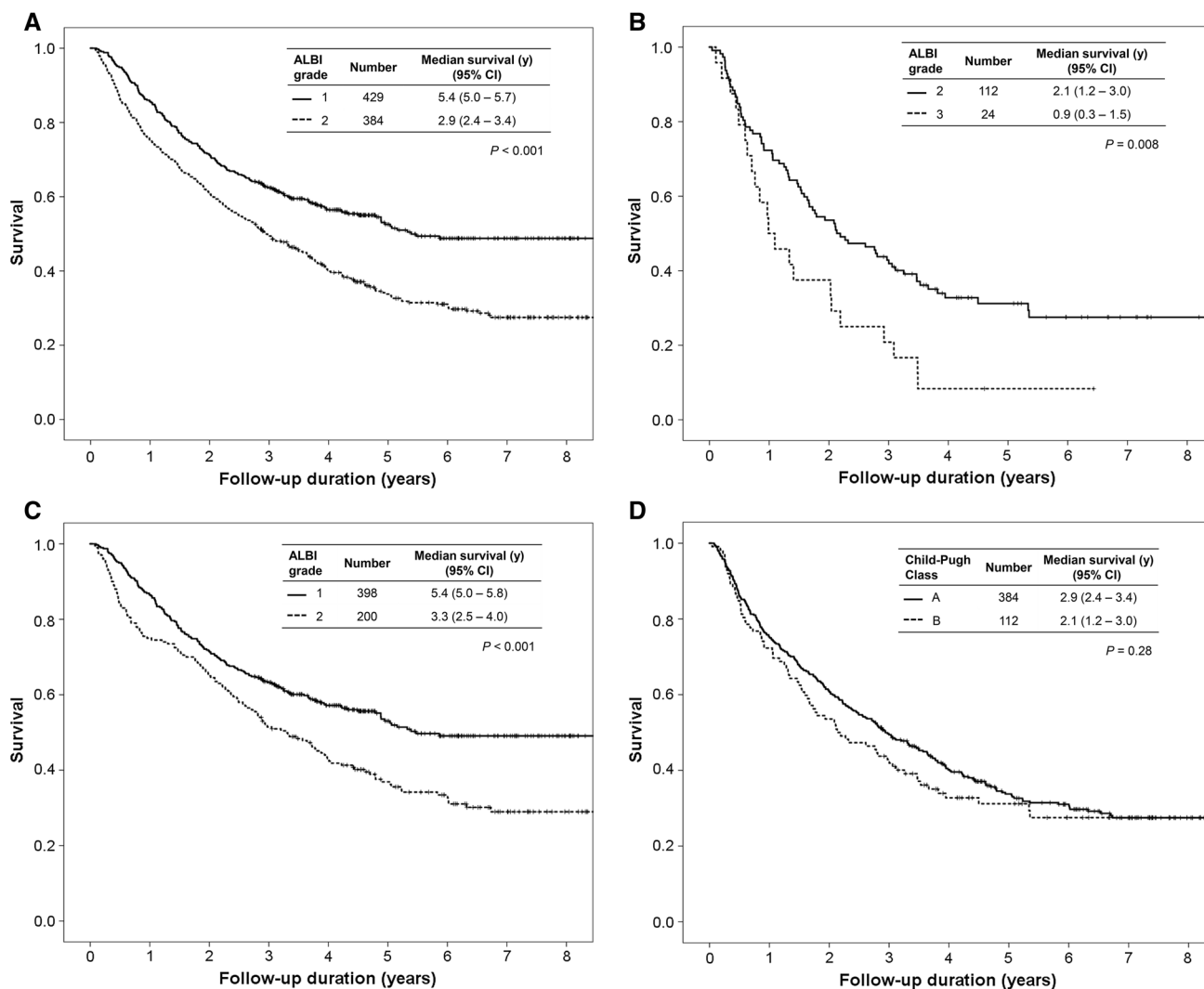
<sup>a</sup>All vascular invasion was limited to minor branch (segmental or subsegmental branch)

**Table 2** Predictive performance of ALBI grade versus Child–Pugh class for overall survival in whole study population (*n* = 951)

Year	ALBI grade	Child–Pugh class	<i>p</i> value
1	0.599 (0.567–0.631)	0.545 (0.513–0.577)	0.004
3	0.584 (0.552–0.615)	0.544 (0.512–0.576)	0.010
5	0.612 (0.580–0.643)	0.545 (0.577–0.613)	< 0.001

HCC (Fig. 1b). There were 599 patients with a Child–Pugh score of 5. The ALBI grade was 1, 2, and 3 for 398 (66.4%), 200 (33.4%), and 1 patient (0.2%), respectively, among

the patients with a Child–Pugh score of 5, and overall survival was also significantly different according to the ALBI grade among patients with a Child–Pugh score of 5 (Fig. 1c). When classified according to the ALBI grade, the Child–Pugh class was A for 429 of 430 patients (99.8%) with an ALBI grade 1, and the Child–Pugh class was B for 24 of 25 patients (96.0%) with an ALBI grade 3. Among 496 patients with ALBI grade 2 HCC, the Child–Pugh class was A for 384 patients (77.4%) and B for 112 patients (22.6%). Overall survival was not different according to Child–Pugh class among patients with an ALBI grade of 2 (Fig. 1d).



**Fig. 1** Overall survival was significantly different according to ALBI grade among patients with the same Child–Pugh score: Child–Pugh class A (a), Child–Pugh class B (b), or Child–Pugh score 5 (c). How-

ever, overall survival was not different according to Child–Pugh class in patients with an ALBI grade of 2 (d)

After TACE, an increase in Child–Pugh score (more than one point) was observed in 207 patients (21.8%). Of them, 53 patients had more than 2-point increase in Child–Pugh score. The survival was worse in patients who had an increase in Child–Pugh score after TACE (40.5 vs. 57.5% at 3 years,  $p < 0.001$ ). There were 163 patients (17.1%) who had an increase in ALBI grade after TACE. Likewise, those with an increase in ALBI grade showed worse survival than those without (41.6 vs. 56.3% at 3 years,  $p = 0.004$ ).

### Development of the Risk Model

Independent predictors of survival were selected from Cox regression analysis of data from the derivation cohort. The performance of the model was tested in the validation set using the C-index (Supplementary Table 1). Finally, ALBI

grade (1 vs. 2 vs. 3), BCLC stage (0/A vs. B vs. C), radiologic response according to mRECIST criteria (CR + PR vs. SD + PD), baseline AFP level ( $\geq 100$  ng/ml vs.  $< 100$  ng/ml), and sex (male vs. female) were selected as variables to generate the risk prediction model. The risk scoring model (ABRAS score: ALBI grade, BCLC stage, radiologic response, AFP level, and sex) was the sum of each score assigned to five risk predictors and ranged from 0 to 8 (Table 3).

### Predictive Performance of the ABRAS Score

The median follow-up duration (range) was 3.0 years (0.1–8.8 years), 3.0 years (0.1–8.7 years), and was 2.2 years (0.1–4.1 years) for derivation, internal validation, and external validation set. In the derivation set, the time-dependent

**Table 3** Prognostic factors for overall survival in the derivation cohort ( $n = 476$ )

Variables	Unadjusted HR (95% CI)	$p$ value	Multivariable adjusted HR (95% CI)	$p$ value	Beta coefficient	Risk score
<i>Sex</i>						
Female	Reference		Reference			0
Male	1.20 (1.03–1.41)	0.019	1.59 (1.15–2.19)	0.005	0.464	1
<i>AFP (ng/ml)</i>						
< 100	Reference		Reference			0
≥ 100	2.07 (1.64–2.62)	< 0.001	1.97 (1.53–2.52)	< 0.001	0.678	1
<i>BCLC stage</i>						
O/A	Reference		Reference			0
B	1.79 (1.37–2.33)	< 0.001	1.71 (1.31–2.24)	< 0.001	0.540	1
C	3.73 (2.65–5.25)	< 0.001	2.80 (1.96–4.02)	< 0.001	1.033	2
<i>ALBI grade</i>						
1	Reference					0
2	1.61 (1.25–2.07)	< 0.001	1.77 (1.37–2.30)	< 0.001	0.576	1
3	4.30 (2.44–7.57)	< 0.001	7.11 (3.98–12.7)	< 0.001	1.962	3
<i>Treatment response</i>						
CR + PR	Reference		Reference			0
SD + PD	1.79 (1.38–2.32)	< 0.001	1.81 (1.39–2.36)	< 0.001	0.596	1

*HBV* hepatitis B virus, *AFP* alpha-fetoprotein, *BCLC stage* Barcelona clinic liver cancer stage, *TACE* transarterial chemoembolization

**Table 4** Predictive performance of prediction model to predict overall survival in the derivation and validation set

Year	Derivation set ( $n = 476$ )	Internal validation set ( $n = 475$ )	External validation set ( $n = 243$ )
1	0.78 (0.73–0.83)	0.78 (0.73–0.84)	0.70 (0.57–0.83)
3	0.73 (0.69–0.78)	0.71 (0.66–0.76)	0.65 (0.55–0.75)
5	0.75 (0.70–0.79)	0.72 (0.67–0.76)	— <sup>a</sup>

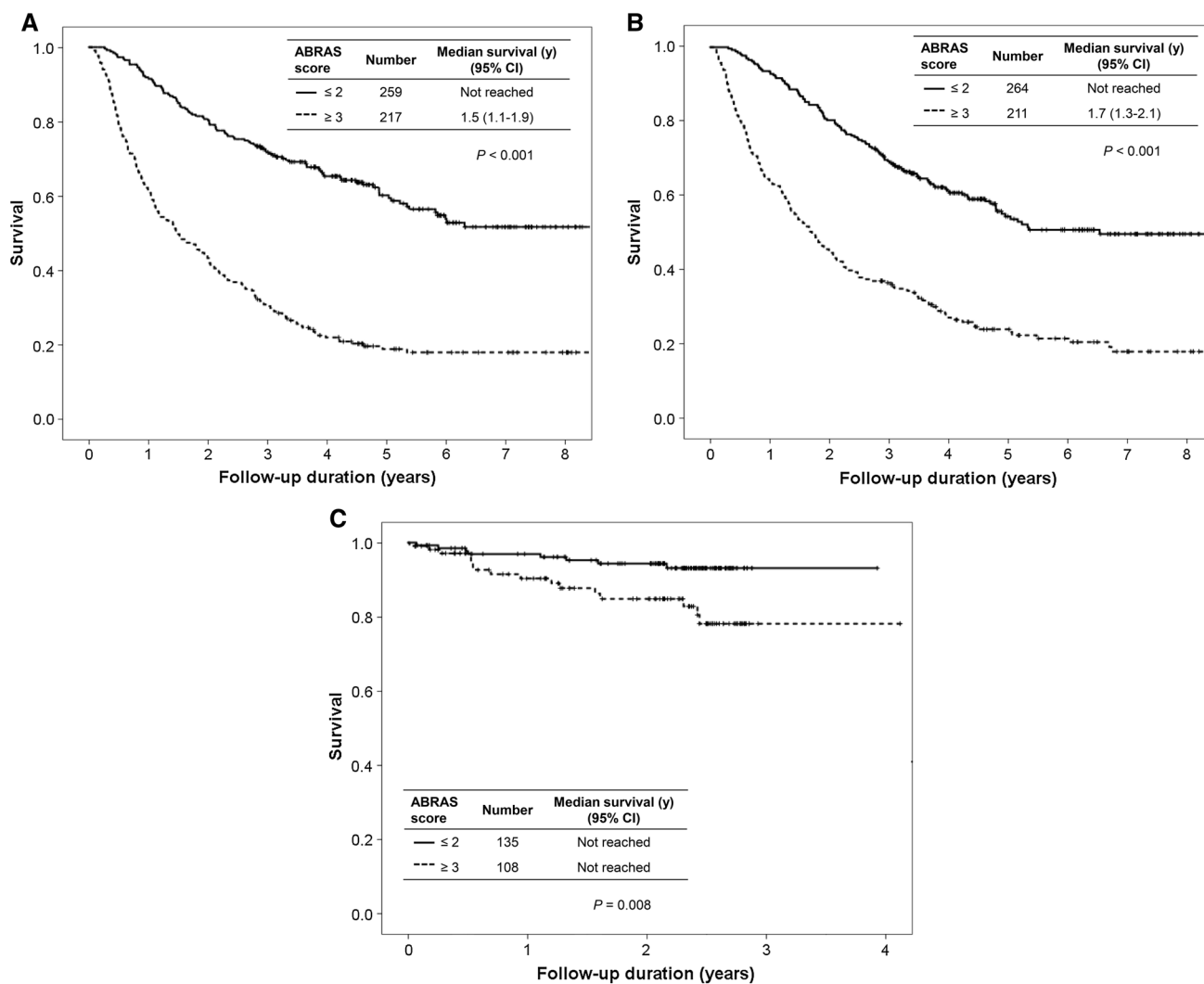
<sup>a</sup>Follow-up duration was short for external validation set to analyze predictive performance at 5 years

AUROC of the ABRAS model for overall survival were 0.78, 0.73, and 0.75 at 1, 3 and 5 years, respectively (Table 4). The time-dependent AUROCs of the model for overall survival were 0.78, 0.71, and 0.72 at 1, 3, and 5 years, respectively, in the internal validation set and were 0.70 and 0.65 at 1 and 3 years in the external validation set (Table 4). Compared to BCLC stage or ALBI grade alone, ABRAS score showed higher AUROCs (Supplementary Table 2). The ABRAS model was also better than treatment response after TACE alone (Supplementary Table 2). According to their risk score, patients were stratified into low (score 0–2) and high risk (score 3–8). Survival was significantly different according to risk group, with a 3-year survival rate of 72 and 30%, respectively, in the derivation set (Fig. 2a), and 69 and 35%, respectively, in the internal validation set (Fig. 2b) ( $p < 0.001$ ). Survival was also significantly different according to risk group in external validation set (2-year survival

rate of 94 and 84% for low and high risk scores,  $p = 0.008$ ) (Fig. 2c). When stratified by BCLC stage, the best cutoff point for the ABRAS score was score of 2 for BCLC stage 0 and stage A, score of 3 for BCLC stage B, and score of 4 for BCLC stage C (Fig. 3), which corresponded to cutoff score of more than 2 using ALBI grade, radiologic response, AFP level, and sex.

## Discussion

Patients who are receiving TACE present with varying tumor burdens and liver functions, and there is considerable variation in the clinical benefit that patients derive from TACE [18, 19]. In this respect, a good risk prediction model is needed for HCC patients undergoing TACE to optimize HCC management strategies. In this study, we developed and validated a new risk prediction model (ABRAS score) that can be used to predict prognosis and guide further treatment strategies for patients undergoing TACE. The five variables used in this model are simple clinical variables that can easily be recognized (ALBI grade, BCLC stage, radiologic response, AFP level, and sex). Age, sex, BCLC stage, and radiologic response after TACE have been reported as important factors associated with prognosis in HCC patients [1, 13, 20]. The uniqueness of this model is that the ABRAS score uses the ALBI grade to assess liver function which is a more objective



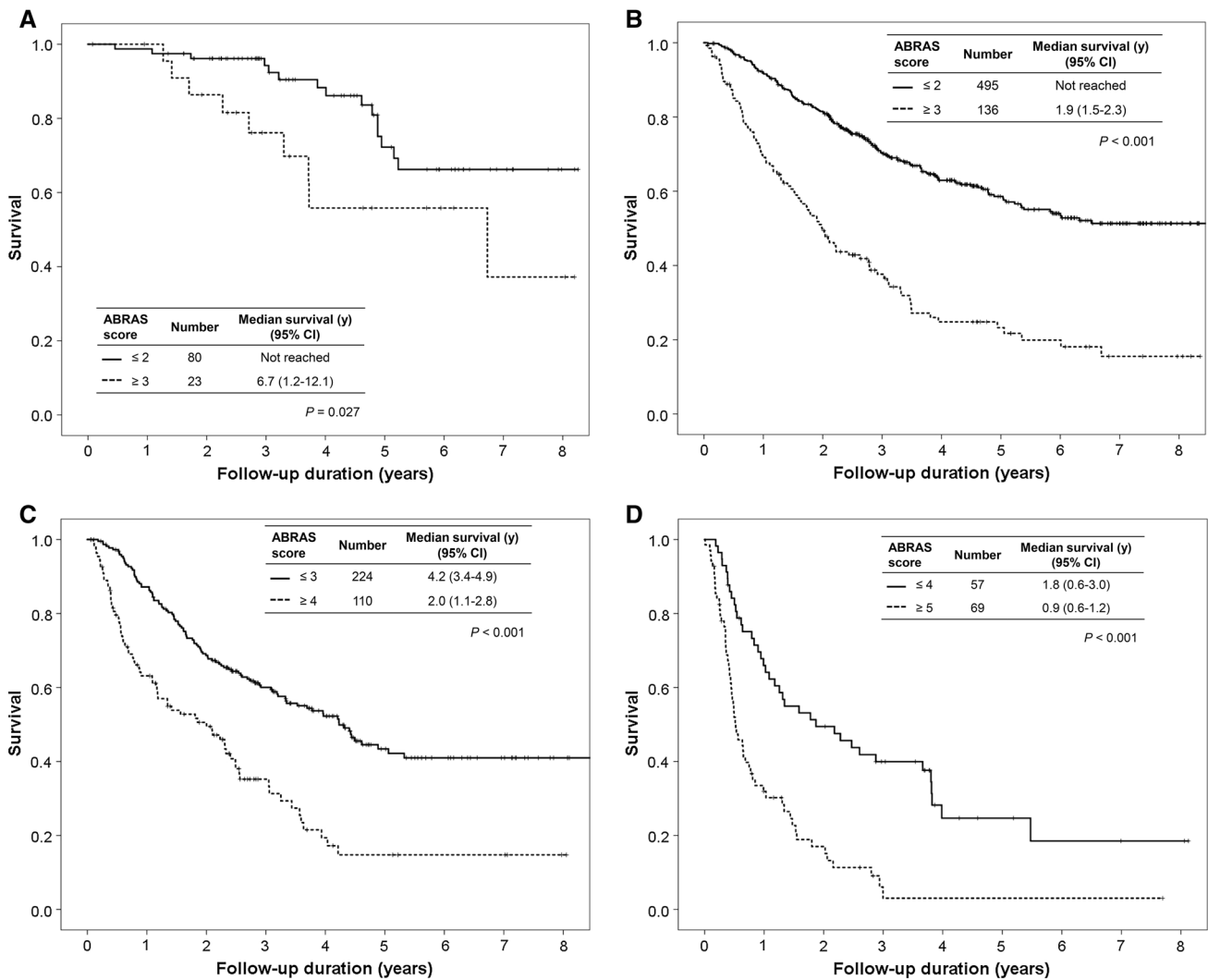
**Fig. 2** Overall survival was significantly different according to the ABRAS score in each study model. Those with a high score showed a lower survival than those with low score: derivation set (a), internal validation set (b), or external validation set (c)

and brief value than Child–Pugh class that is included in previous models to guide the re-treatment strategy of TACE.

Notably, the ALBI grade could discriminate between distinct prognostic groups within the same category of liver function as assessed by Child–Pugh class, while Child–Pugh class could not discriminate between distinct prognostic groups within the same category of ALBI grade. Assessment of liver function is particularly important because cirrhosis is a competing cause of death in HCC patients [13]. In this study, even in patients with a Child–Pugh score of 5, there were significant differences of survival according to ALBI grade, indicating that ALBI can further distinguish a low-risk group from a relatively poor-risk group within the same category based on Child–Pugh score. The ALBI grade has another advantage in that it is entirely composed of objective values. Consistent with the results of this study, ALBI grade

was shown to be a better tool than Child–Pugh score in the prediction of survival in HCC patients [14, 21, 22].

The prognostic performance of the ABRAS model in predicting overall survival was acceptable, expressed as time-dependent AUROCs of 0.73–0.78 in the derivation set and 0.71–0.78 in the internal validation set. The performance of the ABRAS model was better than BCLC stage, ALBI grade, or TACE response alone (Supplementary Table 2). When compared to derivation cohort, there were no significant differences in the baseline characteristics in internal validation cohort, while several characteristics were different in external validation cohort. The external validation cohort was comprised of older patients, poorer liver function, but less advanced tumor with lower serum AFP levels. Follow-up duration was shorter in external validation cohort as well. Initial treatment response was better in external validation cohort with better survival. The differences in



**Fig. 3** Overall survival was significantly different according to the ABRAS score. Those with a high score showed a lower survival than those with low score in each BCLC stage: BCLC stage O (a), BCLC stage A (b), BCLC stage B (c), or BCLC stage C (d)

the several characteristics suggest that external validation cohort is composed of quite different population, although all of them received TACE as an initial treatment strategy. However, ABRAS score was also able to differentiate patient prognosis in the external validation cohort as well, although AUROCs of the model for the overall survival was slightly lower (0.70 and 0.65 at 1 and 3 years) than derivation cohort (0.78 and 0.73 at 1 and 3 years). When patients were stratified into two groups according to each one’s ABRAS score, patients showed a difference in survival without any overlap in the survival curve among high- and low-risk patients. When stratified according to BCLC stage, ABRAS score was able to classify patients into high- and low-risk patients across all the spectrum of BCLC stage (BCLC 0–BCLC C).

Notably, although we tested several on-treatment variables, only radiologic response after the first TACE was selected in this risk prediction model. The other variables,

such as sex, AFP levels, BCLC stage, and ALBI grade, were baseline factors. In ART score and ABCR score, an increase in Child–Pugh score was found to be an independent factor to predict survival [10, 11]. In our cohort, most patients had preserved liver function, and Child–Pugh increase or ALBI increase was not selected as an independent factor for survival, as Child–Pugh increase or ALBI grade increase after TACE was not frequently observed in our cohort. Having baseline factors in this model means this simple risk score can be useful in planning the first TACE. Having radiologic response after TACE in the risk model indicates this model may be useful to estimate further prognosis after the first TACE and may guide further treatment. Prospective validation of ABRAS score is warranted to see whether ABRAS may also help guiding treatment.

Our data have some limitations. This was a retrospective study with several inherent limitations including potential



selection, measurement, and misclassification biases. Because of the long study duration, many radiologists were involved in assessing treatment response after TACE. However, as the radiologists were unaware of the aims of the study, measurement errors in evaluating the response to TACE were independent and non-differential. Several important factors that can be associated with long-term outcome of patients were not assessed. For patients treated with TACE, the level of TACE (lobar, selective, superselective) are an important factor associated with treatment outcome, which we did not have information about [23, 24]. In this cohort, about 75% of patients underwent the second TACE after initial treatment response evaluation, and about 90% of the patients were re-treated with multiple treatment modalities during follow-up (data not shown). Retreatment can be another important factor that determined overall survival of HCC patients [25]. This study is unable to answer whether those with high ABRAS score may benefit from other treatment over TACE. Major strengths of this study are its large sample size, incorporation of the novel ALBI grade in predicting the prognosis of HCC patients receiving TACE, and validation from an independent internal and external cohort. Our real-world cohort included a wide range of BCLC stages (early to advanced) which reflects real-world practices and can be used to all spectra of BCLC stage. Testing of not only baseline, but many on-treatment variables in the survival model is another major strength.

In summary, we developed and validated the ABRAS score, which is entirely composed of objective variables. The novel prediction model can estimate patients' prognosis and help guide treatment strategies for HCC patients planned or received initial treatment with TACE. Although this score needs to be validated in prospective trials and different populations, it can be a useful model in predicting prognosis and helping choose further treatment strategies for HCC patients treated with TACE.

**Author's contribution** Sinn DH and Choi MS contributed to the study design and supervision; Sinn DH and Kim JH were involved in the statistical analysis; Sinn DH, Lee JH, Kim JH, and Hyun D contributed to the writing of the draft manuscript; Kim JH, Lee JH, Cho SK, Shin SW, Chang Y, Kim JY, Yoon JH, Kang W, Gwak GY, Paik YH, Choi MS, Lee JH, Koh KC, and Paik SW collected the data; Lee JH, Chang Y, Kim JY, Yoon JH, Kang W, Gwak GY, Paik YH, Choi MS, Lee JH, Koh KC, Paik SW, and Choi MS were involved in the critical revision of the manuscript. All authors approved the final version of the manuscript.

**Funding** This paper was supported by Research and Business foundation of Sungkyunkwan University (Grant number: S-2015-0047-000).

### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Korean Liver Cancer Study G, National Cancer Center K. KLCSG-NCC Korea practice guideline for the management of hepatocellular carcinoma. *Gut Liver*. 2014;2015(9):267–317.
2. Llovet JM, Real MI, Montana X, et al. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. *Lancet*. 2002;359:1734–1739.
3. Lo CM, Ngan H, Tso WK, et al. Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. *Hepatology*. 2002;35:1164–1171.
4. Bruix J, Sherman M. American Association for the Study of Liver D. 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. Yang HJ, Lee JH, Lee DH, et al. Small single-nodule hepatocellular carcinoma: comparison of transarterial chemoembolization, radiofrequency ablation, and hepatic resection by using inverse probability weighting. *Radiology*. 2014;271:909–918.
7. Chung GE, Lee JH, Kim HY, et al. Transarterial chemoembolization can be safely performed in patients with hepatocellular carcinoma invading the main portal vein and may improve the overall survival. *Radiology*. 2011;258:627–634.
8. 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.
9. Park JW, Chen M, Colombo M, et al. Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. *Liver Int*. 2015;35:2155–2166.
10. Sieghart W, Huckle F, Pinter M, et al. The ART of decision making: retreatment with transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology*. 2013;57:2261–2273.
11. Adhoute X, Penaranda G, Naude S, et al. Retreatment with TACE: the ABCR SCORE, an aid to the decision-making process. *J Hepatol*. 2015;62:855–862.
12. Kim BK, Shim JH, Kim SU, et al. Risk prediction for patients with hepatocellular carcinoma undergoing chemoembolization: development of a prediction model. *Liver Int*. 2016;36:92–99.
13. Johnson PJ, Berhane S, Kagebayashi C, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach—the ALBI grade. *J Clin Oncol*. 2015;33:550–558.
14. Pinato DJ, Sharma R, Allara E, et al. The ALBI grade provides objective hepatic reserve estimation across each BCLC stage of hepatocellular carcinoma. *J Hepatol*. 2016.
15. Kim KM, Sinn DH, Jung SH, et al. The recommended treatment algorithms of the BCLC and HKLC staging systems: Does following these always improve survival rates for HCC patients? *Liver Int*. 2016;36:1490–1497.
16. Korean LC, National CC. Practice guidelines for management of hepatocellular carcinoma 2009. *Korean J Hepatol*. 2009;15(3):391.
17. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis*. 2010;30:52–60.
18. Bolondi L, Burroughs A, Dufour JF, et al. Heterogeneity of patients with intermediate (BCLC B) hepatocellular carcinoma: proposal for a subclassification to facilitate treatment decisions. *Semin Liver Dis*. 2012;32:348–359.

19. Yu SJ. A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010–2016. *Clin Mol Hepatol*. 2016;22:7–17.
20. Sinn DH, Gwak GY, Cho J, Paik SW, Yoo BC. Comparison of clinical manifestations and outcomes between hepatitis B virus- and hepatitis C virus-related hepatocellular carcinoma: analysis of a nationwide cohort. *PLoS ONE*. 2014;9:e112184.
21. Edeline J, Blanc JF, Johnson P, et al. A multicentre comparison between Child Pugh and Albumin–Bilirubin scores in patients treated with sorafenib for hepatocellular carcinoma. *Liver Int*. 2016;36:1821–1828.
22. Liu PH, Hsu CY, Hsia CY et al. ALBI and PALBI grade predict survival for HCC across treatment modalities and BCLC stages in the MELD Era. *J Gastroenterol Hepatol*. 2016.
23. Golfieri R, Cappelli A, Cucchetti A, et al. Efficacy of selective transarterial chemoembolization in inducing tumor necrosis in small (< 5 cm) hepatocellular carcinomas. *Hepatology*. 2011;53:1580–1589.
24. Lee EW, Khan S. Recent advances in transarterial embolotherapies in the treatment of hepatocellular carcinoma. *Clin Mol Hepatol*. 2017;23:265.
25. Dika IE, Abou-Alfa GK. Treatment options after sorafenib failure in patients with hepatocellular carcinoma. *Clin Mol Hepatol*. 2017;23:273–279.