Validation of the “Six-and-Twelve” Prognostic Score in Transarterial Chemoembolization–Treated Hepatocellular Carcinoma Patients

Apichat Kaewdech, MD¹, Pimsiri Sripongpun, MD¹, Natcha Cheewasereechon, MD¹, Sawangpong Jandee, MD¹, Naichaya Chamroonkul, MD¹ and Teerha Piratvisuth, MD¹,²

INTRODUCTION: The “six-and-twelve” prognostic score was proposed recently to predict survival rate in patients with unresectable hepatocellular carcinoma (HCC) treated with transarterial chemoembolization (TACE). However, it has not been validated externally. We validated this score and previous prognostic scores in Thai HCC patients treated with TACE.

METHODS: We identified all HCC patients who underwent TACE between January 2007 and December 2018 at our hospital. The inclusion criteria were treatment-naive, unresectable HCC BCLC-A and BCLC-B; if cirrhosis was present, Child-Pugh score ≤7; and baseline performance status 0–1.

RESULTS: Of 716 HCC patients undergoing TACE, 281 (mean age, 61.1 years; 73.0% men, 92.2% with cirrhosis) were eligible. Approximately half of the patients had hepatitis B virus. Median overall survival was 20.3 (95% confidence interval, 16.4–26.3) months. By stratifying with the “six-and-twelve” score (≤6, >6–12, >12), median (95% confidence interval) overall survival was 35.1 (26.4–53.0), 16.0 (11.6–22.6), and 7.6 (5.4–14.9) months, respectively. Area under the receiver operating characteristic curves (AUROCs) predicting death at 1, 2, and 3 years for the “six-and-twelve” score were 0.714, 0.700, and 0.688, respectively. Compared with the other currently available scores, the AUROC predicting death at 1 year for the “six-and-twelve” score was the most predictive and better than other models except the up-to-seven model.

DISCUSSION: Our study confirms the value of the “six-and-twelve” score to predict survival rate of unresectable HCC treated with TACE. However, in our validation cohort, AUROC of the “six-and-twelve” score was slightly lower than that of the original Chinese cohort (0.73).

Clinical and Translational Gastroenterology 2021;12:e00310. https://doi.org/10.14309/ctg.0000000000000310

INTRODUCTION

Hepatocellular carcinoma (HCC) is a major cancer that spreads globally including in Thailand (1). According to data from the Thai Ministry of Health, liver cancer is a huge burden, since it is the most and fifth most common cancer in Thai men and women, respectively (2). The mortality rate of HCC is high (age-standardized mortality rate, 21.5/100,000 per year). In Western countries, the major causes of HCC are hepatitis C and non-alcoholic fatty liver disease; by contrast, the most common causes in Eastern countries including Thailand are hepatitis B and alcohol (1,3,4).

Treatments of HCC are stratified according to the Barcelona Clinic Liver Cancer (BCLC) stages (3,4). BCLC staging system, which includes tumor burden, liver function, and patient performance status, is presently used. The treatment algorithms are according to the stages of the patients. The therapeutic options are surgical resection, liver transplantation, local ablation therapy, transarterial radioembolization, or systemic therapies (4). Transarterial chemoembolization (TACE) is a modality of choice for HCC patients with BCLC stage B (intermediate stage) (1,3,4). The procedure involves injection of chemotherapeutic agents emulsified with lipoidal agents percutaneously into the hepatic artery that feeds the tumor, followed by injection of embolic agents. Therefore, the procedure yields cytotoxic and ischemic effects to the tumor. TACE initially has been proved to improve survival in HCC (5–8).
According to several reported guidelines, TACE is recommended for HCC with Child A and highly selected Child B cirrhosis (3). Therefore, there were a variety of patients in this group and associated with different outcomes. Finally, pretreatment prognostic models are important tools to decide whether TACE should be performed in patients.

Recently, the “six-and-twelve” prognostic score was introduced to predict the best candidates for TACE (5). This prognostic score was developed from a large, multicenter Chinese cohort study. The study illustrated that the summation of the largest tumor diameter and the numbers of tumor was an independent predictor to estimate overall survival (OS) in patients diagnosed with unresectable HCC, BCLC-A and HCC, BCLC-B disease who underwent TACE. The 3 strata of the score included the summation of scores of ≤6, >6 up to 12, and >12. The higher stratum predicted lower patient survival. An internal validation was performed in the Chinese population, which had a high performance of score. However, there was no external validation in another population.

Currently, to the best of our knowledge, no study exists for external validation of the recently reported prognostic “six-and-twelve score” in HCC patients undergoing TACE (5). However, this prognostic score was derived from a population of Chinese ethnicity in which most patients had chronic hepatitis B, which was different from our population that varied in etiology of HCC. We validated the “six-and-twelve” score and other available prognostic scores in Thai HCC patients who underwent TACE.

METHODS

Patients and study design
A retrospective cohort study was conducted at Songklanagarind Hospital, Prince of Songkla University, Thailand, a supertertiary university hospital located in southern Thailand. An electronic hospital database (Health Informative System) was screened retrospectively. Patients diagnosed with HCC (International Classification of Diseases, 10th revision codes C220—malignant neoplasm of liver cell carcinoma and C229—malignant neoplasm of liver, unspecified) between January 1, 2007, and December 31, 2018, were retrieved. HCC was diagnosed by imaging or histological findings according to the American Association for the Study of Liver Disease (3), European Association for the Study of Liver (4), or Asian Pacific Association for the Study of the Liver (1) guidelines. Inclusion criteria were patients aged ≥18 years, first diagnosis of HCC, unresectable HCC BCLC-A or HCC BCLC-B, Child-Pugh score A5–B7, and TACE performed as monotherapy. Exclusion criteria were HCC with other active malignancies, history of spontaneously tumor rupture, cotreatment with any systemic or locoregional therapies during TACE session or at any time point, and absence of baseline imaging information. The inclusion criteria and exclusion criteria were consistent with the original Chinese cohort (5).

Baseline clinical characteristics, including age at diagnosis, sex, performance status, causes of HCC, Child-Pugh score, BCLC stage, and baseline α-fetoprotein level, were collected. Diagnosis of HCC, and tumor burden, including tumor size and number, was determined by abdominal radiologists and recorded in the Health Informative System.

OS was defined as the interval between initial TACE and death. Patient status at the end of the study (August 31, 2019) was defined as alive or death using data from the Thailand civil registration database.

Treatment procedures
All eligible patients underwent selective conventional TACE. The procedure was performed by 2 interventional radiologists who had at least 5 years of experience. The angiogram was selectively performed at celiac artery and superior mesenteric artery using a 5Fr catheter (Cobra or MIK catheter) and a 0.035-inch J-tip Terumo guidewire. The catheterization was performed at the subsegmental hepatic artery through the femoral artery route. The tumor-feeding hepatic arteries were identified and infused a mixture of lipiodol (4–16 mL) and chemotherapy agents as

Figure 1. Flow chart for patient eligibility. HCC, hepatocellular carcinoma; TACE, transarterial chemoembolization.
adriamycin (10–40 mg) or mitomycin (10–20 mg) under realtime angiography. After that, the feeding hepatic artery was embolized with gelatin sponge particles. On demand, TACE was scheduled at an interval of 4–8 weeks depending on the viability of HCC on follow-up imaging.

Statistical analysis
For baseline characteristics, quantitative variables were presented as median with interquartile range and categorical variables were presented as counts with proportions. Survival curves were estimated using the Kaplan-Meier method and compared by the log-rank test.

The “six-and-twelve” score was validated and compared with prognostic models, including up-to-seven criteria (6), four-and-seven criteria (7), hepatoma arterial-embolization prognostic (HAP) score (8), modified HAP-II (mHAP-II) score (9), and albumin bilirubin (ALBI) score (10). Harrell’s concordance index (C-index) and Akaike’s Information Criterion were used to assess the score discrimination ability. A confirmatory analysis was performed to evaluate the discrimination ability of the scoring systems by estimating the area under the receiver operating characteristic curve (AUROC) for each time point. Pairwise comparison of the AUROC was performed with the Delong test. Sensitivity, specificity, and positive and negative (NPV) predictive values then were calculated at the given cutoffs.

All statistical analyses were performed with R software, version 3.3.2 (R Foundation, Vienna, Austria). P < 0.05 (2-sided) was considered statistically significant.

Compliance with ethical standard
The study protocol was approved by the institutional research boards ethics committee of Faculty of Medicine, Prince of Songkla University, Thailand (REC: 62-314-14-1). The study protocol was conducted in accordance with the ethical guidelines of the 1975 Declaration of Helsinki.
RESULTS

Baseline clinical characteristics

Of 716 patients who underwent TACE during the study period, 281 were eligible for the study; 432 met the exclusion criteria as depicted in Figure 1. Their baseline clinical characteristics are shown in Table 1. Mean age was 61.1 years, 73.0% were men, and 92.2% had cirrhosis. The major etiology of HCC was hepatitis B virus (HBV) (49.8%), followed by hepatitis C virus (18.9%), alcohol (11.7%), nonalcoholic fatty liver disease (9.3%), cryptogenic (9.2%), and autoimmune hepatitis (1.1%). The Child-Pugh stage of liver cirrhosis was classified as A5 (43.8%), A6 (36.6%), and B7 (19.6%).

At initial diagnosis of HCC, the largest tumor size most commonly was **3–7 cm** (37.4%) followed by **≤3 cm** (34.9%). More than half (52.7%) of the patients had a single tumor. According to the BCLC system, the number of cases classified as BCLC stages 0, A, and B was 11.7%, 52.7%, and 35.6%, respectively.

Median values for serum alanine aminotransferase and platelet count were 36.0 (24.0–54.0) U/L and 124.0 (77.0–204.0) x 10^9/L, respectively. Mean serum albumin level was 3.5 (6.0–0.5) g/dL. Most patients had serum alpha-fetoprotein, 400 ng/mL (75.1%). The median number of TACE sessions was 2 per patient (1.0–4.0).

Overall survival

Median OS for the cohort was 20.3 (95% confidence interval [CI], 16.4–26.3) months, with 1-, 2-, and 3-year survivals being 62.5%, 45.6%, and 35.5%, respectively (Figure 2).

Survival according to the “six-and-twelve” score

By stratifying with the “six-and-twelve” score (≤6, >6–12, and >12), median OSs were 35.1 (95% CI, 26.4–53.0), 16.0 (95% CI, 11.6–22.6), and 9.2 (95% CI, 5.8–13.6) months, respectively.

### Table 1. Comparison of the performance and discriminative ability between the “six-and-twelve” score and other models

| Models         | 1-yr AUROC (95% CI) | P value^a | 2-yr AUROC (95% CI) | P value^a | 3-yr AUROC (95% CI) | P value^a | C-index (95% CI) | AIC       |
|----------------|---------------------|-----------|---------------------|-----------|---------------------|-----------|-----------------|-----------|
| Six-and-twelve | 0.714 (0.651–0.776) | Ref.      | 0.700 (0.639–0.761) | Ref.      | 0.688 (0.626–0.750) | Ref.      | 0.699 (0.691–0.706) | 354.171   |
| Up-to-seven    | 0.714 (0.651–0.776) | 1.000     | 0.700 (0.639–0.761) | 1.000     | 0.688 (0.626–0.750) | 1.000     | 0.699 (0.691–0.706) | 354.171   |
| Four-and-seven | 0.644 (0.586–0.701) | <0.001    | 0.645 (0.586–0.704) | 0.193     | 0.638 (0.580–0.696) | 0.245     | 0.611 (0.605–0.617) | 365.187   |
| HAP            | 0.703 (0.644–0.761) | 0.670     | 0.693 (0.635–0.750) | 0.777     | 0.688 (0.630–0.746) | 0.999     | 0.674 (0.666–0.681) | 352.790   |
| mHAP-II        | 0.696 (0.634–0.757) | 0.468     | 0.682 (0.623–0.742) | 0.459     | 0.674 (0.615–0.734) | 0.591     | 0.667 (0.660–0.675) | 355.446   |
| ALBI           | 0.622 (0.557–0.687) | 0.036     | 0.600 (0.534–0.667) | 0.021     | 0.618 (0.551–0.684) | 0.111     | 0.612 (0.603–0.620) | 371.180   |

* All models are continuous data except four-and-seven score.

^aP value for comparison with the “six-and-twelve” score.
and 7.6 (95% CI, 5.4–14.9) months (Figure 3), with 1-year survival probabilities of 75.6%, 57.6%, and 35.8%, respectively.

**Comparison of the performance and discrimination of the “six-and-twelve” score and other scores**

The performance and discrimination of the “six-and-twelve” score and other scores (up-to-seven criteria, four-and-seven criteria, HAP score, mHAP-II score, and ALBI score) were compared (Table 2). The AUROCs predicting death at 1, 2, and 3 years for the “six-and-twelve” score were 0.714, 0.700, and 0.688, respectively (Table 2). The 1-year AUROC value for the “six-and-twelve” score was significantly higher than those for the four-and-seven and ALBI scores and numerically higher than those for the HAP and mHAP-II scores (Table 2). As the backbone formula for calculation of “six-and-twelve” and up-to-seven criteria are the same, not surprisingly, the AUROCs predicting death for the “six-and-twelve” and up-to-seven score criteria displayed identically as 0.714, 0.700, and 0.688 at 1, 2, and 3 years, respectively. In addition, the discriminatory ability of the “six-and-twelve” score and up-to-seven criteria as calculated by the Harrell’s C-index demonstrated the highest value of 0.699. The C-index of the four-and-seven, HAP score, mHAP-II score, and ALBI score were 0.611, 0.674, 0.667, and 0.612, respectively.

The differences in the “six-to-twelve” score and up-to-seven criteria are the cutoffs between 2 scoring systems. Therefore, the sensitivity, specificity, positive predictive value, and NPV of those scores in predicting survival were evaluated further (Table 3). The “six-and-twelve”, model cutoff 6, model cutoff 12, and up-to-seven (model cutoff 7) had NPVs of 76.6%, 69.2%, and 77.1% in predicting 1-year mortality, respectively.

**DISCUSSION**

Recently, the “six-and-twelve” score was proposed to identify the best candidate for TACE (5). The score was the sum of a diameter of the largest tumor size and tumor number, stratified into 3 groups; ≤6, >6 but ≤12, and >12. This study is an external validation study of the “six-and-twelve” score conducted in a different population (11). Our study differed as it had variety etiologies of HCC. This reflected the real-life practice which had heterogeneous causes of HCC.

The performance and discriminatory capacity of this model is the most predictive compared with other currently available

| Models and cutoff values | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|-------------------------|----------------|----------------|---------|---------|
| “Six-and-twelve” (model cutoff 6) | 68.6 | 58.7 | 48.6 | 76.6 |
| “Six-and-twelve” (model cutoff 12) | 29.4 | 90.5 | 63.8 | 69.2 |
| Up-to-seven (model cutoff 7) | 64.7 | 67.6 | 53.2 | 77.1 |

NPV, negative predictive value; PPV, positive predictive value.
models. However, calculation of the “six-and-twelve” score is identical to that for the up-to-seven criteria because both scoring systems use the same parameters of tumor size and number in the score, and the final score is the summation of those parameters. Regarding the performance of these 2 models for 1-year survival prediction before TACE, the up-to-seven criteria with the cutoff value of 7 had the highest NPV. Interestingly, the original “six-and-twelve” score development study showed that the “six-and-twelve” score was better than the up-to-seven criteria. This might be due to the different designated analysis (5). We presupposed that the difference in AUROCs of the “six-and-twelve” score and the up-to-seven criteria in the original study is due to the researchers using the raw continuous score to calculate the AUROC for the “six-and-twelve” score, but used the category of score $\leq 7$ and $>7$ in the latter criteria, which would produce the different AUROC when transforming continuous score to categorical data.

It is well known that tumor burden as well as liver function are associated with survival rate of HCC patients (1,3,4). For the ideal TACE candidate, hepatic function should be preserved. In addition, the survival outcome depends on tumor burden. The “six-and-twelve” score and up-to-seven criteria were based on the largest tumor diameter and number of tumor(s) with different cutoff values. Several reports suggested that tumor burden is an independent predictor of survival (6,12).

The median OS for our cohort was lower than that for the cohort in the original study (20.3 vs 32.9 months) (5). In addition, by stratifying with the “six-and-twelve” score ($\leq 6$, $>6$–12, $>12$), median OS for our cohort again was lower than that in the original cohort in every stratum (our cohort, 35.1, 16.0, and 7.6 months vs original cohort, 49.1, 32.0, and 15.8 months, respectively) (5). Of interest, the reasons to explain the lower survival in our cohort compared with the Chinese cohort (5) might be from the different etiologies of HCC, and the lower proportions of a single tumor, HCC BCLC-A, and cirrhosis Child-Pugh score A in the study population. Of the eligible patients in our study, 49% had HBV infection compared with 85.2% in the original cohort. Notably, HBV-associated HCC patients had better survival times than those with other etiologies from a large database study (13). Moreover, in our study, lower proportions of HCC BCLC-A and cirrhosis Child-Pugh score A were observed in contrast to the original study, which might result in the lower survival rate as tumor stage and preserved liver function of the patients are important predictive factors of survival (14). However, our cohort had a similar median OS (19.8 months) as that from a large systematic review including 10,108 patients who underwent TACE (15). In addition, our data are consistent with those of the previous meta-analyses (16,17).

Our study has notable strengths. First, this study is a real-world practice in which the etiology of HCC was heterogeneous, in contrast to the population in the originally derived “six-and-twelve” score. Most HCC cases were due to HBV infection. Second, to the best of our knowledge, this is the first external validation study of the score outside China (5,18).

We also recognized some limitations in this study. First, our study was conducted based on a single tertiary referral hospital-based cohort, and its results are not generalized to other populations. Secondly, data on viral suppression status of viral hepatitis were not collected in this study, in which recent data suggested that antiviral therapies improved hepatic function in patients with viral-associated HCC (19,20). However, most of our patients who were ideal candidates for TACE had well-preserved liver function at baseline.

In conclusion, the OSs were different in each “six-and-twelve” score strata. Our study confirmed the predictive value of the “six-and-twelve” score in predicting survival of patients with resectable HCC treated with TACE. However, in our validation cohort, the AUROC of the “six-and-twelve” score was slightly lower than that of the original Chinese cohort (AUROC 0.73). The “six-and-twelve” score, despite being the most predictive score currently, may not be the ideal score to predict OS in TACE-treated HCC patients, as the AUROC is in acceptable but not excellent range. Further studies to develop new and more predictive scores should be considered.

CONFLICTS OF INTEREST

Guarantor of the article: Pimsiri Sripongpun, MD.
Specific author contributions: A.K. contributed to the study concept and design, collecting data, analysis and interpretation of data, and drafting of the manuscript. P.S. contributed to the analysis, interpretation of data, and critical revision of the manuscript. N. Cheewasereechon contributed to data collection and drafting the manuscript. S.J. and N. Chamroonkul contributed to the interpretation of data and critical revision of the article. T.P. contributed to the study concept and design and supervised the study. All authors contributed to critical revisions and approved the final manuscript.

Financial support: This work was supported by the grant from Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand.
Potential competing interests: A.K., P.S., N. Cheewasereechon, S.J., and N. Chamroonkul have nothing to declare. T.P. has received research grants from Gilead Sciences, Roche Diagnostic, Janssen, Fibrogen, and VIR and speaker honorarium from Bristol-Myers Squibb, Gilead Sciences, Bayer, Abbott, Esai, Mylan, Ferring, and MSD. All investigators had access to the study data, reviewed, and approved the final manuscript.

Study Highlights

### WHAT IS KNOWN

- The “six-and-twelve” prognostic score had been introduced to predict the best candidates for transarterial chemoembolization (TACE).
- No study exists for external validation of the recently reported prognostic “six-and-twelve score” in hepatocellular carcinoma (HCC) patients in other countries outside of China.

### WHAT IS NEW HERE

- The performance of this model is the most predictive compared to other currently available models except the up-to-seven model.
- The up-to-seven criteria with the cutoff value of 7 had the highest NPV.
- The median overall survival (OS) for our cohort was lower than that for the cohort in the original study.

### TRANSLATIONAL IMPACT

- The “six-and-twelve” score, despite being the most predictive score currently, may not be the ideal score to predict OS in TACE-treated HCC patients.
REFERENCES
1. Omata M, Cheng A-L, Kokudo N, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: A 2017 update. Hepatol Int 2017;11(4):317–70.
2. Chonprasertsuk S, Vilaichone R. Epidemiology and treatment of hepatocellular carcinoma in Thailand. Jpn J Clin Oncol 2017;47(4):294–7.
3. Heimbach JK, Kulik LM, Finn RS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatol Baltim Md 2018;67(1):358–80.
4. Galle PR, Forner A, Llovet JM, et al. EASL clinical practice guidelines: Management of hepatocellular carcinoma. J Hepatol 2018;69(1):182–236.
5. Wang Q, Xia D, Bai W, et al. Development of a prognostic score for recommended TACE candidates with hepatocellular carcinoma: A multicentre observational study. J Hepatol 2019;70(5):893–903.
6. Mazzaferro V, Llovet JM, Miceli R, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: A retrospective, exploratory analysis. Lancet Oncol 2009;10(1):35–43.
7. Yamakado K, Miyayama S, Hirota S, et al. Subgrouping of intermediate-stage (BCLC stage B) hepatocellular carcinoma based on tumor number and size and Child-Pugh grade correlated with prognosis after transarterial chemoembolization. Jpn J Radiol 2014;32(5):260–5.
8. Kadalayil I, Benini R, Pallan L, et al. A simple prognostic scoring system for patients receiving transarterial embolisation for hepatocellular cancer. Ann Oncol 2013;24(10):2565–70.
9. Park Y, Kim SU, Kim BK, et al. Addition of tumor multiplicity improves the prognostic performance of the hepatoma arterial-embolization prognostic score. Liver Int 2016;36(1):100–7.
10. Pinao DJ, Sharma R, Allara E, et al. The ALBI grade provides objective hepatic reserve estimation across each BCLC stage of hepatocellular carcinoma. J Hepatol 2017;66(2):338–46.
11. Wang Z-X, Li J, Wang E-X, et al. Validation of the six-and-twelve criteria among patients with hepatocellular carcinoma and performance score 1 receiving transarterial chemoembolization. World J Gastroenterol 2020;26(15):1805–19.
12. Takayasu K, Arii S, Kudo M, et al. Superselective transarterial chemoembolization for hepatocellular carcinoma. Validation of treatment algorithm proposed by Japanese guidelines. J Hepatol 2012;56(4):886–92.
13. Brar G, McNeel T, McGlynn K, et al. Hepatocellular carcinoma (HCC) survival by etiology: A SEER-Medicare database analysis. J Clin Oncol 2019;37(4 Suppl):201.
14. Wang C-Y, Li S. Clinical characteristics and prognosis of 2887 patients with hepatocellular carcinoma: A single center 14 years experience from China. Medicine (Baltimore) 2019;98(4):e14070.
15. Lencioni R, de Baere T, Soulen MC, et al. Lipiodol transarterial chemoembolization for hepatocellular carcinoma: A systematic review of efficacy and safety data. Hepatology 2016;64(1):106–16.
16. Llovet JM, Bruix J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: Chemoembolization improves survival. Hepatol Baltim Md 2003;37(2):429–42.
17. Cama R, Schepis F, Orlando A, et al. Transarterial chemoembolization for unresectable hepatocellular carcinoma: Meta-analysis of randomized controlled trials. Radiology 2002;224(1):47–54.
18. Wang Z-X, Wang E-X, Bai W, 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(6):657–69.
19. Colombo M, Lleo A. The impact of antiviral therapy on hepatocellular carcinoma epidemiology. Hepatic Oncol 2018;5(1):HEP03.
20. Laursen TL, Sandahl TD, Kazankov K, et al. Liver-related effects of chronic hepatitis C antiviral treatment. World J Gastroenterol 2020;26(22):2931–47.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.