Received: 6 May 2021 | Revised: 27 June 2021 | Accepted: 1 July 2021

DOI: 10.1002/lci2.33

ORIGINAL ARTICLE

Liver Cancer INTERNATIONAL WILEY

Recalibrating survival prediction among patients receiving trans-arterial chemoembolization for hepatocellular carcinoma

Alessandro Cucchetti[1] | Edoardo G. Giannini[2] | Cristina Mosconi[3] | Maria Corina Plaz Torres[2] | Giulia Pieri[2] | Fabio Farinati[4] | Gian Ludovico Rapaccini[5] | Maria Di Marco[6] | Eugenio Caturelli[7] | Rodolfo Sacco[8] | Giuseppe Cabibbo[9] | Claudia Campani[10] | Andrea Mega[11] | Maria Guarino[12] | Antonio Gasbarrini[13] | Gianluca Svegliati-Baroni[14] | Francesco Giuseppe Foschi[15] | Gabriele Missale[16] | Alberto Masotto[17] | Gerardo Nardone[18] | Giovanni Raimondo[19] | Gianpaolo Vidili[20] | Maurizia Rossana Brunetto[21] | Vito Sansone[22] | Marco Zoli[23] | Francesco Azzaroli[24] | Franco Trevisani[25] | the ITA.LI.CA Study Group

[1]Department of Medical and Surgical Sciences, DIMEC, Alma Mater Studiorum, University of Bologna, Bologna, Italy

[2]Gastroenterology Unit, Department of Internal Medicine, University of Genova, IRCCS Ospedale Policlinico San Martino, Genova, Italy

[3]Department of Specialist, Diagnostic and Experimental Medicine, Radiology Unit, IRCCS, Azienda Ospedaliero-Universitaria di Bologna, Sant'Orsola-Malpighi Hospital, Bologna, Italy

[4]Gastroenterology Unit, Department of Surgery, Oncology and Gastroenterology, University of Padova, Padova, Italy

[5]Gastroenterology Unit, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Roma, Italy

[6]Medicine Unit, Bolognini Hospital, Seriate, Italy

[7]Gastroenterology Unit, Belcolle Hospital, Viterbo, Italy

[8]Gastroenterology and Digestive Endoscopy Unit, Foggia University Hospital, Foggia, Italy

[9]Gastroenterology & Hepatology Unit, Department of Health Promotion, Mother & Child Care, Internal Medicine & Medical Specialties, PROMISE, University of Palermo, Palermo, Italy

[10]Internal Medicine and Hepatology Unit, Department of Experimental and Clinical Medicine, University of Firenze, Firenze, Italy

[11]Gastroenterology Unit, Bolzano Regional Hospital, Bolzano, Italy

[12]Gastroenterology Unit, Department of Clinical Medicine and Surgery, University of Napoli "Federico II", Napoli, Italy

[13]Internal Medicine and Gastroenterology Unit, Policlinico Gemelli, Università Cattolica del Sacro Cuore, Roma, Italy

[14]Gastroenterology Unit, Polytechnic University of Marche, Ancona, Italy

[15]Department of Internal Medicine, Ospedale per gli Infermi di Faenza, Faenza, Italy

[16]Infectious Diseases and Hepatology Unit, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy

[17]Gastroenterology Unit, Ospedale Sacro Cuore Don Calabria, Negrar, Italy

[18]Hepato-Gastroenterology Unit, Department of Clinical Medicine and Surgery, University of Napoli "Federico II", Napoli, Italy

[19]Clinical and Molecular Hepatology Unit, Department of Clinical and Experimental Medicine, University of Messina, Messina, Italy

[20]Clinica Medica Unit, Department of Medical, Surgical and Experimental Sciences, University of Sassari, Azienda Ospedaliero-Universitaria di Sassari, Sassari, Italy

[21]Department of Clinical and Experimental Medicine, Hepatology and Liver Physiopathology Laboratory and Internal Medicine, University of Pisa, Pisa, Italy

[22]Internal Medicine–Piscaglia Unit, Department of Medical and Surgical Sciences, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

[23]Internal Medicine–Zoli Unit, Department of Medical and Surgical Sciences, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

Abbreviations: AFP, alpha-fetoprotein; CI, confidence intervals; HAP, hepatoma arterial-embolization prognostic; HCC, hepatocellular carcinoma; ITA.LI.CA, Italian Liver Cancer; TACE, trans-arterial chemoembolization.

A. Cucchetti and E. G. Giannini share first co-authorship.

[24]Gastroenterology Unit, Department of Surgical and Medical Sciences, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
[25]Medical Semeiotics Unit, Department of Medical and Surgical Sciences, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

Correspondence
Edoardo G. Giannini, Gastroenterology Unit, Department of Internal Medicine, University of Genoa, Viale Benedetto XV, no. 6, 16132, Genoa, Italy.
Email: egiannini@unige.it

Abstract

Background & Aims: The Pre-TACE-Predict model was devised to assess prognosis of patients treated with trans-arterial chemoembolization (TACE) for hepatocellular carcinoma (HCC). However, before entering clinical practice, a model should demonstrate that it performs a useful role.

Methods: We performed an independent external validation of the Pre-TACE model in a cohort that differs in setting and time period from the one that generated the original model. Data from 826 patients treated with TACE for naïve HCC (2008-2018) were used to assess calibration and discrimination of the Pre-TACE-Predict model.

Results: The four risk-categories identified by the Pre-TACE-Predict model had gradient monotonicity, with median survivals of 52.0, 36.2, 29.9, and 14.1 months respectively. However, predicted survivals systematically underestimated observed survivals (R^2: 0.667). A recalibration was adopted maintaining fixed the prognostic index and modifying the baseline survival function. This resulted in an almost perfect calibration (R^2: 0.995) in all the four risk categories. Cox regressions showed that aetiology and macrovascular invasion, included in the Pre-TACE-Predict model, had no prognostic impact in the present study population, and that coefficients for tumour size and multiplicity were overestimated. The *c*-index was similar to that of the m-HAP-III, but higher than those of HAP, m-HAP-II and the six-and-twelve models.

Conclusions: The recalibration of Pre-TACE-Predict model improved the estimation of survival probabilities of HCC patients treated with TACE. The highest discriminatory ability of the Pre-TACE-model in comparison to other available models, together with risk stratification and recalibration, makes it the best prognostic tool currently available for these patients.

KEYWORDS
hepatocellular carcinoma, prediction model, recalibration in the large, survival, trans-arterial chemoembolization

1 | INTRODUCTION

The prognostic classification of patients with hepatocellular carcinoma (HCC) is problematic since overall survival is influenced by tumour burden, residual liver function, and patient performance status. Over the last few years, several systems were developed to predict the prognosis of patients undergoing trans-arterial chemoembolization (TACE).[1-9] Among all the proposed, the Pre-TACE-Predict model is the largest and most comprehensive model, applicable prior to treatment fulfilling decision-making needs.[9]

The Pre-TACE-Predict model includes tumour features [number and largest diameter, presence/absence of macrovascular invasion, and alpha-fetoprotein (AFP) values], liver tests (albumin and bilirubin) and, unique among the available prognostic models, aetiology of liver disease (hepatitis virus B, C, and alcohol). This peculiarity is important considering that the Pre-TACE-Predict model was generated including both Eastern and Western patients, thus potentially accounting for the variability related to the worldwide epidemiologic features of HCC. However, one model may not necessarily fit every different population. Indeed, in its external validation process, the Pre-TACE-Predict model showed better accuracy in Eastern than in Western patients. In this regard, given the potential use of this clinical decision aid, its generalized use in the clinical practice should follow the demonstration of a reliable prognostic capability in independent series, especially in Western populations.[10-12]

In the present study, we performed an independent external validation applying the most severe level of stringency, that differ

in terms of investigators, location and recruitment calendar period from those which produced the original model.[11,12]

2 | METHODS

2.1 | Study population

Data derived from the Italian Liver Cancer (ITA.LI.CA) registry, which prospectively collected data of HCC patients diagnosed and treated in 24 centres across Italy since 1987. The ITA.LI.CA database conforms to the current Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data, and to the ethical guidelines of the Declaration of Helsinki. From a total of 7816 registry' entries, 5196 diagnosed between January 2008 and December 2018 were retained for the present analysis. Of these, 4142 patients who did not receive TACE as the main therapy for a naïve HCC were excluded. This exclusion step contemplates that TACE was the main therapy adopted in a hierarchical therapeutic order, so that patients who received TACE before other potentially curative therapies, including liver transplantation, were not included in the analysis, since these patients have not TACE registered as the main therapy.[13,14] Of the 1054 selected patients, 193 were excluded because data were shared with the pre-TACE model published by Han et al, and additional 35 cases were excluded because of an incomplete data entry.[9] Therefore, the study population consisted of 826 patients treated with TACE for naïve HCC, not shared with the previously published model and with all necessary data for the planned validation study.

2.2 | Description of the pre-TACE model

The Pre-TACE-Predict model includes variables available before treatment such as tumour number, size, AFP, albumin, bilirubin, absence/presence of macrovascular invasion, hepatitis B, hepatitis C and alcohol as aetiologic factors.[9] The model derived from a set of 1714 patients, internally validated in another 1714 patients, and externally validated in 407 Western and 786 Eastern patients. Details of the formula are described in the Appendix section (Supplementary material) and it can be accessed at: https://prediction-models.lirerpool.ac.uk/tace. In its continuous form, the Harrell's *C* index in the external Western cohort was 0.613 [standard error (SE): 0.017] with a Gönen & Heller's *K*-index of 0.587 (SE: 0.016).

On the basis of the linear predictor derived from the model, four risk categories were identified: ≤0.94 (risk category #1), >0.94 to ≤1.47 (risk category #2), >1.47 to ≤2.10 (risk category #3) and >2.10 (risk category #4). In the external Western validation, the corresponding median survivals were 34.8, 24.0, 17.1 and 8.3 months respectively. Predicted survival of each category remained well within the 95% confidence intervals (CI) of the observed Kaplan–Meier rates, providing evidence for good calibration of the stratification.[9]

KEY POINTS

Trans-arterial chemoembolization (TACE) is a palliative treatment commonly used in clinical practice for the treatment of primary liver cancer. The prognosis of patients undergoing TACE is quite heterogeneous due to a number of factors, mainly related to both patients and tumour characteristics. The Pre-TACE is a prognostic model that has shown to be able to accurately predict prognosis in a large series of patients undergoing TACE. However, before entering clinical practice, a model should demonstrate that it performs a useful role. In this study, we performed an independent external validation of the Pre-TACE model in a cohort that differs in setting and time period from the one that generated the original model. We observed that a recalibration of Pre-TACE-Predict model improved the estimation of survival probabilities of HCC patients treated with TACE, and that, in comparison to other available models, it is the best prognostic tool currently available for these patients.

3 | METHODOLOGY

In prognostic models, patients and clinicians are interested in the future risk of the disease rather than the probability of a positive test.[15,16] Discrimination serves to separate patients with and without the outcome, and the discriminatory ability is important in diagnostic settings, where a separation between people with or without the disease according to the test result or model score is needed. Calibration concerns the agreement between the observed and the predicted risk, and a good calibration is very important in prognostic settings where a prediction of the future risk of a specific patient group is desirable. Models with adequate calibration by the predicted risk strata provide useful information for medical decision making.[17]

On this background, our primary endpoint was to assess calibration of the Pre-TACE-Predict model in a large, independent, external validation cohort. In order to do so, the predicted 6-, 12-, 24- and 36-month survival rates were computed for all patients and median values of risk categories were compared with survival rates from Kaplan–Meier curves. Relationships between median predicted and observed values were graphically inspected and measured through R^2 calculation. Eventual need for recalibration was accomplished through the 'calibration-in-the-large', avoiding modifying the linear predictor achieved by refitting the baseline survival rates for the present data as described by Royston and Demler.[11,18]

The second aim of the study was to verify the discrimination ability of the Pre-TACE-Predict model. To this end, the non-stratified linear predictor was measured with the common Harrell's *C* index and Gönen & Heller's *K* index. In particular, this latter indicator is considered a more robust concordance index than the Harrell's *C*

index, since it relaxes the proportion of censored cases.[19] In the presence of an elevated number of censored events, the Harrell's *C* index increases deceptively.[20]

Regression coefficients for one or more covariates in the Pre-TACE-Predict model may differ between original derivation and the present external validation datasets. This was formally tested by running a first Cox regression on the covariates forming the Pre-TACE-Predict model in the current dataset (using clustering to account for the multi-institutional origin of data) and comparing the obtained coefficients to those previously published. Subsequently, a second identical Cox regression was adopted 'offsetting' the original PI evaluated in the validation dataset, so that the coefficient of PI was constrained to equal 1. In a hypothetical perfect model, these latter variable coefficients would be 0.[11,18] All the analyses were conducted in STATA (16.1, StataCorp LLC).

4 | RESULTS

Baseline characteristics of the 826 patients included in the study are reported in Table 1. The main differences in our series as compared to the original series by Han et al were the higher prevalence of HCV patients, the lower prevalence of macroscopic vascular invasion, and the overall smaller tumour burden.[9] During a median follow-up of 44.3 months [interquartile range (IQR): 23.2-70.4 months], 426 patients died (51.6%) with 400 censored events (48.4%). The median survival was 37.2 months (IQR, 19.1-77.3 months), and the 6-, 12-, 24- and 36-month survival rates were 97.1%, 88.1%, 67.6% and 51.9% respectively.

The median value of Pre-TACE-Predict linear predictor (PI) was 1.23 (IQR, 0.92-1.57). Applying the proposed risk classes, 229 patients were in category #1 (27.7%), 347 in category #2 (42.0%), 199 in category #3 (24.1%) and 51 in category #4 (6.2%). Kaplan–Meier survival rates are reported in Figure 1. Median survivals for each category were 52.0 (IQR, 29—not reached), 36.2 (IQR, 20.1-77.3), 29.9 (IQR, 15.1-52.8) and 14.1 months (IQR, 8.1-7.1) respectively.

4.1 | Calibration

Calibration of the Pre-TACE-Predict model is depicted in Figure 2A. The R^2 calculated between the observed survival rates and those predicted by Pre-TACE-Predict model was 0.667. Notably, observed survival rates were always higher than the predicted ones, suggesting a systematic underestimation by the prediction model. Predicted versus observed survival rates by risk groups are reported in Figure S1.

Recalibration through 'calibration-in-the-large' was applied refitting the baseline survival for the present data and the following $S_0(t)$ values were obtained: 0.97, 0.86, 0.63 and 0.48 for probabilities at 6-, 12-, 24- and 36-month respectively. Applying these new $S_0(t)$ values, calibration remarkably ameliorated (Figure 2B) and the R^2 increased to 0.995. Predicted versus observed survival rates by risk groups after recalibration are reported in the Figure S2.

TABLE 1 Clinical characteristics at diagnosis of 826 patients treated with TACE for HCC between 2008 and 2018

Variables	All patients (n = 826)
Age [years (median; IQR)]	70 (63-76)
Male (%)	645 (78.1)
Aetiology[a](%)	
Hepatitis C	428 (51.8)
Hepatitis B	83 (10.1)
Alcohol	258 (31.2)
Other	126 (15.3)
Symptoms at diagnosis	64 (7.9)
ECOG PS (%)	
0	597 (72.3)
1	186 (22.5)
2	43 (5.2)
Albumin [g/L (median; IQR)]	36 (32-39)
Bilirubin [μmol/L (median; IQR)]	18.8 (13.7-27.4)
MELD (median; IQR)	9 (8-11)
Ascites at diagnosis (%)	909 (19.0)
Child-Pugh class (%)	
A	590 (71.4)
B	221 (26.8)
C	15 (1.8)
Nodular pattern (%)	
Single lesion	414 (50.1)
Up to 3 lesions	360 (43.6)
>3 lesions	52 (6.3)
Largest diameter [cm (median; IQR)]	3.2 (2.4-4.5)
Macrovascular invasion (%)	33 (4.0)
AFP [ng/mL (median; IQR)]	25 (6-90)

Note: Abbreviations: AFP, alpha-fetoprotein; ECOG PS, Eastern Cooperative Oncology Group Performance Status; MELD, model for end-stage liver disease.

[a]One patient can have more than one cause of liver disease, so the sum of aetiologies did not necessarily sum to 100%.

4.2 | Discrimination

In the present study population, the Harrell's *C* index of the linear predictor (PI) of the Pre-TACE-Predict model was 0.625 and the Gonen and Heller's *K* was 0.599. Hazard ratios (HRs) of components of the Pre-TACE-Predict model estimated here and compared to those published from the original model are reported in Table 2, together with coefficients obtained when the linear predictor was also included and constrained to 1.

Alpha-fetoprotein (HR: 1.27; 95% CI: 1.12-1.45) and albumin (HR: 0.97; 95%CI: 0.95-0.99) confirmed their prognostic role with

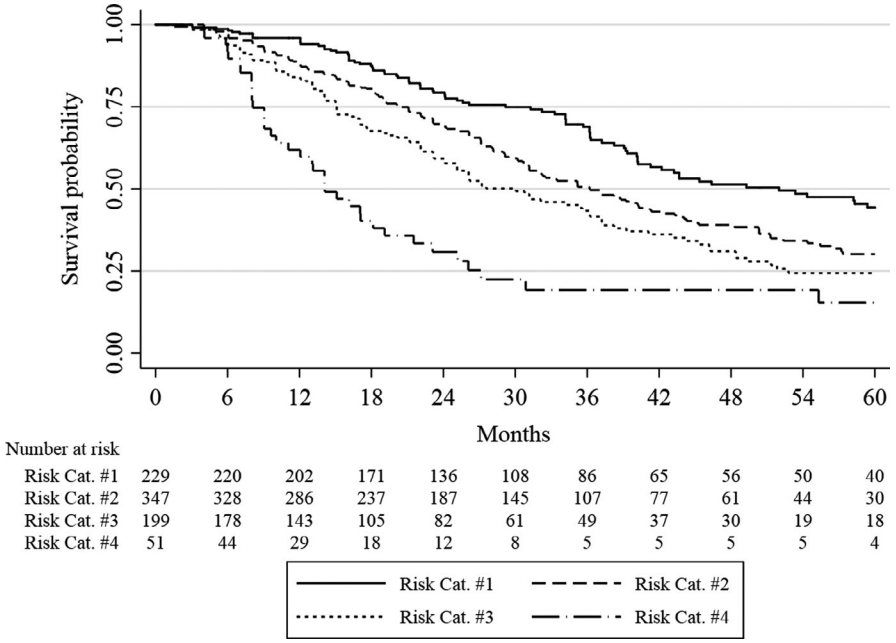

FIGURE 1 Kaplan–Meier survival curves of the four risk groups identified by the Pre-TACE-Prediction model

values very similar to those published. Offsetting the linear predictor returned coefficients of 0.01 for both, indicating that the Pre-TACE-Predict model had an almost perfect prognostic accuracy for these two parameters. Tumour size (HR: 2.27; 95% CI, 1.39-3.71) and multifocality (HR: 1.19; 95% CI, 1.02-1.49) also determined overall survival, but their HRs were lower than those of the Pre-TACE-Predict model (coefficients: −0.43 and −0.18, respectively). Bilirubin was another prognostic indicator (HR: 1.78; 95% CI, 1.28-2.46) but with a HR higher than that of the Pre-TACE-Predict model (coefficient: 0.12). Conversely, macrovascular invasion and aetiology of liver disease had no prognostic relevance in the present external validation.

4.3 | Comparison with other scores

To evaluate if other available models discriminate better than the Pre-TACE-Predict model ($C = 0.625$, $K = 0.599$), the linear predictors of hepatoma arterial-embolization prognostic (HAP), m-HAP-II, m-HAP-III and of the six-&-twelve models were also evaluated in this series (Table 3).

Compared to the Pre-TACE-Predict model, the linear predictor of the m-HAP-III model showed very similar Harrell's index ($C = 0.630$) but lower Gonen and Heller's value ($K = 0.579$). The HAP model showed both lower Harrell's index ($C = 0.587$) and Gonen and Heller's value ($K = 0.574$), and these figures were similar to those of the m-HAP-II model ($C = 0.582$, $K = 0.573$). Lastly, the six-&-twelve model showed a Harrell's index ($C = 0.588$) similar to HAP and m-HAP-II models but had the lowest Gonen and Heller's value ($K = 0.565$).

The comparison of Harrell's C indexes among the models evaluated disclosed that the Pre-TACE-Predict and the m-HAP-III models had similar discriminant ability, which was superior to those of HAP, m-HAP-II and six-&-twelve models (Table S1).

5 | DISCUSSION

Prognostic models should meet several needs of clinical care. They can be used to inform patients and their families about the expected disease course, to drive decisions on the treatment that can benefit most the patient and to stratify patients by disease severity when planning clinical trials.[11,12] Suitable prognostic models in the setting of TACE are especially useful in the choice of the treatment, considering the available therapeutic alternatives.[21]

Options for validation depend on the available information describing results from the derivation datasets. The current literature pertinent on TACE survival prediction models showed the availability of four levels of information: (a) the description of PI resulting from multivariable regression coefficients; (b) the generation of risk categories; (c) the presence of stratified Kaplan–Meier curves; (d) the description of baseline survival function to allow for model reproducibility. Among all published models, only the Pre-TACE-Predict one fulfils all of them,[9] whereas the remaining models lack one or more of these aspects: HAP and m-HAP-II only described the PI and then formed four different risk groups, without calibrations,[2,5] and the m-HAP-III lacks risk categorization and consequently stratified survival curves.[6] Lastly, the six-&-twelve model provided a nomogram for survival prediction without reporting baseline survival function.[8] Therefore, the Pre-TACE-Predict model is the best described one, allowing for both external discrimination and, more importantly, calibration.

The distinction of the four risk categories provided by the Pre-TACE-Predict model well stratified patients with different survival probabilities (Figure 1). However, we noted that the corresponding median survivals were consistently lower than those published in the external Western validation cohort.[9] In that cohort, median survivals of the four categories were 34.8, 24.0, 17.1 and 8.3 months, respectively, while our corresponding median survivals were 52.0,

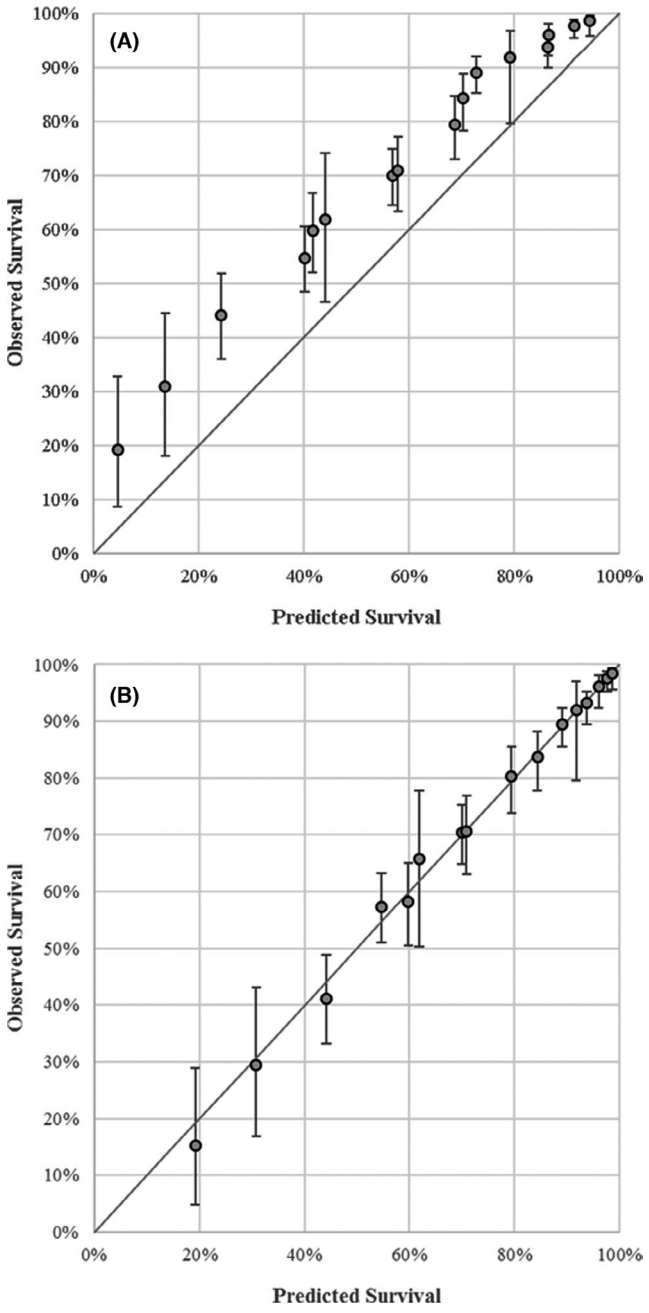

FIGURE 2 A, *R*-squared of observed versus predicted survival rates. As noticeable, predicted values were always lower than the observed suggesting a systematic underestimation of the prediction model. B, After recalibration, predicted values were very near to the observed ones

36.2, 29.9 and 14. months. This was not an unexpected finding, as the therapeutic landscape in any condition progresses (eg improvement in treatment of the underlying liver disease, or of TACE technique) the predicted survival of a model may underestimate the observed survival in more recent populations. This difference, together with the observed monotonicity of our Kaplan–Meier curves, indicates that the Pre-TACE-Predict model was miscalibrated in our external validation cohort. In this regard, it should be observed that calibration performance of prediction models based on regression algorithms commonly receives little attention, despite the fact that poorly calibrated algorithms can be misleading and potentially harmful for clinical decision-making. In such situations, the model update is indicated.[22] The observation that predicted values were systematically lower than the observed ones (Figure 2A and Figure S1) allowed us to use the 'calibration-in-the-large' which provided the necessary accuracy in survival prediction to the model.[23] Final results indicated that baseline survival rates at 6, 12, 24 and 36 months after TACE are 0.97, 0.88, 0.63 and 0.48 respectively, and that they can adequately predict our observed events. Nevertheless, it is possible that these expected survival rates are inaccurate for the published internal and external cohorts of the original model devised in the Han et al publication.[9]

Therefore, we thought that some solutions might be suggested in order to improve the fine tuning of the model. A practical solution would rely on the overall patient survival after TACE observed in the health care system in which the clinician should decide whether or not to submit the patient to TACE. The median overall survival of the whole Pre-TACE-Predict cohort was 19.9 months, whereas it was 37.2 months in our cohort. The knowledge of median survival of their TACE patients allows clinicians to decide to adopt the original Pre-TACE-Predict model (in the case of similar low median survival), or the recalibrated model (in the case of higher median survival). In other words, if the decision-maker feels that the original Pre-TACE-Predict model may underestimate survival probabilities, he/she can shift to the recalibrated model. Both calculators can be found here: https://jscalc.io/calc/2omTfeWrmOLc41ei.

An interesting finding of our study is the improved survival of our patients compared to the one observed in the cohort that generated the Pre-TACE-Predict model, which was the main cause of the systematic underestimation of survival in each risk stratum. There are several possible explanations for this finding. Firstly, when deciding the level of stringency of our external validation, we examined a time period different from the one of the original model and in particular 75.5% of data used to generate the Pre-TACE-Predict model refer to HCC cases managed prior to January 2012, whereas this proportion was 29.3% in our cohort. The survival of the present cohort is even longer than the one reported in the m-HAP-III cohort (24.6 months), including similar patients but dated to the period 2000-2012.[6] These observations remark the improved survival observed after TACE over time, thanks to a better patient selection and the continuous technical refinements of intra-arterial treatments.[24,25] Moreover, our cohort was formed by 51.8% of patients with chronic hepatitis C infection, and about half of them received TACE in the Direct Acting Antivirals (DAA) era. This is the highest proportion of patients with hepatitis C virus infection among centres included in the Pre-TACE-Predict model so that it is conceivable that a higher proportion of our HCC patients benefited from the effects of the sustained viral response provided by DAA (administered before tumour detection or after a complete radiological response) on cirrhosis progression and amenability to either re-treatment or sorafenib in the case of disease progression. Although the survival benefit of sorafenib after

TABLE 2 Results from multivariable Cox regressions of the pre-TACE-predict parameters in the present study population

Pre-TACE-Predict parameters	Present validation HR (95% CI)	Published data HR (95% CI)	Offsetting linear predictor Coeff. (95% CI)
Multifocal	1.19 (1.02-1.49)	1.35 (1.13-1.61)	−0.18 (−0.36 to 0.09)
Size (log 10, cm)	2.27 (1.39-3.71)	4.08 (3.16-5.27)	−0.43 (−0.92 to 0.06)
AFP (log 10, ng/ml)	1.27 (1.12-1.45)	1.34 (1.27-1.43)	−0.01 (−0.11 to 0.14)
Albumin (g/L)	0.97 (0.95-0.99)	0.99 (0.97-1.01)	−0.01 (−0.03 to 0.01)
Bilirubin (log 10, μmol/L)	1.78 (1.28-2.46)	1.43 (1.02-2.09)	0.12 (−0.21 to 0.44)
Vascular invasion	1.05 (0.58-1.88)	2.38 (1.86-3.05)	−0.48 (−1.07 to 0.10)
Aetiology			
HCV	Ref.	Ref.	Ref
HBV	1.13 (0.85-1.49)	1.46 (1.21-1.77)	−0.03 (−0.31 to 0.25)
Alcohol	1.07 (0.79-1.46)	1.47 (1.19-1.83)	−0.26 (−0.57 to 0.04)
Other	1.15 (0.89-1.50)	1.55 (1.22-1.97)	−0.07 (−0.34 to 0.19)

Abbreviation: HR, hazard ratio.

Present HR were compared with published data. Offsetting the linear predictor means that the PI from the Pre-TACE-Predict model was constrained to 1 to evaluate the deviance of the variable coefficients from the 0. In a hypothetical perfect model, these latter coefficients would be 0.

TABLE 3 Discrimination indexes of various pretreatment models

Model	Harrell's *C*	Gonen and Heller's *K*	6-month AUC	12-month AUC	24-month AUC	36-month AUC
Pre-TACE-Predict	0.625 (0.015)	0.599 (0.012)	0.655 (0.063)	0.698 (0.029)	0.654 (0.023)	0.631 (0.033)
m-HAP-III	0.630 (0.015)	0.579 (0.010)	0.698 (0.057)	0.715 (0.028)	0.664 (0.022)	0.654 (0.023)
HAP	0.587 (0.016)	0.574 (0.013)	0.722 (0.059)	0.676 (0.030)	0.599 (0.023)	0.576 (0.032)
m-HAP-II	0.582 (0.016)	0.573 (0.013)	0.718 (0.055)	0.651 (0.032)	0.613 (0.023)	0.586 (0.032)
Six-&-Twelve	0.588 (0.016)	0.565 (0.012)	0.606 (0.064)	0.634 (0.031)	0.622 (0.023)	0.612 (0.024)

Note: Abbreviation: HAP, hepatoma arterial-embolization prognostic.

Models are ordered on the basis of their *K*-values. Standard errors are reported in parentheses.

TACE is currently under investigation, an improvement in time-to-progression has been reported.[26-29] Therefore, the combination of technique refinement, a high prevalence of hepatitis C virus positive patients and the DAA-induced clearance of the infection in some of them, and, eventually, the sequential use of sorafenib can be considered the reasons for the long survival observed in our TACE population.

Another cause of underestimation by the Pre-TACE-Predict model might be identified in the low prevalence of macroscopic vascular invasion in our cohort (4%). The inclusion of this unfavourable variable into the Pre-TACE-Predict model, together with the null effect of aetiology, could have upsized the relative 'effect size' of the remaining parameters, producing higher coefficients and hence lowering predicted survivals. The recalculation of the HRs for tumour size and multiplicity yielded considerably lower values than the original ones (Table 2), being only the HRs of AFP and albumin very similar to those of the Pre-TACE-Predict model. Conversely, a more weight was observed for bilirubin, but despite this the Pre-TACE-Predict model underestimated, thus tumour features probably affect prognosis more than liver function tests. All in all, we feel that the lower weight attributed by the recalibrated model to macroscopic vascular invasion and aetiology of liver disease might more closely reflect the population of patients who in the future will be treated with TACE, as the widespread treatment of HCV patients with DAA will likely equalize this variable by avoiding the risk of death due to liver decompensation following successful oncological treatment, while the advent of other therapeutic options with competitive overall survival rates and a lower likelihood of post-treatment hepatic decompensation—such as trans-catheter arterial radio-emobolization and new systemic treatments—will decrease the proportion of patients with macroscopic vascular invasion undergoing TACE.[30-33]

With the exception of aetiology and macrovascular invasion, the remaining prognostic factors (multiplicity, size, AFP, albumin and bilirubin) are present even in the m-HAP-III model. Consequently, it was not surprising that the discrimination ability of the Pre-TACE-Predict model and of the m-HAP-III was very similar. However, the Pre-TACE-Predict model overcomes the limit of the lack of risk categorization of the m-HAP-III model, maintaining a similar *C* index (and better Heller's *K*) and providing good prediction after recalibration. Moreover, considering that the Pre-TACE-Predict model was also superior to HAP, m-HAP-II and the six-&-twelve models, it can be stated that it is the best prognostic model currently available for HCC patients undergoing TACE.

The present study has some limitations to be acknowledged. First, data on DAA administration and achievement of sustained virological response in patients with chronic hepatitis C were not available. Therefore, its favourable impact on the survival of our patients remains speculative. Second, we considered only the patient status before the first TACE, as it was the indispensable pre-requisite to test the Pre-TACE-Predict model. The eventual availability of data regarding radiological response would have allowed the evaluation of the Post-TACE-Predict model. However, such radiological data were not available for most of the patients, referring this task to another possible study. Last, we excluded patients who underwent liver transplantation after TACE which was used as a 'bridge' to surgery. This selection, necessary to fulfil the Pre-TACE-Predict inclusion criteria, can have modified the characteristics of patients who potentially benefit from TACE. However, it can be confidently assumed that the proportion of patients within the registry who were transplanted after TACE was minimal.[13]

In conclusion, our study provided an external validation of the Pre-TACE-Predict model, which is the most recently proposed and comprehensive prognostic model for HCC patients treated with TACE. However, due to the improved outcome of TACE treatment observed in recent years, a recalibration of the baseline survival function was needed to optimize the estimation of survival probabilities. As a result, the highest discriminatory ability of the Pre-TACE-Predict model in comparison to the other models, together with its risk stratification and recalibration, makes it the best prognostic instrument we currently have to predict the TACE outcome in HCC patients.

ACKNOWLEDGEMENTS

Others members of the ITA.LI.CA group: Semeiotics Unit, Department of Medical and Surgical Sciences and Azienda Ospedaliero-Universitaria di Bologna: Maurizio Biselli, Paolo Caraceni, Annagiulia Gramenzi, Davide Rampoldi, Nicola Reggidori, Valentina Santi, Benedetta Stefanini. Internal Medicine, Department of Medical and Surgical Sciences and Azienda Ospedaliero-Universitaria di Bologna: Alessandro Granito, Luca Muratori, Fabio Piscaglia, Francesco Tovoli, Donatella Magalotti. Gastroenterology Unit, Department of Surgical and Medical Sciences and Azienda Ospedaliero-Universitaria di Bologna: Elton Dajti, Giovanni Marasco, Federico Ravaioli. Department of Specialist, Diagnostic and Experimental Medicine, Radiology Unit, University of Bologna: Alberta Cappelli, Rita Golfieri, Cristina Mosconi, Matteo Renzulli. Gastroenterology Unit, Department of Surgery, Oncology and Gastroenterology, University of Padova: Filippo Pelizzaro, Barbara Penzo. Gastroenterology and Digestive Endoscopy Unit, Foggia University Hospital, Foggia: Ester Marina Cela, Antonio Facciorusso. Department of Internal Medicine, Gastroenterology Unit, University of Genova, IRCCS Ospedale Policlinico San Martino, Genova: Valentina Cacciato, Edoardo Casagrande. Gastroenterology Unit, Fondazione Policlinico Universitario A. Gemelli, IRCCS, Roma: Nicoletta de Matthaeis. Liver Injury and Transplant Unit, Polytechnic University of Marche, Ancona: Gloria Allegrini. Gastroenterology Unit, Belcolle hospital, Viterbo: Valentina Lauria, Giorgia Ghittoni, Giorgio Pelecca. Vascular and Interventional Radiology Unit, Belcolle Hospital, Viterbo: Fabrizio Chegai, Fabio Coratella, Mariano Ortenzi. Infectious Disease Unit, Belcolle hospital, Viterbo: Serena Dell'Isola. Department of Medicine and Surgery, Infectious Diseases and Hepatology Unit, University of Parma and Azienda Ospedaliero-Universitaria of Parma: Elisabetta Biasini, Andrea Olivani. Gastroenterology Unit, IRCCS Sacro Cuore Don Calabria hospital, Negrar: Alessandro Inno, Fabiana Marchetti. Department of Health Promotion, Mother & Child Care, Internal Medicine & Medical Specialties, PROMISE, Gastroenterology & Hepatology Unit, University of Palermo: Ciro Celsa, Mauro Grova, Caterina Stornello, Anita Busacca, Calogero Cammà, Giacomo Emanuele Maria Rizzo. Department of Clinical and Experimental Medicine, Clinical and Molecular Hepatology Unit, University of Messina: Maria Stella Franzè, Carlo Saitta. Department of Medical, Surgical and Experimental Sciences, Azienda Ospedaliero-Universitaria of Sassari: Assunta Sauchella. Department of Internal Medicine, Ospedale per gli Infermi di Faenza: Lucia Napoli, Vittoria Bevilacqua, Dante Berardinelli, Alberto Borghi, Andrea Casadei Gardini, Fabio Conti, Alessandro Cucchetti, Anna Chiara Dall'Aglio, Giorgio Ercolani. Department of Experimental and Clinical Medicine, Internal Medicine and Hepatology Unit, University of Firenze: Fabio Marra, Chiara Di Bonaventura, Stefano Gitto, Valentina Adotti. Department of Clinical Medicine and Surgery, Hepato-Gastroenterology Unit, University of Napoli "Federico II": Pietro Coccoli, Antonio Malerba. Department of Clinical Medicine and Surgery, Gastroenterology Unit, University of Napoli "Federico II": Mario Capasso, Filomena Morisco. Department of Clinical and Experimental Medicine, Hepatology and Liver Physiopathology Laboratory, University Hospital of Pisa, Pisa, Italy: Filippo Oliveri, Veronica Romagnoli.

CONFLICT OF INTEREST

None to declare.

AUTHOR CONTRIBUTIONS

EGG, MCPT, GP, FF, GLR, MDM, EC, RS, GC, CC, AM, MG, AG, GSB, FGF, GM, AM, GN, GR, GV, MRB, VS, MZ, FA and FT performed the research. AC, EGG, FT designed the research study. AC analysed the data. AC, EGG, and FT wrote the paper. All authors have read and approved the final manuscript.

ORCID

Edoardo G. Giannini https://orcid.org/0000-0001-8526-837X
Giuseppe Cabibbo https://orcid.org/0000-0002-0946-3859

REFERENCES

1. Sieghart W, Hucke F, Pinter M, et al. The ART of decision making: retreatment with transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology*. 2013;57:2261-2273.
2. Kadalayil L, Benini R, Pallan L, et al. A simple prognostic scoring system for patients receiving transarterial embolisation for hepatocellular cancer. *Ann Oncol*. 2013;24:2565-2570.
3. Hucke F, Pinter M, Graziadei I, et al. How to STATE suitability and START transarterial chemoembolization in patients with intermediate stage hepatocellular carcinoma. *J Hepatol*. 2014;61:1287-1296.

4. 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.
5. 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;36:100-107.
6. Cappelli A, Cucchetti A, Cabibbo G, et al. Refining prognosis after trans-arterial chemo-embolization for hepatocellular carcinoma. *Liver Int*. 2016;36:729-736.
7. 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.
8. 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;70:893-903.
9. Han G, Berhane S, Toyoda H, et al. Prediction of survival among patients receiving transarterial chemoembolization for hepatocellular carcinoma: a response-based approach. *Hepatology*. 2020;72:198-212.
10. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009;4(338):b606.
11. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol*. 2013;13:33.
12. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19:453-473.
13. Vitale A, Farinati F, Pawlik TM, et al. The concept of therapeutic hierarchy for patients with hepatocellular carcinoma: a multicenter cohort study. *Liver Int*. 2019;39:1478-1489.
14. Vitale A, Trevisani F, Farinati F, Cillo U. Treatment of hepatocellular carcinoma in the precision medicine era: from treatment stage migration to therapeutic hierarchy. *Hepatology*. 2020;72:2206-2218.
15. Moons KG, Harrell FE. Sensitivity and specificity should be de-emphasized in diagnostic accuracy studies. *Acad Radiol*. 2003;10:670-672.
16. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem*. 2008;54:17-23.
17. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA*. 2017;318:1377-1384.
18. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. *Stat Med*. 2015;34:1659-1680.
19. Han X, Zhang Y, Shao Y. On comparing 2 correlated C indices with censored survival data. *Stat Med*. 2017;36:4041-4049.
20. Gönen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika*. 2005;92:965-970.
21. Galle PR, Tovoli F, Foerster F, Wörns MA, Cucchetti A, Bolondi L. The treatment of intermediate stage tumours beyond TACE: From surgery to systemic therapy. *J Hepatol*. 2017;67:173-183.
22. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17:230.
23. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128-138.
24. Trevisani F, Golfieri R. Lipiodol transarterial chemoembolization for hepatocellular carcinoma: Where are we now? *Hepatology*. 2016;64:23-25.
25. Giannini EG, Bodini G, Corbo M, et al. Impact of evidence-based medicine on the treatment of patients with unresectable hepatocellular carcinoma. *Aliment Pharmacol Ther*. 2010;31:493-501.
26. Kudo M, Imanaka K, Chida N, et al. Phase III study of sorafenib after transarterial chemoembolisation in Japanese and Korean patients with unresectable hepatocellular carcinoma. *Eur J Cancer*. 2011;47:2117-2127.
27. Lencioni R, Llovet JM, Han G, et al. Sorafenib or placebo plus TACE with doxorubicin-eluting beads for intermediate stage HCC: The SPACE trial. *J Hepatol*. 2016;64:1090-1098.
28. Meyer T, Fox R, Ma YT, et al. Sorafenib in combination with transarterial chemoembolisation in patients with unresectable hepatocellular carcinoma (TACE 2): a randomised placebo-controlled, double-blind, phase 3 trial. *Lancet Gastroenterol Hepatol*. 2017;2:565-575.
29. Kudo M, Ueshima K, Ikeda M, et al. Randomised, multicentre prospective trial of transarterial chemoembolisation (TACE) plus sorafenib as compared with TACE alone in patients with hepatocellular carcinoma: TACTICS trial. *Gut*. 2020;69:1492-1501.
30. Bucci L, Garuti F, Lenzi B, et al. The evolutionary scenario of hepatocellular carcinoma in Italy: an update. *Liver Int*. 2017;37:259-270.
31. Cabibbo G, Petta S, Barbara M, et al. Hepatic decompensation is the major driver of death in HCV-infected cirrhotic patients with successfully treated early hepatocellular carcinoma. *J Hepatol*. 2017;67:65-71.
32. Giannini EG, Aglitti A, Borzio M Overview of immune checkpoint inhibitors therapy for hepatocellular carcinoma, and The ITA.LI.CA cohort derived estimate of amenability rate to immune checkpoint inhibitors in clinical practice. *Cancers*. 2019;11(11):1689.
33. Guarino M, Cucchetti A, Pontillo G, et al. Pattern of macrovascular invasion in hepatocellular carcinoma. *Eur J Clin Invest*. 2021;51:e13542. https://doi.org/10.1111/eci.13542

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Cucchetti A, Giannini EG, Mosconi C, et al; the ITA.LI.CA Study Group. Recalibrating survival prediction among patients receiving trans-arterial chemoembolization for hepatocellular carcinoma. *Liver Cancer Int*. 2021;2:45–53. https://doi.org/10.1002/lci2.33Graphical Abstract