Proceedings of ICBioMed 2026 Symposium: Biomedical Engineering and Public Health: Real-Time Wearables for Disease Detection
DOI: 10.54254/2753-8818/2026.ES38691

Nested Elastic-Net Cox Modeling for Post-Recurrence Survival Risk Stratification in Hepatocellular Carcinoma: Development and Internal Validation

Chengyi Wang

Department of Medical Laboratory Technology, College of Laboratory Medicine, Hebei North University, Zhangjiakou, China
13877863307@163.com

Abstract. Post-recurrence survival (PRS) after hepatocellular carcinoma (HCC) recurrence is heterogeneous and may reflect the combined effects of tumor burden, liver reserve, treatment characteristics, and routine laboratory measures. This secondary analysis aimed to develop and internally validate a penalized Cox model for PRS risk stratification. A total of 431 patients with recurrent HCC from a previously reported prospective cohort were analyzed using 30 baseline candidate predictors. A nested five-fold elastic-net Cox framework was applied, with inner cross-validation for lambda selection and outer cross-validation for strictly out-of-fold (OOF) prediction. The mixing parameter was fixed at alpha = 0.5; missing values were imputed using training-fold medians or modes, and tied event times were handled with the Breslow approximation. Among 431 patients, 183 experienced the predefined death event and 248 were censored. Eight predictors were retained in at least four of five outer folds: ablative modality, tumor number, gamma-glutamyl transferase, albumin, Child-Pugh class, drinking, smoking, and tumor size. The OOF C-index was 0.5576, with 1-, 3-, and 5-year AUCs of 0.5560, 0.5745, and 0.5934. Median OOF risk score stratification separated 216 low-risk and 215 high-risk patients, with median PRS of 44.07 versus 24.03 months (HR = 1.500, 95% CI 1.120-2.009; log-rank P = 0.00622). The framework provided significant risk stratification but limited discrimination and calibration, supporting its interpretation as an exploratory internally validated model rather than a stand-alone clinical prediction tool.

Keywords: hepatocellular carcinoma, post-recurrence survival, elastic-net Cox regression, nested cross-validation, risk stratification

1. Introduction

Hepatocellular carcinoma (HCC) frequently recurs after potentially curative or locoregional treatment, and outcomes after recurrence vary substantially between patients. Tumor burden, underlying liver reserve, treatment characteristics, and circulating biomarkers can all contribute to prognosis. In the source prospective cohort used for the present analysis, 431 of 583 patients experienced recurrence, providing a clinically relevant population in which to investigate post-

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.ES38691

recurrence survival [1]. Contemporary reviews and treatment frameworks likewise emphasize the importance of tumor burden, liver function, and treatment context in HCC prognosis [2, 3].

Conventional Cox proportional hazards regression is widely used for time-to-event analysis, but models that include many correlated clinical and laboratory variables may be unstable or overfit when variable selection and coefficient estimation are performed in the same dataset. Elastic-net regularization combines L1 and L2 penalties and can retain groups of correlated predictors while shrinking less informative coefficients toward zero [4]. Penalized Cox models extend this regularization framework to survival outcomes [5].

A second methodological challenge is optimistic performance estimation. If predictor selection or tuning is informed by observations that are later used for model evaluation, estimated discrimination can be biased upward. Nested cross-validation addresses this problem by separating model development and hyperparameter tuning in an inner loop from unbiased performance estimation in an outer loop. This separation is particularly important when the same dataset is used for both model development and internal validation. More broadly, resampling-based internal validation is recommended to reduce optimism and to separate model development from performance assessment [6, 7].

The present study therefore aimed to develop and internally validate a nested five-fold elastic-net Cox model for PRS among patients with recurrent HCC. Baseline variables available before locoregional treatment were used to generate strictly OOF predictions for all patients, evaluate discrimination and calibration, and examine whether the OOF risk score provided clinically interpretable survival stratification. The overall study design and nested modeling workflow are summarized in Figure 1.

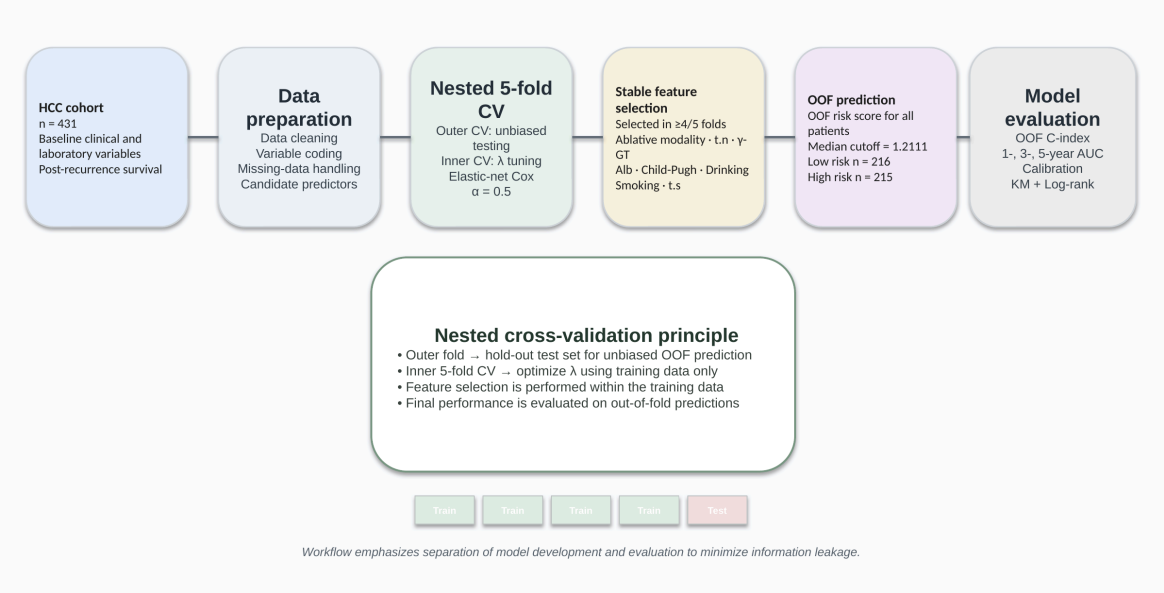


Figure 1. Study design and nested five-fold elastic-net Cox modeling workflow for post-recurrence survival in hepatocellular carcinoma. The analytical cohort consisted of 431 recurrent HCC patients. Baseline variables entered the nested cross-validation pipeline. Inner five-fold cross-validation selected λ using training data only, whereas the held-out outer fold generated unbiased OOF predictions. Predictors retained in at least four of five outer folds were considered stable. The median OOF risk score was used for low- versus high-risk stratification

DOI: 10.54254/2753-8818/2026.ES38691

2. Materials and methods

2.1. Data source, study design, and participants

This study was a secondary prognostic-modeling analysis of the recurrent-patient subset from a previously reported prospective cohort of virus-related HCC treated at Beijing You'an Hospital, Capital Medical University [1]. The source cohort prospectively enrolled 583 adults from January 1, 2012 to December 31, 2016. Patients were 18-75 years old, received locoregional treatment for curative purposes, and were required to achieve complete ablation. The source study excluded patients with Child-Pugh class C, severe coexisting disease, secondary liver cancer, or advanced HCC.

A total of 431 patients in the source cohort developed recurrence and constituted the analytical cohort for the present study. All 431 had valid PRS time and event information; 183 experienced the predefined death event and 248 were censored. The source study was approved by the Ethics Committee of Capital Medical University affiliated Beijing You'an Hospital (No. LL-2019-004-K). Written informed consent for participation was not required according to the original report [1].

2.2. Candidate predictors

Candidate predictors were restricted to baseline demographic, clinical, tumor-related, treatment-related, and routinely available laboratory variables. Thirty variables entered the modeling pipeline: age, sex, hypertension, diabetes, antiviral treatment, smoking, drinking, family history, etiology, ALD (dataset-coded variable), cirrhosis, Child-Pugh class, tumor number (t.n), tumor size (t.s), fractional ablation, ablative modality, platelet count, lymphocyte count, monocyte count, neutrophil count, ALT, AST, total bilirubin, albumin, γGT, AFP level, NLR, PLR, MLR, and INR. Variables measured at recurrence and outcome-derived variables were excluded from model development.

2.3. Outcome definition

The primary outcome was PRS, defined as the interval from first documented recurrence to HCC-related death or last follow-up, consistent with the recurrence-death survival definition in the source cohort [1]. The analytical survival-time variable was recurrencedeathmonth and the event indicator was Survival.status.analysis. Patients without the predefined death event during follow-up were censored.

2.4. Data preprocessing

Continuous predictors were entered as numerical variables and categorical predictors were factor encoded. Within every outer cross-validation iteration, missing continuous values were imputed using medians calculated exclusively from the outer training data, while missing categorical values were imputed using the corresponding training-set mode. The training-derived imputation values were then applied unchanged to the held-out outer fold. Predictor standardization was performed internally during glmnet fitting (standardize = TRUE). This preprocessing strategy avoided information leakage from the outer validation fold.

DOI: 10.54254/2753-8818/2026.ES38691

2.5. Nested five-fold elastic-net Cox modeling

Nested five-fold cross-validation was used for model development and internal validation. The cohort was split into five stratified outer folds according to event status. In each outer iteration, four folds were used for model development and the remaining fold was held out for evaluation. Within the outer training data, a stratified inner five-fold cross-validation procedure selected λ by minimum cross-validated partial-likelihood deviance (lambda.min). The elastic-net mixing parameter was fixed at $\alpha = 0.5$. Across the five outer folds, selected λ values ranged from 0.0573 to 0.1552, with a mean of 0.0872.

After λ selection, the elastic-net Cox model was fitted to the complete outer training set and applied only to the corresponding held-out outer fold. Repeating this process across all outer folds produced one OOF linear predictor for every patient. Tied event times in the penalized Cox models were handled using the Breslow approximation, explicitly fixed to ensure reproducibility across glmnet versions [5].

2.6. Predictor-selection stability

Predictor-selection stability was summarized across the five outer training models at the parent-variable level. Variables selected in at least four of five outer folds were classified as the final stable predictor set. Selection in at least three folds was recorded descriptively but was not used to define the final stable set.

2.7. OOF performance and calibration

Primary performance measures were calculated from true OOF predictions. Overall discrimination was quantified using Harrell's concordance index (C-index). Time-dependent AUCs at 1, 3, and 5 years were estimated using the timeROC approach [8]. For calibration, each outer training fold was used to estimate a fold-specific baseline cumulative hazard under the Breslow approximation. The fold-specific baseline hazard and held-out OOF linear predictor were combined to obtain absolute event probabilities at 1, 3, and 5 years. Calibration was summarized by comparing mean predicted event probability with Kaplan-Meier observed event probability across quintiles of OOF predicted risk.

2.8. OOF risk stratification and Kaplan-Meier analysis

The OOF risk score was calculated as exp(OOF linear predictor). Patients were divided at the cohort median OOF risk score (1.2111), yielding 216 low-risk and 215 high-risk patients. Kaplan-Meier curves were generated on the PRS time scale and compared using the log-rank test. A Cox model containing the OOF-derived risk-group indicator was used to estimate the HR and 95% CI for high versus low risk. The Breslow approximation was used for tied event times.

2.9. Conventional multivariable Cox model

For clinical interpretation, a conventional multivariable Cox proportional hazards model was fitted using only variables that met the ≥4/5-fold stability criterion. This model was not used to generate OOF performance estimates and therefore served as an association-oriented interpretation step rather than a second predictive validation analysis. Tied event times were handled using the Breslow approximation.

DOI: 10.54254/2753-8818/2026.ES38691

2.10. Statistical software and reporting

Two-sided P values <0.05 were considered statistically significant for conventional association analyses. Analyses were performed in R version 4.5.1 using readxl 1.5.0, survival 3.8-3, glmnet 5.0, and timeROC 0.4.1. The reporting strategy distinguished model development, predictor selection, and internal validation in accordance with general TRIPOD principles [9].

3. Results

3.1. Study population

The analytical cohort comprised 431 patients with recurrent HCC. During PRS follow-up, 183 patients experienced the predefined death event and 248 were censored. Baseline characteristics are summarized in Table 1.

Table 1. Baseline characteristics of the analytical cohort (n = 431)

Variable	Overall cohort (n = 431)
Age, mean (SD), years	56.93 (8.26)
Female, n (%)	69 (16.0)
Etiology, n (%)	
HBV	368 (85.4)
HCV	43 (10.0)
co-infection	20 (4.6)
ALD variable = 1, n (%)	71 (16.5)
Cirrhosis, n (%)	369 (85.6)
Child-Pugh class B, n (%)	117 (27.1)
Hypertension, n (%)	99 (23.0)
Diabetes, n (%)	74 (17.2)
Antiviral therapy, n (%)	217 (50.6)
Smoking, n (%)	197 (45.7)
Drinking, n (%)	147 (34.1)
Family history, n (%)	181 (42.1)
Multiple tumors (t.n = 1), n (%)	157 (36.4)
Tumor size >30 mm (t.s = 1), n (%)	148 (34.3)
Fractional ablation, n (%)	66 (15.3)
Ablative modality, n (%)	
RFA (category 1)	229 (53.1)
MWA (category 2)	111 (25.8)
AHC (category 3)	91 (21.1)
PLT, mean (SD)	119.16 (57.13)
Lymphocytes, mean (SD)	1.25 (0.62)
Monocytes, mean (SD)	0.44 (0.27)
Neutrophils, mean (SD)	3.33 (1.84)

DOI: 10.54254/2753-8818/2026.ES38691

Table 1. (continued)

ALT, mean (SD)	40.24 (25.94)
AST, mean (SD)	37.70 (20.56)
TBIL, mean (SD)	19.49 (10.23)
Albumin, mean (SD)	36.71 (4.37)
γGT, mean (SD)	76.18 (61.47)
AFP level = 1, n (%)	259 (60.5)
NLR, mean (SD)	3.31 (2.74)
PLR, mean (SD)	108.62 (55.96)
MLR, mean (SD)	0.40 (0.23)

Values are presented as mean (SD) or n (%), as appropriate. Percentages for variables with missing observations are based on available data.

3.2. Stable predictors selected by nested elastic-net regression

Eight predictors met the stricter stability criterion of selection in at least four of five outer folds: ablative modality, tumor number, γGT, albumin, Child-Pugh class, drinking status, smoking status, and tumor size. Ablative modality, tumor number, and γGT were selected in all five outer folds; albumin, Child-Pugh class, drinking, smoking, and tumor size were selected in four of five folds. AFP level, ALT, antiviral treatment, AST, and hypertension were selected in three of five folds but did not meet the final ≥4/5 stability threshold.

Across folds in which the eight stable predictors had non-zero coefficients, coefficient direction was consistent: tumor number, γGT, Child-Pugh category, drinking, smoking, tumor size, and the selected ablative-modality contrast had positive coefficients, whereas albumin had negative coefficients. These patterns support relative stability of both selection and direction across the nested resampling procedure.

3.3. OOF discrimination and calibration

The nested elastic-net Cox model achieved an OOF C-index of 0.5576. Time-dependent AUCs were 0.5560 at 1 year, 0.5745 at 3 years, and 0.5934 at 5 years. Discrimination therefore remained limited across all assessed horizons, although the AUC increased modestly over time.

OOF calibration showed deviation from the ideal 45° line, especially at higher predicted event probabilities. The higher-risk strata generally showed overestimation of absolute event probability. These findings indicate imperfect absolute-risk calibration and argue against interpreting the current model as a calibrated individual-level survival prediction tool. The principal OOF performance and risk-stratification metrics are summarized in Table 2, while the corresponding discrimination, calibration, risk-score distribution, and Kaplan-Meier results are shown in Figure 2.

Table 2. Out-of-fold performance and risk-stratification results

Metric	Result
OOF C-index	0.5576
1-year time-dependent AUC	0.5560
3-year time-dependent AUC	0.5745

DOI: 10.54254/2753-8818/2026.ES38691

Table 2. (continued)

5-year time-dependent AUC	0.5934
Median OOF risk-score cutoff	1.2111
Low-risk group, n	216
High-risk group, n	215
Median PRS, low-risk group (months)	44.07
Median PRS, high-risk group (months)	24.03
High vs Low risk HR (95% CI)	1.500 (1.120-2.009)
Log-rank χ^2	7.486
Log-rank P	0.00622

3.4. OOF risk stratification and Kaplan-Meier analysis

The median OOF risk score of 1.2111 divided the cohort into 216 low-risk and 215 high-risk patients. Median PRS was 44.07 months in the low-risk group versus 24.03 months in the high-risk group, a difference of approximately 20.03 months.

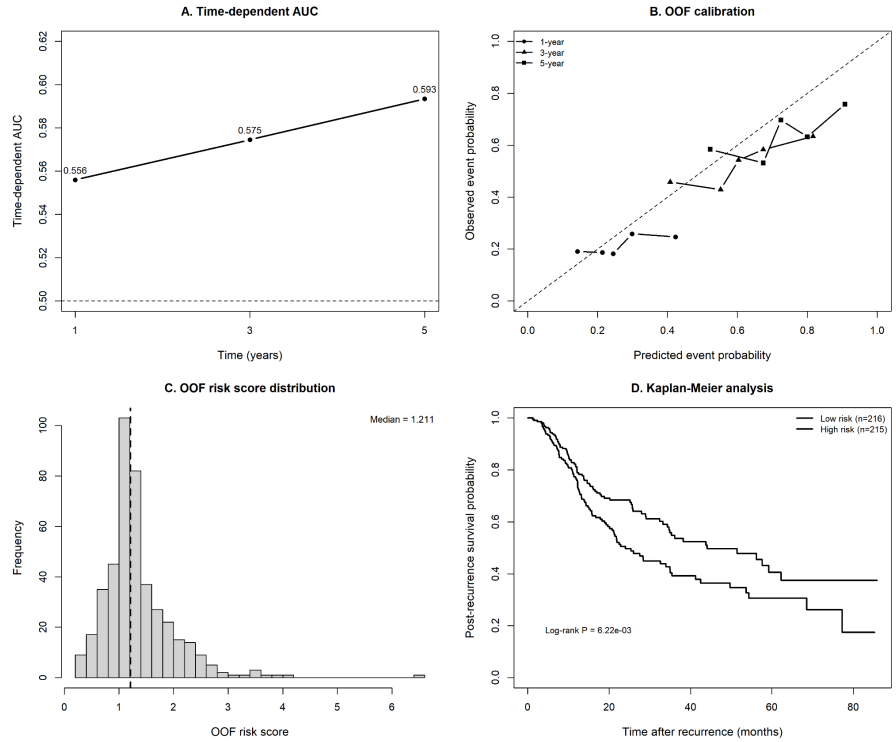


Figure 2. Internal validation and OOF risk stratification of the nested elastic-net Cox model for post-recurrence survival in HCC. (A) Time-dependent AUCs at 1, 3, and 5 years (0.556, 0.575, and 0.593). (B) OOF calibration comparing predicted and observed event probabilities; the dashed diagonal line represents ideal calibration. (C) Distribution of OOF risk scores; the median score (1.211) defined the risk-group cutoff. (D) Kaplan-Meier curves for low-risk (n = 216) and high-risk (n = 215) patients. Median PRS was 44.07 versus 24.03 months; HR = 1.500 (95% CI 1.120-2.009); log-rank P = 0.00622

DOI: 10.54254/2753-8818/2026.ES38691

High-risk patients had a higher mortality hazard than low-risk patients (HR = 1.500, 95% CI 1.120-2.009). Kaplan-Meier survival distributions differed significantly between groups (log-rank χ^2 = 7.486, P = 0.00622). Thus, despite limited individual-level discrimination, the OOF score provided statistically significant broad risk stratification.

3.5. Conventional multivariable Cox interpretation

In the conventional multivariable Cox model including the eight stable predictors, tumor number and γGT were statistically significant. Higher tumor number was associated with increased mortality hazard (HR = 1.461, 95% CI 1.074-1.987; P = 0.0157), and γGT was also associated with higher mortality risk (HR = 1.0039 per unit, 95% CI 1.0016-1.0061; P < 0.001). Tumor size showed a borderline association (HR = 1.375, 95% CI 0.990-1.911; P = 0.0576). The remaining stable predictors did not reach conventional statistical significance. The complete conventional Cox regression results are presented in Table 3.

The difference between stable selection in penalized regression and conventional P-value significance is expected because the two procedures answer different questions: elastic-net selection assesses joint predictive contribution under coefficient shrinkage, whereas conventional Cox regression estimates unpenalized associations and their hypothesis tests.

Table 3. Conventional multivariable Cox regression of predictors retained in at least four of five outer folds

Variable	HR (95% CI)	P value
MWA vs RFA	1.381 (0.970-1.966)	0.073
AHC vs RFA	0.979 (0.658-1.455)	0.915
Multiple vs single tumors	1.461 (1.074-1.987)	0.016
γGT (per unit)	1.004 (1.002-1.006)	<0.001
Albumin (per unit)	0.977 (0.942-1.014)	0.225
Child-Pugh class B vs A	1.308 (0.929-1.841)	0.123
Drinking (1 vs 0)	1.222 (0.881-1.695)	0.229
Smoking (1 vs 0)	1.308 (0.960-1.782)	0.089
Tumor size >30 vs ≤30 mm	1.375 (0.990-1.911)	0.058

Categorical contrasts are shown relative to their model-matrix reference levels. The conventional Cox model was used for interpretation and not for OOF performance estimation.

4. Discussion

4.1. Principal findings

This secondary analysis developed and internally validated a nested five-fold elastic-net Cox framework for PRS risk stratification in recurrent HCC. Eight baseline variables were relatively stable across outer training folds: ablative modality, tumor number, γGT, albumin, Child-Pugh class, drinking, smoking, and tumor size. The strictly OOF C-index was 0.5576, and time-dependent AUCs ranged from 0.5560 to 0.5934, demonstrating limited discrimination. Nevertheless, the

DOI: 10.54254/2753-8818/2026.ES38691

median OOF risk score separated patients into groups with materially different median PRS (44.07 vs 24.03 months) and a significant difference in mortality hazard.

The central interpretation is therefore not that the model achieves high individual-level predictive accuracy. Rather, the results support an exploratory risk-stratification framework that captures a reproducible multivariable signal while providing a conservative estimate of internal performance through nested OOF evaluation.

4.2. Interpretation of the stable predictors

Tumor burden was prominent in the stable set. Tumor number was selected in all five outer folds and remained statistically significant in the conventional multivariable Cox model, while tumor size was selected in four folds and showed a borderline conventional association. These findings are clinically plausible because greater tumor burden may reflect more aggressive disease biology and a higher likelihood of subsequent progression after recurrence.

Liver-related variables also contributed. γGT was selected in all five outer folds and remained significantly associated with mortality in conventional Cox regression. Albumin and Child-Pugh class were selected in four folds, consistent with the importance of hepatic functional reserve for outcomes after HCC recurrence. However, these variables should be interpreted as components of a multivariable prognostic pattern rather than isolated causal effects.

Ablative modality, smoking, and drinking were also stably selected. Their coefficients may reflect treatment selection, baseline risk profiles, or residual clinical differences that are not fully captured by the available covariates. The analysis does not establish causal effects of these exposures or treatment categories. Notably, AFP level was selected in three of five folds but did not meet the stricter ≥4/5 stability threshold. This does not imply that AFP lacks prognostic value; rather, its incremental predictive contribution was less stable than that of the final eight variables under the present resampling scheme.

4.3. Methodological value of nested cross-validation

The use of nested cross-validation was a major methodological feature. Predictor selection and λ tuning were confined to the training data within each outer iteration, and performance was evaluated only in the corresponding held-out fold. OOF predictions were then pooled across outer folds. This design reduces the optimism that can occur when feature selection and evaluation are performed on the same observations.

The training-fold-only imputation strategy served the same purpose. Missing-data summaries derived from validation patients were not used during model development. In addition, the tie-handling method was explicitly locked to Breslow to improve reproducibility across software versions. These implementation details are important because seemingly minor preprocessing or software defaults can change apparent model performance.

4.4. Discrimination, calibration, and risk stratification

The model's discrimination was limited. An OOF C-index of 0.5576 and AUCs below 0.60 at all three horizons indicate that the model should not be presented as a high-accuracy clinical prediction tool. The modest increase in AUC from 1 to 5 years does not change this interpretation.

Calibration also deviated from ideal, particularly in higher predicted-risk strata, where event probabilities tended to be overestimated. Because the analysis used only baseline variables, the

DOI: 10.54254/2753-8818/2026.ES38691

model may not capture disease status, treatment response, or laboratory changes occurring at the time of recurrence. This likely limits accurate estimation of absolute post-recurrence risk.

Despite these limitations, risk-group separation was statistically significant. Median PRS differed by approximately 20.03 months, and the high-risk group had a 1.50-fold higher mortality hazard. This distinction is important: statistically significant separation of broad risk groups can coexist with limited individual-level discrimination. Accordingly, the model is better characterized as a preliminary stratification tool than as a personalized survival calculator.

4.5. Strengths and limitations

The strengths of this analysis include the use of a clinically defined recurrent-HCC cohort, penalized survival modeling for correlated predictors, training-fold-only preprocessing, nested cross-validation, true OOF performance estimation, time-dependent discrimination assessment, and OOF calibration based on held-out absolute risk predictions.

Several limitations should also be acknowledged. First, this was a secondary analysis of a single center-derived cohort and therefore requires external validation. Second, the analytical cohort included only patients who experienced recurrence; results are conditional on recurrence and should not be generalized to all treated HCC patients. Third, only baseline variables were used, whereas recurrence-time tumor characteristics, treatment after recurrence, and dynamic biomarkers may contain important additional prognostic information. Fourth, the sample size and 183 observed events constrain the complexity and precision of multivariable prediction. Fifth, the median risk-score cutoff was data-derived and should not be considered a clinically established threshold. Finally, no external recalibration was performed, and the observed calibration deviations limit immediate clinical use.

5. Conclusion

A nested five-fold elastic-net Cox framework identified eight relatively stable baseline predictors of post-recurrence survival in patients with recurrent HCC. Strict OOF evaluation showed limited discrimination and imperfect calibration, but the OOF risk score significantly stratified patients into groups with different post-recurrence survival. The model should therefore be regarded as an exploratory internally validated risk-stratification framework. External validation, incorporation of recurrence-time information, and recalibration are required before clinical application.

Ethics and data source statement

This study used the recurrent-patient subset of a previously published prospective cohort [1]. The source study was approved by the Ethics Committee of Capital Medical University affiliated Beijing You'an Hospital (No. LL-2019-004-K), and written informed consent for participation was not required according to the original report. The present work represents a secondary analysis of the supplied cohort data.

Data availability

The present analysis was performed on the recurrent-patient dataset derived from the previously published cohort. The source publication states that original contributions are included in the article and supplementary material and that further inquiries may be directed to the corresponding authors [1].

DOI: 10.54254/2753-8818/2026.ES38691

References

[1] Wang, Q., Liu, B., Qiao, W., Li, J., Yuan, C., Long, J., Hu, C., Zang, C., Zheng, J., & Zhang, Y. (2021). The dynamic changes of AFP from baseline to recurrence as an excellent prognostic factor of hepatocellular carcinoma after locoregional therapy: A 5-year prospective cohort study. *Frontiers in Oncology, 11*, 756363. https://doi.org/10.3389/fonc.2021.756363

[2] Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., Lencioni, R., Koike, K., Zucman-Rossi, J., & Finn, R. S. (2021). Hepatocellular carcinoma. *Nature Reviews Disease Primers, 7*(1), 6. https://doi.org/10.1038/s41572-020-00240-3

[3] Reig, M., Forner, A., Rimola, J., Ferrer-Fàbrega, J., Burrel, M., Garcia-Criado, Á., Kelley, R. K., Galle, P. R., Mazzaferro, V., Salem, R., Sangro, B., Singal, A. G., Vogel, A., Fuster, J., Ayuso, C., & Bruix, J. (2022). BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *Journal of Hepatology, 76*(3), 681–693. https://doi.org/10.1016/j.jhep.2021.11.018

[4] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

[5] Simon, N., Friedman, J., Hastie, T., & Tibshirani, R. (2011). Regularization paths for Cox's proportional hazards model via coordinate descent. *Journal of Statistical Software, 39*(5), 1–13. https://doi.org/10.18637/jss.v039.i05

[6] Harrell, F. E., Jr. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis* (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-19425-7

[7] Steyerberg, E. W. (2019). *Clinical prediction models: A practical approach to development, validation, and updating* (2nd ed.). Springer. https://doi.org/10.1007/978-3-030-16399-0

[8] Blanche, P., Dartigues, J. F., & Jacqmin-Gadda, H. (2013). Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in Medicine, 32*(30), 5381–5397. https://doi.org/10.1002/sim.5958

[9] Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine, 162*(1), 55–63. https://doi.org/10.7326/M14-0697