

Development and External Validation of a TNM plus CEA Nomogram for Overall Survival after Surgery in Early-Onset Colorectal Cancer

Junting Guan

The Second Affiliated Hospital of Harbin Medical University

Huan Xiong

The Second Affiliated Hospital of Nanchang University

Zewen Chang

The Fourth Affiliated Hospital of Harbin Medical University

Jiaqi Wang

The Second Affiliated Hospital of Harbin Medical University

Xingyuan Li

The Second Affiliated Hospital of Harbin Medical University

Shengqi Pan

The Second Affiliated Hospital of Harbin Medical University

Qi Li

The Second Affiliated Hospital of Harbin Medical University

Qingchao Tang

tangqingchao@hrbmu.edu.cn

The Second Affiliated Hospital of Harbin Medical University

Research Article

Keywords: early-onset colorectal cancer, nomogram, carcinoembryonic antigen, overall survival, Cox regression, external validation

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10625905/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background: Early-onset colorectal cancer (EOCRC) requires postoperative prognostic tools that add clinically useful information beyond anatomic stage while remaining transparent enough for external use. We developed and externally validated a parsimonious overall-survival (OS) nomogram for surgically treated EOCRC and assessed the incremental value of pretreatment carcinoembryonic antigen (CEA).

Methods: SEER patients diagnosed from 2011 to 2015 were assigned to development and internal validation cohorts, and patients treated at the Second Affiliated Hospital of Harbin Medical University from 2014 to 2017 formed an external validation cohort. Candidate predictors were screened with LASSO Cox regression. LASSO-selected models were evaluated using nested model comparison, proportional-hazards diagnostics, a grade-stratified sensitivity model, Harrell's C-index, time-dependent AUC, calibration, decision curve analysis, and Kaplan-Meier risk stratification.

Results: The analysis included 7164 patients, including 3447 in development, 3446 in internal validation, and 271 in external validation. LASSO retained tumor grade, AJCC T, AJCC N, AJCC M, and CEA. The grade-containing model fitted better but violated the global proportional-hazards assumption; therefore, the primary model retained AJCC T/N/M and CEA, with grade handled in sensitivity analysis. C-indexes were 0.813, 0.805, and 0.700 in the three cohorts. External 1-, 3-, and 5-year AUCs were 0.604, 0.705, and 0.726.

Conclusions: This externally validated TNM plus CEA nomogram provides a simple risk-stratification tool for OS after surgery in EOCRC. Absolute long-term predictions should be interpreted cautiously in external settings.

Background

Colorectal cancer (CRC) remains a major contributor to cancer incidence and mortality worldwide, including in China [1–3]. Although screening and prevention have reduced CRC burden in some older populations, CRC diagnosed before 50 years of age has increased in several regions. This epidemiologic shift has drawn attention to early-onset colorectal cancer (EOCRC), which differs from later-onset CRC in stage distribution and tumor biology, and which carries a distinct survivorship burden [4–13].

Prognostic accuracy matters here. Most of these patients are diagnosed at a point in life when family and career commitments are heaviest. The AJCC TNM staging system remains the foundation of prognostic assessment and treatment planning [14], but outcomes can vary within the same stage. A prediction model based on routinely available clinicopathologic variables may therefore help refine individualized survival estimates and identify patients who require closer follow-up after surgery.

Several SEER-based EOCRC nomograms have been reported [15–17]. However, a prediction model that mainly extends TNM staging must show what is gained beyond TNM alone, not only that a multivariable model can be fitted. Prognostic Cox models also require transparent assessment of PH assumptions, because a predictor may be clinically informative and improve likelihood-based fit while still behaving as a time-dependent factor in a standard Cox nomogram. These issues are especially relevant when a model is intended for external use rather than for explaining one registry dataset.

We therefore developed and externally validated a Cox-based nomogram for OS in surgically treated EOCRC using SEER development and internal validation cohorts and an independent hospital cohort from China. The

study was reported in accordance with TRIPOD/TRIPOD + AI guidance [18, 29]. Rather than maximizing apparent fit, we audited each step of model building against a TNM-only benchmark, with particular attention to the PH assumption. The final aim was to produce a reproducible and clinically interpretable TNM plus CEA model while being explicit about transportability and calibration limitations.

Methods

Study design and data sources

This retrospective prognostic prediction study developed and externally validated a nomogram for OS in EOCRC. SEER records for patients diagnosed from January 2011 to December 2015 were used for model development and internal validation. Eligible SEER patients were divided at an approximately 1:1 ratio into development and internal validation cohorts during the R-based data-preparation workflow. The external validation cohort included patients treated at the Second Affiliated Hospital of Harbin Medical University between January 2014 and June 2017. The same eligibility criteria were applied to the external cohort.

Ethics

All methods were carried out in accordance with relevant guidelines and regulations. The SEER database contains de-identified publicly available data and does not require additional institutional review board approval for approved users. The study adhered to the World Medical Association's Declaration of Helsinki for ethical human research. The need for ethics approval and consent was waived by the Ethics Committee of the Second Affiliated Hospital of Harbin Medical University due to the retrospective nature of the study.

Participants and outcome

Patients were eligible if primary tumor resection was recorded as the first course of treatment, age at diagnosis was younger than 50 years, and colorectal cancer was pathologically confirmed. Patients were excluded if they received preoperative chemoradiotherapy, had incomplete data required for survival analysis or candidate predictors, or had multiple primary tumors. Patients with M1 disease were retained if they otherwise met the eligibility criteria because AJCC M category was a prespecified candidate predictor and final-model variable; clinically, these operations may include resection for symptom control in selected patients with obstruction, bleeding, or pain.

The endpoint was OS, defined as the time from diagnosis to death from any cause or last follow-up. For the external cohort, survival time from diagnosis and vital status were obtained from institutional follow-up records and were recorded in months. Death was coded as the event, and patients alive at last follow-up were censored. Records with zero-month survival time were replaced by 0.5 month to retain these patients while ensuring strictly positive survival times for Cox regression; the available preprocessing audit is provided in Supplementary Material S1.

Candidate predictors and variable processing

Candidate predictors were selected according to clinical relevance and availability in both SEER and external validation data. The candidate set included age, race, sex, marital status, primary tumor site, histological subtype, tumor grade, AJCC T stage, AJCC N stage, AJCC M stage, pretreatment CEA, radiotherapy status, chemotherapy status, lymph-node-yield category, and tumor size. Lymph-node-yield category was retained in the candidate set because examination of at least 12 regional lymph nodes is a recognized colorectal cancer quality indicator, although it was not selected into the final model.

Categorical variables were harmonized before analysis. Primary tumor site was grouped as right colon, left colon, or rectum. The right colon included the cecum, ascending colon, hepatic flexure, and transverse colon; the left colon included the splenic flexure, descending colon, and sigmoid colon. Histology was grouped as adenocarcinoma, mucinous/signet-ring cell carcinoma, or other histology. AJCC T and N were coded as T1-T4 and N0-N2, respectively; AJCC M was coded as M0, M1a, and M1b. Because 2011–2015 SEER cases do not provide AJCC 8th-edition M1c information, these variables were treated as SEER-derived AJCC 7th-edition-compatible T/N/M categories [14]. Pretreatment CEA was dichotomized as < 5 ng/mL or ≥ 5 ng/mL.

Model development and validation

Predictor screening was performed in the development cohort using LASSO Cox regression with ten-fold cross-validation. The λ_{1se} rule was selected to favor parsimony [21]. LASSO selection is shown in Fig. 1. The full LASSO-selected Cox model (Model A) included tumor grade, AJCC T, AJCC N, AJCC M, and CEA. The simplified candidate model (Model B) omitted tumor grade. Model A and Model B were compared using log-likelihood, AIC, BIC, apparent C-index, nested likelihood-ratio testing, and Schoenfeld residual PH diagnostics [22].

Model selection weighed likelihood-based fit and PH diagnostics against the practical requirements of a point-based nomogram. Because tumor grade was selected by LASSO but was the main source of PH violation in Model A, an additional grade-stratified sensitivity model (Model C) was fitted as $\text{Surv}(\text{months}, \text{status}) \sim \text{AJCC T} + \text{AJCC N} + \text{AJCC M} + \text{CEA} + \text{strata}(\text{Grade})$. In Model C, grade defined stratum-specific baseline hazards rather than a single grade coefficient; therefore, it did not receive a conventional hazard ratio or point assignment. Model C coefficients and grade-specific baseline survival values are provided in Supplementary Material S1.

The final model was also compared with a prespecified TNM-only benchmark to quantify the incremental value of pretreatment CEA. Discrimination was assessed using Harrell's C-index and time-dependent AUC at 12, 36, and 60 months [23, 25, 26]. Internal calibration was assessed by bootstrap calibration in the development cohort. Exploratory external calibration used prediction quintiles in the hospital validation cohort and compared mean predicted survival with Kaplan-Meier observed survival. Three-year DCA compared the final TNM plus CEA model with TNM-only, treat-all, and treat-none strategies [28]. Risk stratification used the development-cohort median linear predictor as the cutoff; Kaplan-Meier curves, log-rank tests, and Cox models compared high-risk and low-risk groups.

For reproducible absolute survival prediction, the linear predictor was calculated as the sum of the final-model coefficients multiplied by the corresponding indicator variables. Predicted OS at time t was calculated as $S(t|X) = S_0(t)^{\exp(LP)}$, where $S_0(t)$ is the development-cohort baseline survival. The final coefficients and baseline

survival values at 12, 36, and 60 months are provided in Supplementary Material S1; the nomogram is a graphical representation of the same fixed Cox model.

No formal sample-size calculation was performed because the study used existing registry and institutional cohorts. The development cohort included 787 deaths for eight final-model regression coefficients, exceeding common minimum-event considerations for Cox prediction modeling [24]. Analyses were performed using R version 4.5.1 with survival, glmnet, rms, timeROC, dcurves, tableone, ggplot2, and survminer packages.

Results

Study population

The final analysis included 7164 patients with EOCRC: 3,447 in the development cohort, 3,446 in the internal validation cohort, and 271 in the external validation cohort. Deaths occurred in 787 patients (22.8%) in the development cohort, 794 patients (23.0%) in the internal validation cohort, and 87 patients (32.1%) in the external validation cohort. Median follow-up by reverse Kaplan-Meier analysis was 43, 45, and 68 months, respectively. Baseline characteristics are summarized in Table 1.

Table 1
Baseline clinicopathologic characteristics of the development, internal validation, and external validation cohorts.

Characteristic	Level	Development cohort	Internal validation cohort	External validation cohort	P value
Number of patients		3447	3446	271	
Age, years	Median [IQR]	44 [39–47]	44 [39–47]	44 [40–47]	0.994
Sex	Male	1666 (48.3)	1678 (48.7)	160 (59.0)	0.003
	Female	1781 (51.7)	1768 (51.3)	111 (41.0)	
Race	White	2499 (72.5)	2532 (73.5)	0 (0.0)	< 0.001
	Black	519 (15.1)	508 (14.7)	0 (0.0)	
	Other	429 (12.4)	406 (11.8)	271 (100.0)	
Marital status	Married	2012 (58.4)	2066 (60.0)	241 (88.9)	< 0.001
	Unmarried	1435 (41.6)	1380 (40.0)	30 (11.1)	
Primary tumor site	Right colon	1223 (35.5)	1268 (36.8)	56 (20.7)	< 0.001
	Left colon	1484 (43.1)	1449 (42.0)	71 (26.2)	
	Rectum	740 (21.5)	729 (21.2)	144 (53.1)	
Histological subtype	Adenocarcinoma	3090 (89.6)	3058 (88.7)	208 (76.8)	< 0.001
	Other histological subtypes	75 (2.2)	75 (2.2)	12 (4.4)	
	Mucinous adenocarcinoma or signet-ring cell carcinoma	282 (8.2)	313 (9.1)	51 (18.8)	
Tumor grade	Well differentiated	198 (5.7)	187 (5.4)	43 (15.9)	< 0.001
	Moderately differentiated	2502 (72.6)	2517 (73.0)	191 (70.5)	
	Poorly differentiated	616 (17.9)	594 (17.2)	28 (10.3)	
	Undifferentiated	131 (3.8)	148 (4.3)	9 (3.3)	
AJCC T category	T1	231 (6.7)	234 (6.8)	25 (9.2)	< 0.001
	T2	372 (10.8)	402 (11.7)	26 (9.6)	
	T3	2056 (59.6)	2034 (59.0)	93 (34.3)	
	T4	788 (22.9)	776 (22.5)	127 (46.9)	

Characteristic	Level	Development cohort	Internal validation cohort	External validation cohort	P value
AJCC N category	N0	1420 (41.2)	1382 (40.1)	175 (64.6)	< 0.001
	N1	1086 (31.5)	1129 (32.8)	57 (21.0)	
	N2	941 (27.3)	935 (27.1)	39 (14.4)	
AJCC M category	M0	2659 (77.1)	2633 (76.4)	252 (93.0)	< 0.001
	M1a	476 (13.8)	478 (13.9)	12 (4.4)	
	M1b	312 (9.1)	335 (9.7)	7 (2.6)	
CEA level	< 5 ng/mL	2034 (59.0)	1989 (57.7)	181 (66.8)	0.012
	≥5 ng/mL	1413 (41.0)	1457 (42.3)	90 (33.2)	
Radiotherapy	No	3171 (92.0)	3203 (92.9)	244 (90.0)	0.109
	Yes	276 (8.0)	243 (7.1)	27 (10.0)	
Chemotherapy	No	1179 (34.2)	1156 (33.5)	155 (57.2)	< 0.001
	Yes	2268 (65.8)	2290 (66.5)	116 (42.8)	
Lymph-node-yield category	< 12 examined regional lymph nodes	33 (1.0)	40 (1.2)	49 (18.1)	< 0.001
	≥12 examined regional lymph nodes	3414 (99.0)	3406 (98.8)	222 (81.9)	
Tumor size	< 5 cm	1664 (48.3)	1660 (48.2)	150 (55.4)	0.070
	≥5 cm	1783 (51.7)	1786 (51.8)	121 (44.6)	

Predictor selection and model choice

LASSO retained five parent predictors at λ_{1se} : tumor grade, AJCC T, AJCC N, AJCC M, and pretreatment CEA (Fig. 1). Model A, which included all five predictors, fitted better than Model B by likelihood criteria (likelihood-ratio chi-square 41.093, $df = 3$, $P < 0.001$), but Model A violated the global PH assumption ($P < 0.001$), mainly because of tumor grade ($P < 0.001$). Model B satisfied the global PH test ($P = 0.095$), although AJCC N showed a mild predictor-specific signal ($P = 0.031$).

The grade-stratified sensitivity model (Model C) resolved the PH concern (global $P = 0.390$) and also removed the AJCC N signal ($P = 0.196$). Its discrimination was slightly higher than Model B in all three cohorts, but the absolute gains were small (C-index differences approximately 0.007–0.010), and a routine nomogram would require grade-specific baseline survival scales. Therefore, Model B was retained as the primary point-based model, with Model C presented as sensitivity analysis rather than as evidence that grade is unimportant.

Final model performance and incremental value beyond TNM alone

The final model included AJCC T, AJCC N, AJCC M, and pretreatment CEA. Advanced T stage, nodal involvement, metastatic disease, and elevated CEA were associated with poorer OS (Table 2). The nomogram is shown in Fig. 2. Compared with a TNM-only benchmark, adding CEA significantly improved likelihood-based fit in the development cohort (likelihood-ratio chi-square 57.429, df = 1, P < 0.001). The C-index increased from 0.806 to 0.813 in the development cohort, from 0.799 to 0.805 in the internal validation cohort, and from 0.684 to 0.700 in the external validation cohort.

Table 2
Final model specification and validation performance. A. Final multivariable Cox model coefficients.

Predictor	Coefficient	HR (95% CI)	P value
AJCC T2 vs T1	-0.043	0.958 (0.514–1.787)	0.893
AJCC T3 vs T1	0.440	1.552 (0.936–2.575)	0.089
AJCC T4 vs T1	0.853	2.347 (1.402–3.927)	0.001
AJCC N1 vs N0	0.522	1.685 (1.358–2.091)	< 0.001
AJCC N2 vs N0	0.949	2.582 (2.086–3.196)	< 0.001
AJCC M1a vs M0	1.337	3.808 (3.184–4.554)	< 0.001
AJCC M1b vs M0	1.703	5.488 (4.514–6.672)	< 0.001
CEA > = 5 ng/mL vs < 5 ng/mL	0.610	1.840 (1.565–2.163)	< 0.001

B. Validation performance summary.

Cohort	N/events	Median follow-up, months	C-index (95% CI)	AUC 1/3/5 year	Calibration summary	High-vs low-risk HR (95% CI)
Development cohort	3447/787	43	0.813 (0.798–0.828)	0.841/0.857/0.844	Internal MAE 0.0134/0.0242/0.0265	7.454 (6.114–9.088), < 0.001
Internal validation cohort	3446/794	45	0.805 (0.790–0.820)	0.810/0.841/0.845	Not separately assessed	8.915 (7.220–11.010), < 0.001
External validation cohort	271/87	68	0.700 (0.646–0.755)	0.604/0.705/0.726	Overall error 0.0073/0.0203/0.0809	3.558 (2.276–5.563), < 0.001

The final model C-indexes were 0.813 (95% CI, 0.798–0.828), 0.805 (95% CI, 0.790–0.820), and 0.700 (95% CI, 0.646–0.755) in the development, internal validation, and external validation cohorts, respectively. External 1-, 3-, and 5-year AUCs were 0.604, 0.705, and 0.726 (Fig. 3). Bootstrap internal calibration showed mean absolute errors of 0.0134, 0.0242, and 0.0265 at 1, 3, and 5 years (Fig. 4). Exploratory external calibration suggested close agreement at 1 and 3 years but optimism at 5 years; the 5-year mean predicted OS was 0.7888 compared with observed Kaplan-Meier OS of 0.7079 (Fig. 5).

Three-year DCA suggested that TNM plus CEA provided net benefit over treat-all and treat-none strategies across most clinically relevant thresholds and showed modest benefit over TNM-only across many thresholds (Fig. 6). Risk stratification by the development-cohort median linear predictor separated survival in all cohorts (Fig. 7). High-risk versus low-risk HRs were 7.454 (95% CI, 6.114–9.088), 8.915 (95% CI, 7.220–11.010), and 3.558 (95% CI, 2.276–5.563) in the development, internal validation, and external validation cohorts, respectively (all $P < 0.001$).

Discussion

This study developed and externally validated a simple TNM plus CEA nomogram for OS after surgery in EOCRC. The final model used four routinely available predictors: AJCC T stage, AJCC N stage, AJCC M stage, and pretreatment CEA. It showed good discrimination in SEER development and internal validation cohorts and moderate discrimination in an independent hospital validation cohort. Risk stratification consistently identified groups with markedly different survival across all three cohorts.

The main contribution of this study is not a larger set of predictors, but a more transparent handling of a staging-based prediction question. Because three of the four final predictors were AJCC T/N/M categories, TNM-only was the most clinically relevant benchmark. Adding CEA improved likelihood-based fit and modestly improved C-index and time-dependent AUC in SEER cohorts, while the external incremental advantage was smaller and less stable. This supports describing the model as a pragmatic TNM plus CEA tool with modest added value, rather than as a broad replacement for TNM staging.

Tumor grade required particular attention. LASSO selected grade, and the full model including grade fitted better by likelihood criteria. Grade should therefore not be described as non-prognostic. The problem was that grade behaved as a time-dependent prognostic factor in a standard Cox model and drove a major global PH violation. A grade-stratified Cox model resolved this concern, including the mild AJCC N signal seen in Model B, suggesting that some apparent N-stage time dependence may have reflected grade-related heterogeneity in baseline hazards. However, the discrimination gains from Model C were small, and routine use would require grade-specific baseline survival scales. We therefore treated Model B as the pragmatic primary point-based model and Model C as a sensitivity analysis.

The final predictors are clinically plausible. AJCC T stage reflects depth of invasion, AJCC N stage reflects regional nodal involvement, and AJCC M stage reflects metastatic burden [14]. Pretreatment CEA is a widely available biomarker with recognized prognostic relevance in CRC [20]. The direction and magnitude of hazard ratios followed expected clinical patterns, supporting face validity of the model.

External validation is a strength of this study, but it also highlights limitations of transportability. The model discriminated well in SEER but less strongly in the external hospital cohort. The hospital cohort was smaller and

came from a different country and care setting. Its baseline characteristics also differed from those of the SEER cohorts. Calibration was also less uniform externally, particularly at 5 years. Thus, the model may be most useful for postoperative risk ranking and follow-up discussion, while absolute long-term survival estimates should be interpreted cautiously until larger multicenter external validation or recalibration is performed.

This study has limitations. It was retrospective and subject to selection bias and residual confounding. SEER lacks molecular subtype, mismatch repair status, family history, comorbidities, margin status, recurrence, detailed systemic therapy, and treatment sequence. The external cohort was single-center and relatively small. The available audit describes the finalized analysis dataset rather than a complete original source-level exclusion flow. AJCC M category lacked an M1c level, so peritoneal metastasis as a distinct AJCC 8th-edition category could not be evaluated. Lymph-node-yield category was considered during candidate screening because it is clinically expected in colorectal cancer studies, but it was not selected by LASSO and was not included in the final model. Finally, although Model B satisfied the global PH test, the Model C sensitivity analysis shows that grade-related baseline-hazard heterogeneity remains an important modeling consideration.

Conclusions

A four-predictor nomogram based on AJCC T stage, AJCC N stage, AJCC M stage, and pretreatment CEA provided a simple and externally validated tool for OS prediction in surgically treated EOCRC. The model should be interpreted as a pragmatic TNM plus CEA model with modest incremental value over TNM-only staging. Grade-stratified sensitivity analysis supported the handling of tumor grade as a modeling tradeoff rather than evidence that grade lacks prognostic relevance. Further multicenter external validation and possible recalibration are needed before using the model for absolute long-term risk communication.

Abbreviations

AJCC

American Joint Committee on Cancer

AUC

area under the receiver operating characteristic curve

CEA

carcinoembryonic antigen

CI

confidence interval

CRC

colorectal cancer

DCA

decision curve analysis

EOCRC

early-onset colorectal cancer

HR

hazard ratio

LASSO

least absolute shrinkage and selection operator

OS
overall survival
PH
proportional hazards
SEER
Surveillance,Epidemiology,and End Results
TNM
tumor-node-metastasis
TRIPOD
Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki. The SEER database contains de-identified publicly available data and does not require additional institutional review board approval for approved users. For the external validation cohort, the Ethics Committee of the Second Affiliated Hospital of Harbin Medical University waived both the requirement for ethics approval and the requirement for informed consent, given the retrospective design of the study and the use of de-identified data.

Consent for publication

Not applicable. This manuscript does not contain any individual person's identifiable data.

Availability of data and materials

The SEER data used in this study are available from the SEER Program after approval of a data-use agreement. The external validation data from the Second Affiliated Hospital of Harbin Medical University are not publicly available because of institutional and privacy restrictions, but may be available from the corresponding author upon reasonable request and after appropriate institutional approval. The R scripts used for data preprocessing, model development, validation, and figure generation are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the Chen Xiao-Ping Foundation for the Development of Science and Technology of Hubei Province, grant number CXPJJH121001-2021028.

Authors' contributions

JG, HX, ZC and QT conceived and designed the study. JG, JW, XL, SP and QL collected and curated the data. JG, HX and ZC performed the statistical analysis and prepared the figures and tables. All authors contributed to

interpretation of the data. JG, HX and ZC wrote the main manuscript text. JW, XL, SP, QL and QT critically revised the manuscript for important intellectual content. QT supervised the study and obtained funding. All authors read and approved the final manuscript.

Acknowledgements

The authors acknowledge the SEER Program for providing publicly available cancer registry data.

Use of generative AI

Generative AI tools were used solely for language editing. Specifically, Claude (Anthropic) was used to improve the grammar, clarity, and readability of English text originally written by the authors. No generative AI tool was used to design the study, generate or analyse data, produce or interpret results, create figures or tables, or draft original scientific content. All statistical analyses were performed by the authors using R version 4.5.1. The authors reviewed and edited all AI-assisted text and take full responsibility for the accuracy and integrity of the manuscript.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74:229–63.
2. Han B, Zheng R, Zeng H, Wang S, Sun K, Chen R, et al. Cancer incidence and mortality in China, 2022. *J Natl Cancer Cent.* 2024;4:47–53.
3. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. *CA Cancer J Clin.* 2023;73:233–54.
4. Patel SG, Karlitz JJ, Yen T, Lieu CH, Boland CR. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. *Lancet Gastroenterol Hepatol.* 2022;7:262–74.
5. Akimoto N, Ugai T, Zhong R, Hamada T, Fujiyoshi K, Giannakis M, et al. Rising incidence of early-onset colorectal cancer: a call to action. *Nat Rev Clin Oncol.* 2021;18:230–43.
6. Bailey CE, Hu CY, You YN, Bednarski BK, Rodriguez-Bigas MA, Skibber JM, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975–2010. *JAMA Surg.* 2015;150:17–22.
7. Montminy EM, Zhou M, Maniscalco L, Penrose H, Yen T, Patel SG, et al. Contributions of adenocarcinoma and carcinoid tumors to early-onset colorectal cancer incidence rates in the United States. *Ann Intern Med.* 2021;174:157–66.
8. Stoffel EM, Murphy CC. Epidemiology and mechanisms of the increasing incidence of colon and rectal cancers in young adults. *Gastroenterology.* 2020;158:341–53.
9. Zaborowski AM, Abdile A, Adamina M, Aigner F, d'Allens L, Allmer C, et al. Characteristics of early-onset vs late-onset colorectal cancer: a review. *JAMA Surg.* 2021;156:865–74.
10. Bhandari A, Woodhouse M, Gupta S. Colorectal cancer is a leading cause of cancer incidence and mortality among adults younger than 50 years in the USA: a SEER-based analysis with comparison to other young-

- onset cancers. *J Investig Med*. 2017;65:311–5.
11. Willauer AN, Liu Y, Pereira AAL, Lam M, Morris JS, Raghav KPS, et al. Clinical and molecular characterization of early-onset colorectal cancer. *Cancer*. 2019;125:2002–10.
 12. Andric F, Al-Fairouzi A, Wettergren Y, Szeponik L, Bexe-Lindskog E, Cusack JC, et al. Immune microenvironment in sporadic early-onset versus average-onset colorectal cancer. *Cancers*. 2023;15:1457.
 13. Demb J, Gomez SL, Canchola AJ, Qian AS, Earles A, Liu L, et al. Racial and ethnic variation in survival in early-onset colorectal cancer. *JAMA Netw Open*. 2024;7:e2446820.
 14. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A, editors. *AJCC cancer staging manual*. 7th ed. New York: Springer; 2010.
 15. Yu C, Zhang Y. Development and validation of a prognostic nomogram for early-onset colon cancer. *Biosci Rep*. 2019;39:BSR20181781.
 16. Wu J, Lu L, Chen H, Lin Y, Zhang H, Chen E, et al. Prognostic nomogram to predict the overall survival of patients with early-onset colorectal cancer: a population-based analysis. *Int J Colorectal Dis*. 2021;36:1981–93.
 17. Liu J, Liu Z, Li J, Tian S, Dong W. Personalizing prognostic prediction in early-onset colorectal cancer. *J Cancer*. 2020;11:6727–36.
 18. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med*. 2015;13:1.
 19. Hankey BF, Ries LA, Edwards BK. The Surveillance, Epidemiology, and End Results Program: a national resource. *Cancer Epidemiol Biomarkers Prev*. 1999;8:1117–21.
 20. Becerra AZ, Probst CP, Tejani MA, Aquina CT, González MG, Hensley BJ, et al. Evaluating the prognostic role of elevated preoperative carcinoembryonic antigen levels in colon cancer patients: results from the National Cancer Database. *Ann Surg Oncol*. 2016;23:1554–61.
 21. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16:385–95.
 22. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81:515–26.
 23. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–87.
 24. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE Jr, Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: part II—binary and time-to-event outcomes. *Stat Med*. 2019;38:1276–96.
 25. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56:337–44.
 26. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32:5381–97.
 27. Pencina MJ, D'Agostino RB, Sr, Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27:157–72.
 28. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak*. 2006;26:565–74.
 29. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*.

Figures

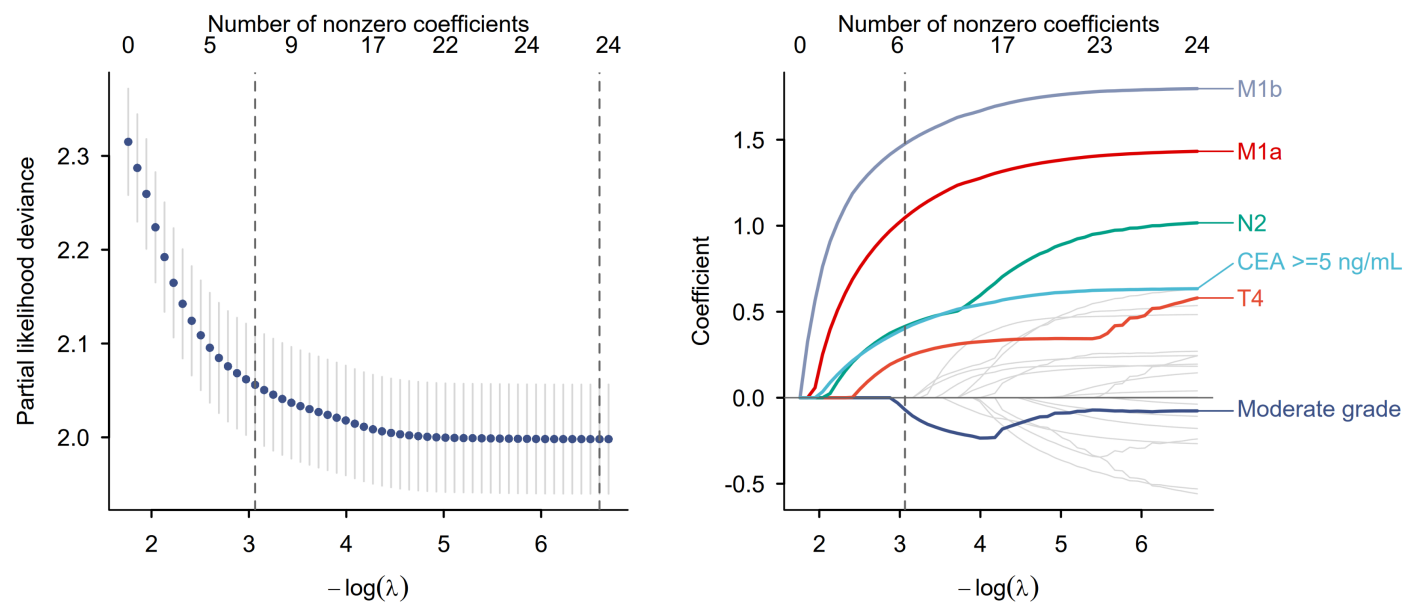


Figure 1

LASSO Cox predictor selection. Ten-fold cross-validation and coefficient paths are shown for candidate-predictor screening in the development cohort. The lambda.1se rule was used to favor a parsimonious candidate set.

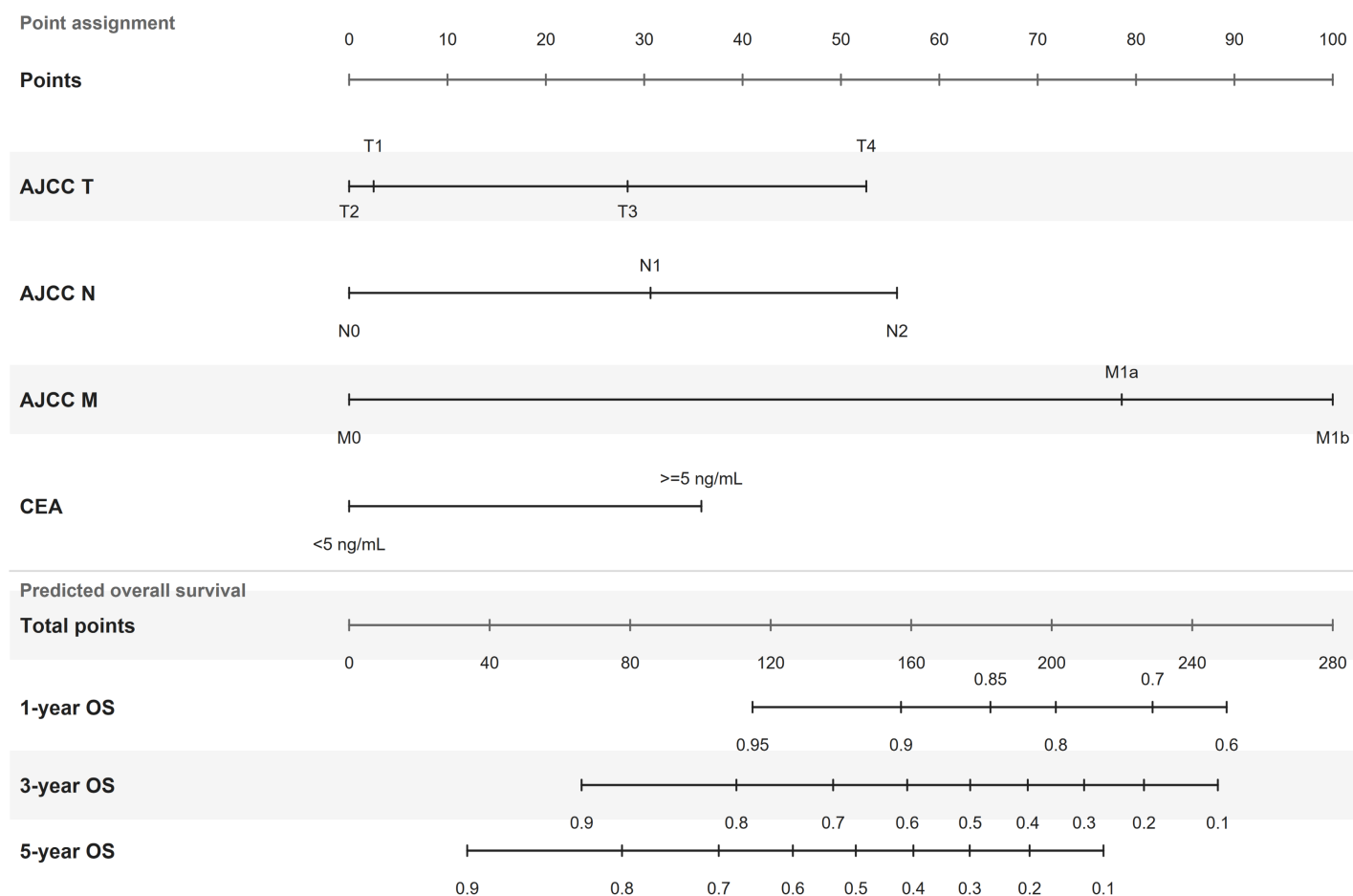


Figure 2

Nomogram for individualized prediction of 1-, 3-, and 5-year overall survival. The nomogram was constructed from the final fixed Cox model including AJCC T stage, AJCC N stage, AJCC M stage, and pretreatment CEA. Higher total points indicate lower predicted survival.

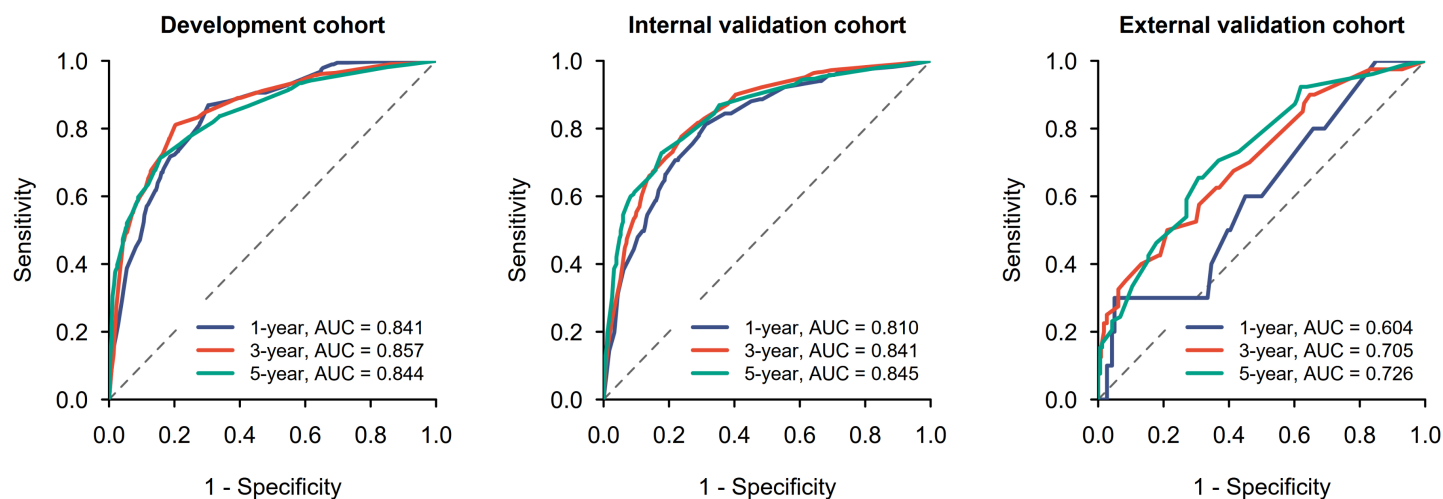


Figure 3

Time-dependent discrimination of the final model. Receiver operating characteristic curves are shown at 1, 3, and 5 years for the development cohort, internal validation cohort, and external validation cohort.

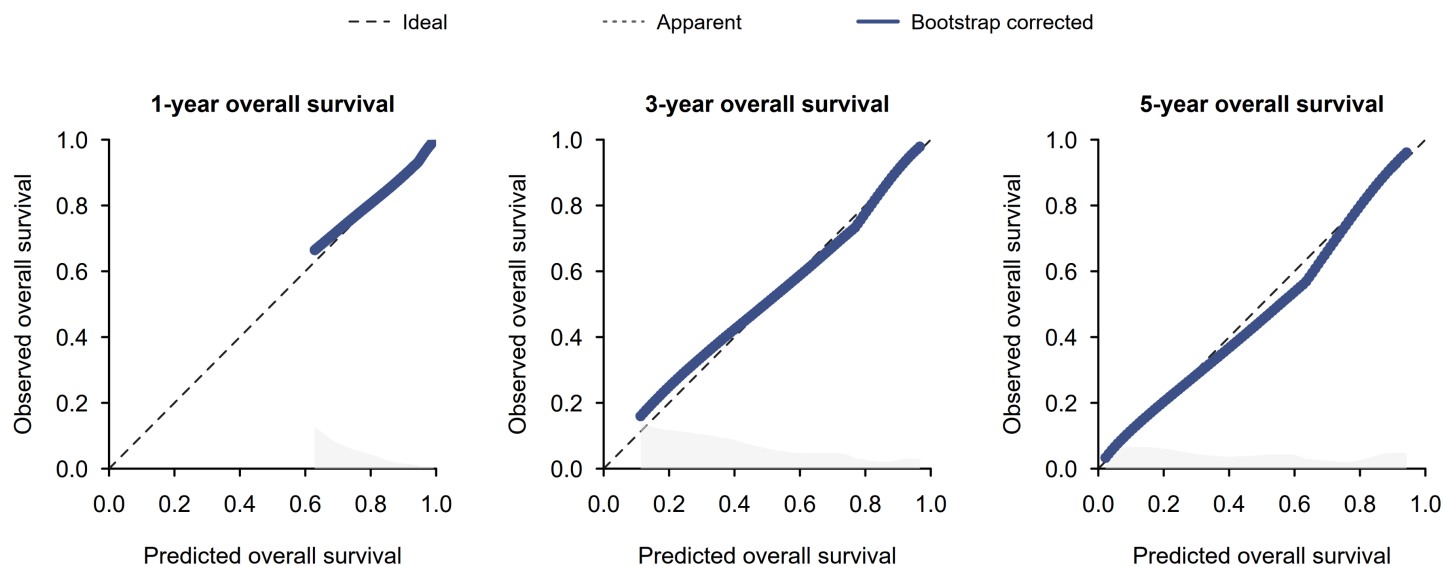


Figure 4

Bootstrap internal calibration of the final model. Calibration curves are shown for predicted 1-, 3-, and 5-year overall survival in the development cohort.

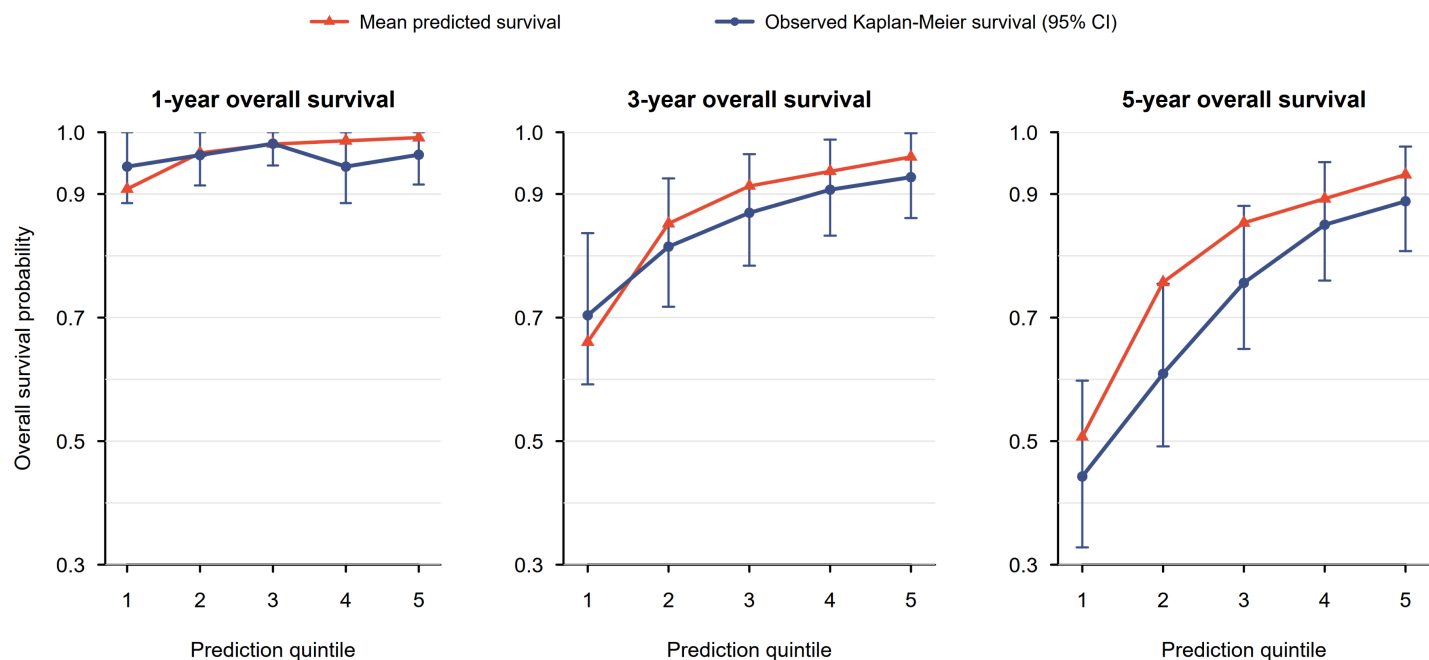


Figure 5

Exploratory grouped external calibration. Mean predicted 1-, 3-, and 5-year overall survival is compared with Kaplan-Meier observed survival by prediction quintile in the external validation cohort.

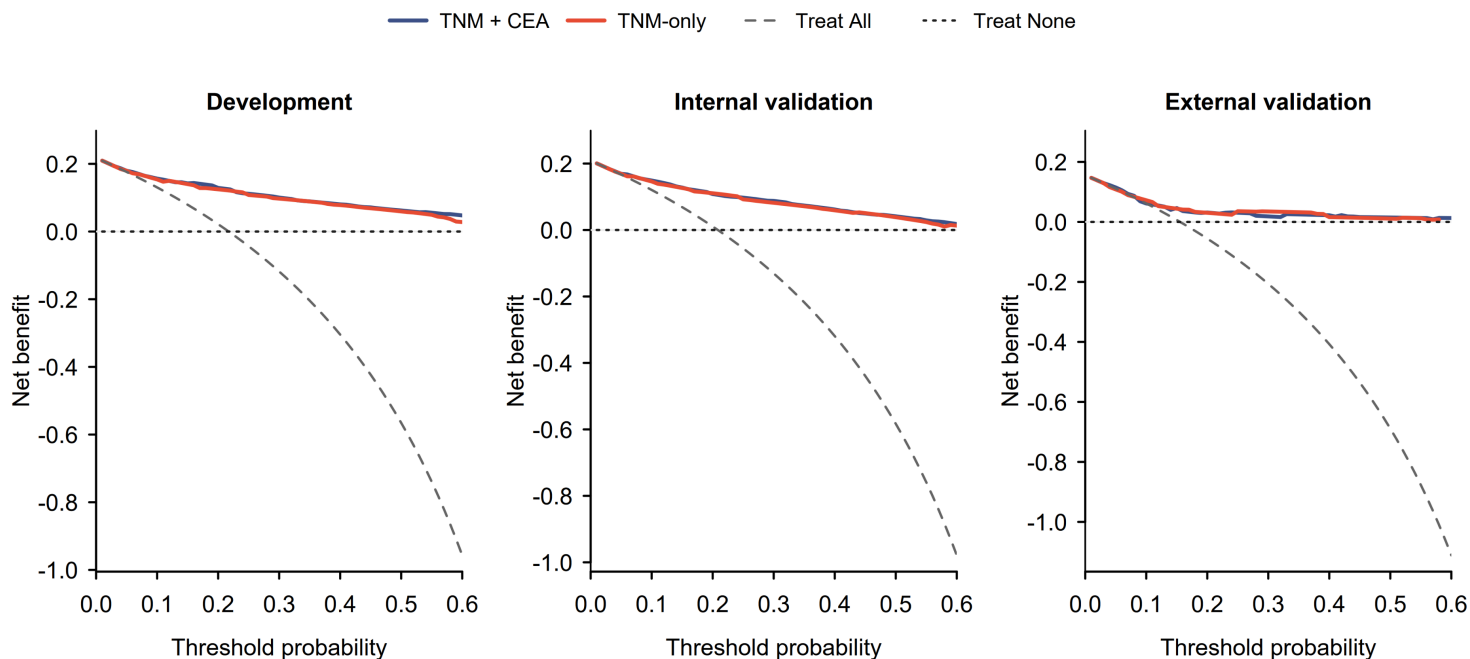


Figure 6

Three-year decision curve analysis. The final TNM plus CEA model is compared with TNM-only, treat-all, and treat-none strategies in the development cohort, internal validation cohort, and external validation cohort.

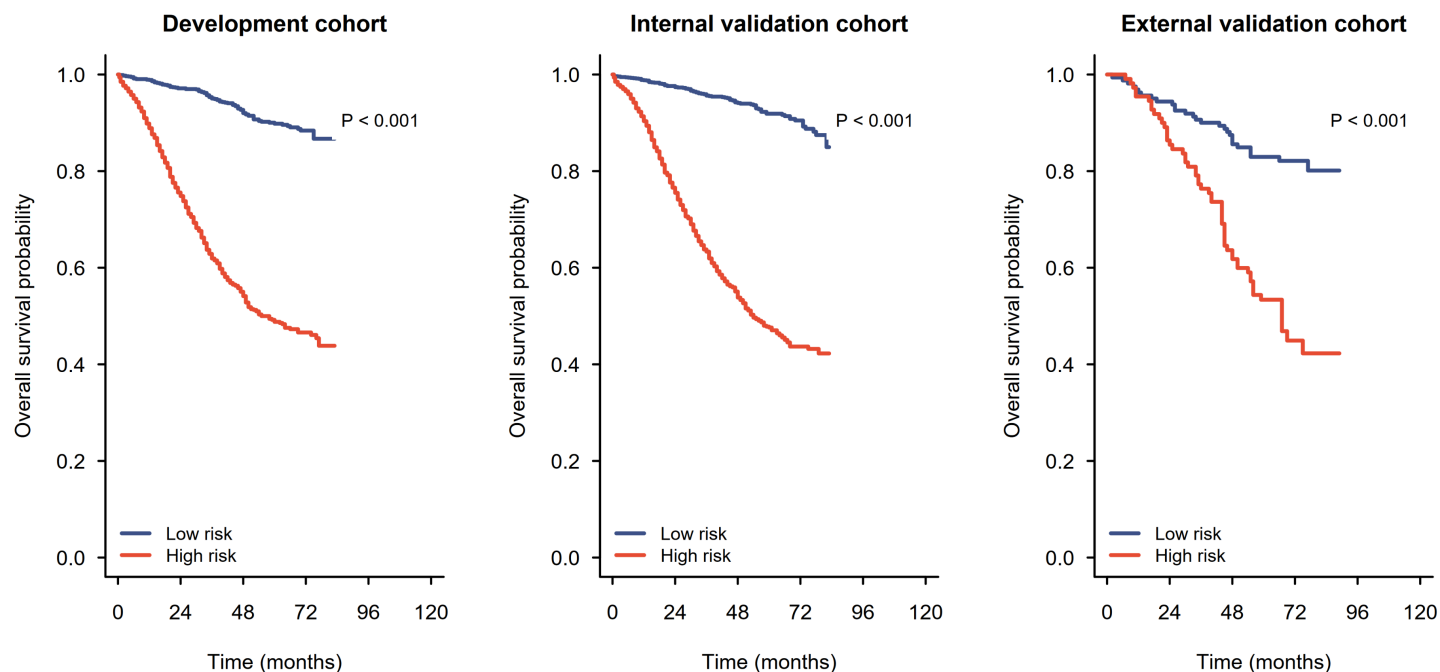


Figure 7

Kaplan-Meier risk stratification. Survival curves are shown for high-risk and low-risk groups defined by the development-cohort median linear predictor.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterialS1BMC.docx](#)