



Prognostic and Predictive Model of the Log Odds of the Negative Lymph Node/T Stage Ratio in Resectable Gastric Adenocarcinoma Patients

Wei Yang¹ · Shuaibing Lu¹ · Fusheng Ge¹ · Yawei Hua¹  · Xiaobing Chen^{2,3}

Received: 28 April 2022 / Accepted: 24 June 2022 / Published online: 14 July 2022
© The Society for Surgery of the Alimentary Tract 2022

Abstract

Purpose There are few reports on disease-specific survival (DSS) prediction systems for resected gastric cancer (GC) patients. The aim of this study was to create a nomogram based on the log odds of the negative lymph node/T stage ratio (LONT) for individual risk prediction.

Methods We applied the Surveillance, Epidemiology, and End Results (SEER) Program database released in 2021 to screen GC patients from 2010 to 2015. Using a competitive risk model, we plotted the cumulative risk curve of variables for gastric cancer–specific death and death from other causes at each time point. According to the minimum BIC, we constructed and assessed a nomogram for the 12-month, 36-month, and 60-month cumulative mortality probabilities assessed by time-dependent ROC curves (time-AUCs), the C-index, Brier scores, decision curve analysis (DCA), and calibration curves.

Results A total of 3895 patients were ultimately included and randomly assigned to two sets: the training set ($n = 2726$, 70%) and the validation set ($n = 1169$, 30%). The LONT was a remarkable independent predictor of gastric cancer–specific death (high versus low: 0.705, 95% CI 0.524–0.95, $p = 0.021$). The variables selected based on the minimum BIC were as follows: location, AJCC, AJCC.T, AJCC.N, radiotherapy, LONT.cat, and chemotherapy. According to the time-AUC, C-index, Brier score, DCA, and calibration curves, the nomogram risk score had excellent survival prediction ability for DSS.

Conclusions A low LONT was associated with a high cumulative incidence of DSS. A prognostic nomogram model based on the LONT could effectively predict DSS for resectable GC patients.

Keywords Nomogram · Gastric cancer · Negative lymph nodes · Negative lymph nodes/T stage · Prognosis

Wei Yang and Shuaibing Lu contributed equally to this work and should be considered co-first authors.

✉ Yawei Hua
huayawei1965@163.com

✉ Xiaobing Chen
zlychenxb0807@zzu.edu.cn

¹ Department of General Surgery, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, No. 127 Dongming Rd, Zhengzhou 450008, Henan, China

² Department of Medical Oncology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China

³ Zhengzhou Key Laboratory for Precision Therapy of Gastrointestinal Cancer, No. 127 Dongming Rd, Zhengzhou 450008, China

Introduction

Gastric cancer (GC) is the fifth most common cancer worldwide and a serious threat to human health, being the third most common cause of cancer death.^{1,2} A diagnosis is made histologically after endoscopic biopsy and staging is performed using CT, ultrasound id, PET, and laparoscopy. Surgical treatment for locally advanced operable gastric cancer should include D2 lymphadenectomy (including removal of the perigastric lymph nodes and lymph nodes along the branches of the celiac artery).

In recent years, an increasing number of studies have shown that the number of lymph nodes (ELNs) examined^{3,4} and the number of LNS-negative lymph nodes (NLNs)^{5–7} are independent risk factors affecting prognosis. Indeed, T staging, which is based on the depth of tumor invasion, can represent the main features of the tumor^{8–10}; therefore, it is a reliable prognostic factor for gastric cancer. However, whether a combination of the two aspects can be used as a novel and more accurate prognostic

indicator to reflect the outcome of patients remains unclear, and there are no reports in this regard. Therefore, in our study, a new evaluation method was proposed: the logarithmic ratio of the number of negative lymph nodes removed (NLN) to the tumor size (T) ($\text{Log}[(\text{NLNs} + 1)/\text{T stage}]$, LONT).

The Surveillance, Epidemiology, and End Results (SEER) database is a widely applied dataset of clinical data on cancer patients. In this study, we selected gastric cancer patients from the SEER database over a long time period (2010–2015). First, based on the competitive risk model, we investigated the prognostic value of the LONT for resectable gastric adenocarcinoma

without metastasis at first diagnosis. Second, we built a new prognostic nomogram model to assess the prognosis of GC patients undergoing surgical treatment.

Materials and Methods

Data Sources

We used the SEER*Stat 8.3.8 software released in 2021 for screening GC patients from 2010 to 2015. We used the

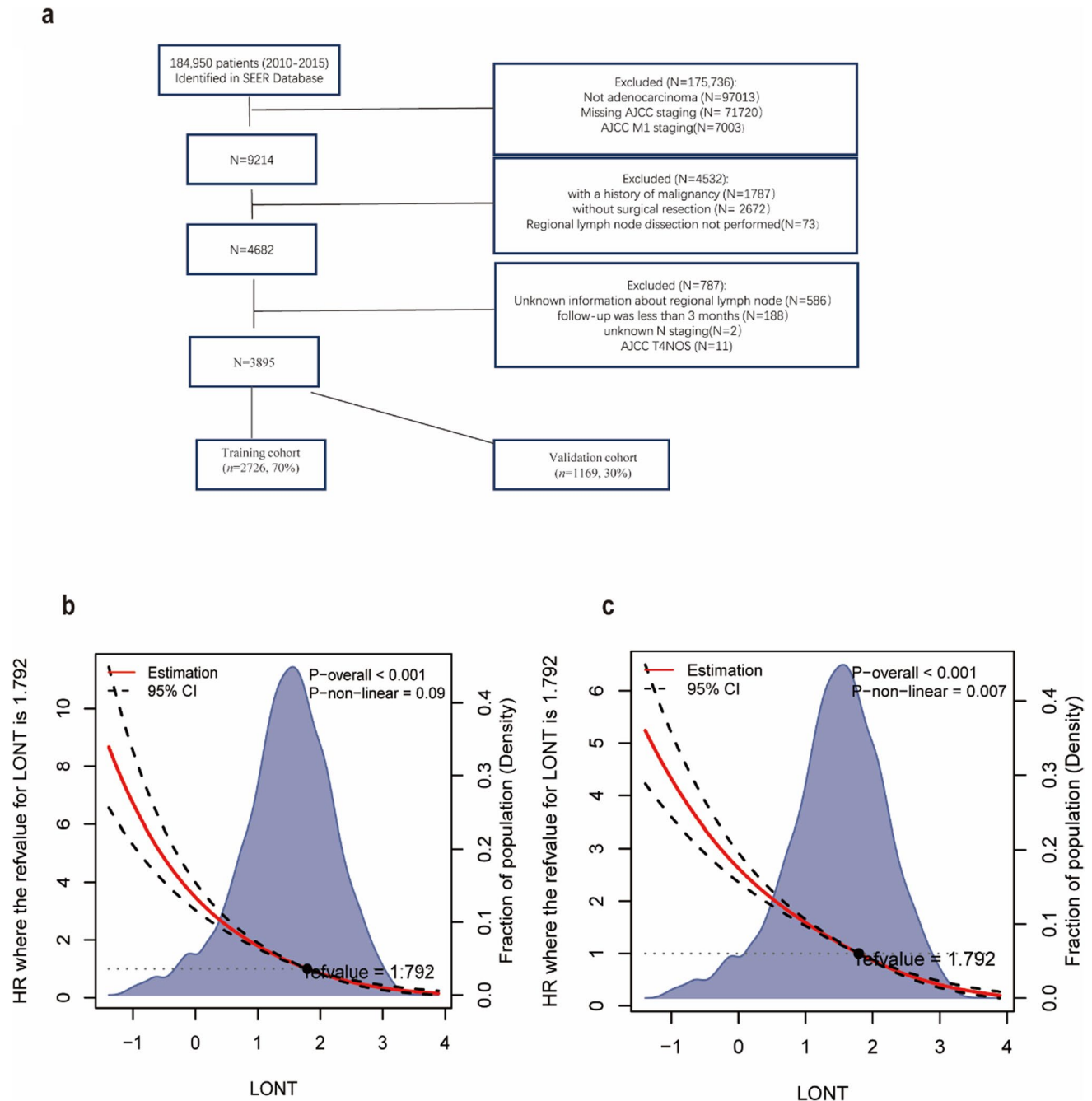


Fig. 1 Flowchart of patient selection (a); restricted cubic spline (RCS) of the LONT for OS (b) and DSS (c)

Table 1 Baseline clinicopathological characteristics of patients in the training and validation cohort

Variables	Total (n = 3895)	Train (n = 2726)	Validation (n = 1169)	p
Survival months, Median (Q1, Q3)	27 (15, 47)	27 (15, 47)	27 (16, 48)	0.263
Outcome, n (%)				0.586
Death	1632 (42)	1134 (42)	498 (43)	
Live	2263 (58)	1592 (58)	671 (57)	
Outcome3, n (%)				0.096
Death from gastric cancer	771 (20)	555 (20)	216 (18)	
Death from Others	861 (22)	579 (21)	282 (24)	
Live	2263 (58)	1592 (58)	671 (57)	
Age, median (Q1,Q3)	66 (58, 74)	66 (57, 74)	66 (58, 74)	0.173
Sex, n (%)				0.899
Male	2775 (71)	1940 (71)	835 (71)	
Female	1120 (29)	786 (29)	334 (29)	
Race, n (%)				0.49
Others	692 (18)	472 (17)	220 (19)	
Black	447 (11)	311 (11)	136 (12)	
White	2756 (71)	1943 (71)	813 (70)	
Diagnosis, n (%)				0.143
2010	639 (16)	462 (17)	177 (15)	
2011	665 (17)	457 (17)	208 (18)	
2012	649 (17)	431 (16)	218 (19)	
2013	647 (17)	448 (16)	199 (17)	
2014	671 (17)	485 (18)	186 (16)	
2015	624 (16)	443 (16)	181 (15)	
Location, n (%)				0.784
Cardia	1858 (48)	1299 (48)	559 (48)	
Antrum/Pylorus	841 (22)	599 (22)	242 (21)	
Middle	830 (21)	572 (21)	258 (22)	
Others	366 (9)	256 (9)	110 (9)	
AJCC, n (%)				0.779
III	1709 (44)	1204 (44)	505 (43)	
I	1082 (28)	758 (28)	324 (28)	
II	1104 (28)	764 (28)	340 (29)	
AJCC.T, n (%)				0.527
T4b	148 (4)	109 (4)	39 (3)	
T1	946 (24)	651 (24)	295 (25)	
T2	586 (15)	423 (16)	163 (14)	
T3	1801 (46)	1258 (46)	543 (46)	
T4a	414 (11)	285 (10)	129 (11)	
AJCC.N, n (%)				0.828
N3	527 (14)	373 (14)	154 (13)	
N0	1672 (43)	1157 (42)	515 (44)	
N1	1059 (27)	746 (27)	313 (27)	
N2	637 (16)	450 (17)	187 (16)	
Radiotherapy, n (%)				0.127
NO	2118 (54)	1511 (55)	607 (52)	
Unknown	54 (1)	36 (1)	18 (2)	
YES	1723 (44)	1179 (43)	544 (47)	
Chemotherapy, n (%)				0.263
NO	1456 (37)	1035 (38)	421 (36)	
YES	2439 (63)	1691 (62)	748 (64)	
Differentiation, n (%)				0.191

Table 1 (continued)

Variables	Total (n=3895)	Train (n=2726)	Validation (n=1169)	p
G3~G4	2077 (53)	1430 (52)	647 (55)	
G1	277 (7)	207 (8)	70 (6)	
G2	1328 (34)	936 (34)	392 (34)	
Unknown	213 (5)	153 (6)	60 (5)	
NLNs, median (Q1,Q3)	13 (7, 20)	13 (7, 20)	12 (7, 20)	0.324
Tumor size, n (%)				0.527
1	946 (24)	651 (24)	295 (25)	
2	586 (15)	423 (16)	163 (14)	
3	1801 (46)	1258 (46)	543 (46)	
4	414 (11)	285 (10)	129 (11)	
5	148 (4)	109 (4)	39 (3)	
LONT, median (Q1, Q3)	1.79 (1.1, 2.37)	1.79 (1.1, 2.37)	1.73 (1.18, 2.37)	0.679
Age.cat, n (%)				0.058
Age ≥ 60	2738 (70)	1891 (69)	847 (72)	
Age < 60	1157 (30)	835 (31)	322 (28)	
LONT.cat, n (%)				0.189
Low	2055 (53)	1419 (52)	636 (54)	
High	1840 (47)	1307 (48)	533 (46)	
NLNs.cat, n (%)				0.224
0~12	1876 (48)	1291 (47)	585 (50)	
13~24	1385 (36)	992 (36)	393 (34)	
25~89	634 (16)	443 (16)	191 (16)	

primary site codes C16.0–C16.9 for gastric cancer and the International Classification of Diseases for Oncology, Third Edition (ICDO-3) histologic codes 8140/3 (adenocarcinoma, NOS).¹¹ Finally, we identified a total of 184,950 cases between 2010 and 2015.

Follow-up data were collected up to the most recent date (December 31, 2020). SEER-Stat (version 8.3.5) was used to filter and collect the data of the representative patients (<http://seer.cancer.gov/>).

Ethics Review

The public data did not involve patient-specific data, so according to Swedish law, the consent of the ethics committee was not required.

Inclusion and Exclusion Criteria

Only cases who met the following criteria were included: (1) more than 18 years of age; (2) pathologically confirmed disease; (3) known GC molecular subtypes; (4) only one primary malignancy; and (5) active follow-up with date, survival duration, and known cause of death.

Cases were excluded if they encompassed the following criteria: (1) lacked a detailed description of the surgery; (2) had stage I to III GC and did not undergo radical excision;

and (3) contained any missing data for the baseline information (e.g., race, site, and grade).

Data Extraction and Preprocessing

Information on the following variables of interest was collected: age, sex, race, time of diagnosis, location of primary focus, AJCC stage, AJCC T stage, size of primary tumor, AJCC N stage, radiotherapy, chemotherapy, number of negative lymph nodes (NLNs), pathological grade, survival time, and outcome (OS/DSS and competitive risk outcome). The TNM status of each case was reassessed according to the 8th edition of the American Joint Committee on Cancer (AJCC) Cancer Staging Manual based on CS extension and regional node positivity in the SEER database.

Outcomes were defined as follows: survival was defined as 0, disease-specific death (gastric cancer) was defined as 1, and death not caused by gastric cancer was defined as 2 (competitive event). The time of diagnosis or case inclusion period was 2010 to 2015. Continuous age values were transformed to dichotomous variables according to the median method: < 60 years old vs. ≥ 60 years old. Race was divided into three levels by frequency distribution: white vs. black vs. others. Location was converted into four levels on the basis of frequency distribution: cardia vs. middle (gastric fundus and gastric body

and small bend and big bend) vs. antrum and pylori vs. others (cross region and NOS). Differentiation was transformed into three classifications on the basis of frequency distribution: G3 + 4 (Ref) vs. G2 vs. G1. AJCC staging was translated into three groups according to frequency: I vs. II vs. III; AJCC. T was classified into five categories according to frequency: T1 vs. T2 vs. T3 vs. T4a vs. T4b; tumor size was designated as 1, 2, 3, 4, and 5 according to the AJCC classification (T: T1–4b; AJCC). N was classified into five categories according to frequency: N0 vs. N1 vs. N2 vs. N3a vs. N3b; whether radiotherapy and

chemotherapy were used was converted to two categories: Yes vs. No; and NLN was divided into three grades in terms of the change in RCS curve (see note): 0 ~ 12 vs. 13 ~ 24 vs. 25 ~ 89. Calculation of the main variable lot was as follows: T1-4b were designated as 1, 2, 3, 4, and 5, respectively, and then the LONT was calculated as $LONT = \text{Log}[(NLNs + 1)/T \text{ stage}]$. The primary endpoint, OS, was the time from the date of diagnosis to the time of death from any cause, and the secondary endpoint, DSS, was determined from the time of confirmed diagnosis to the date of death due to gastric cancer.

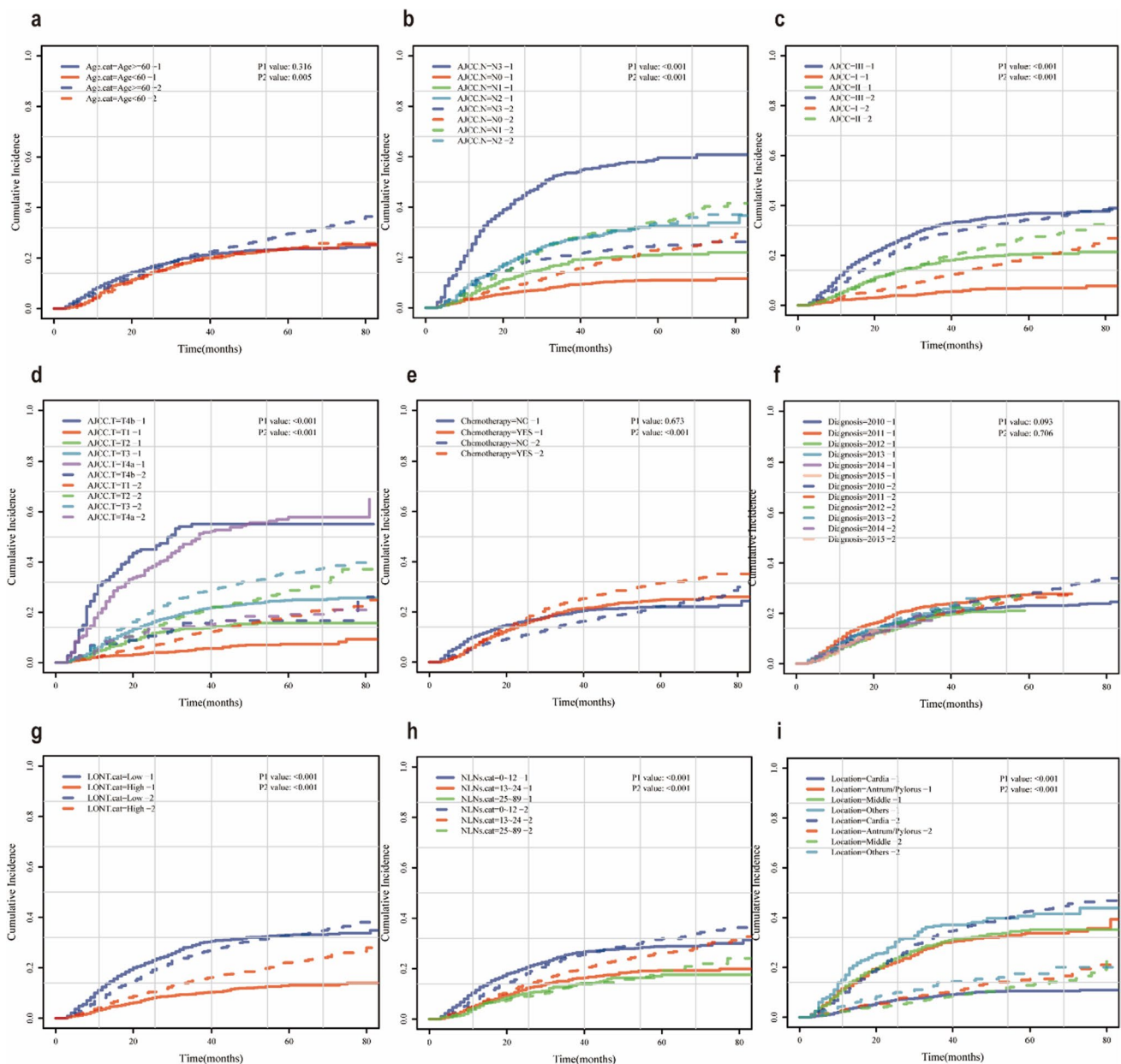


Fig. 2 Cumulative incidence function (CIF) curves for variables. **a** Age; **b** AJCC. N; **c** AJCC. T; **d** AJCC. T; **e** chemotherapy; **f** diagnosis; **g** LONT. cat; **h** NLNs.cat; **i** location

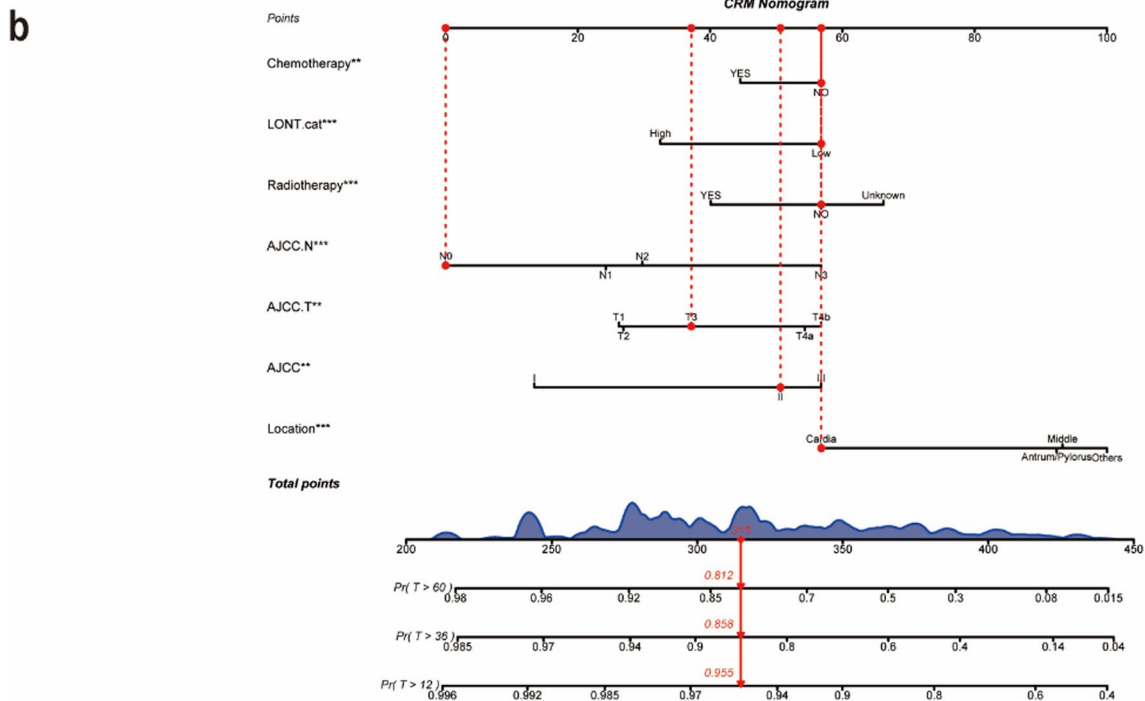
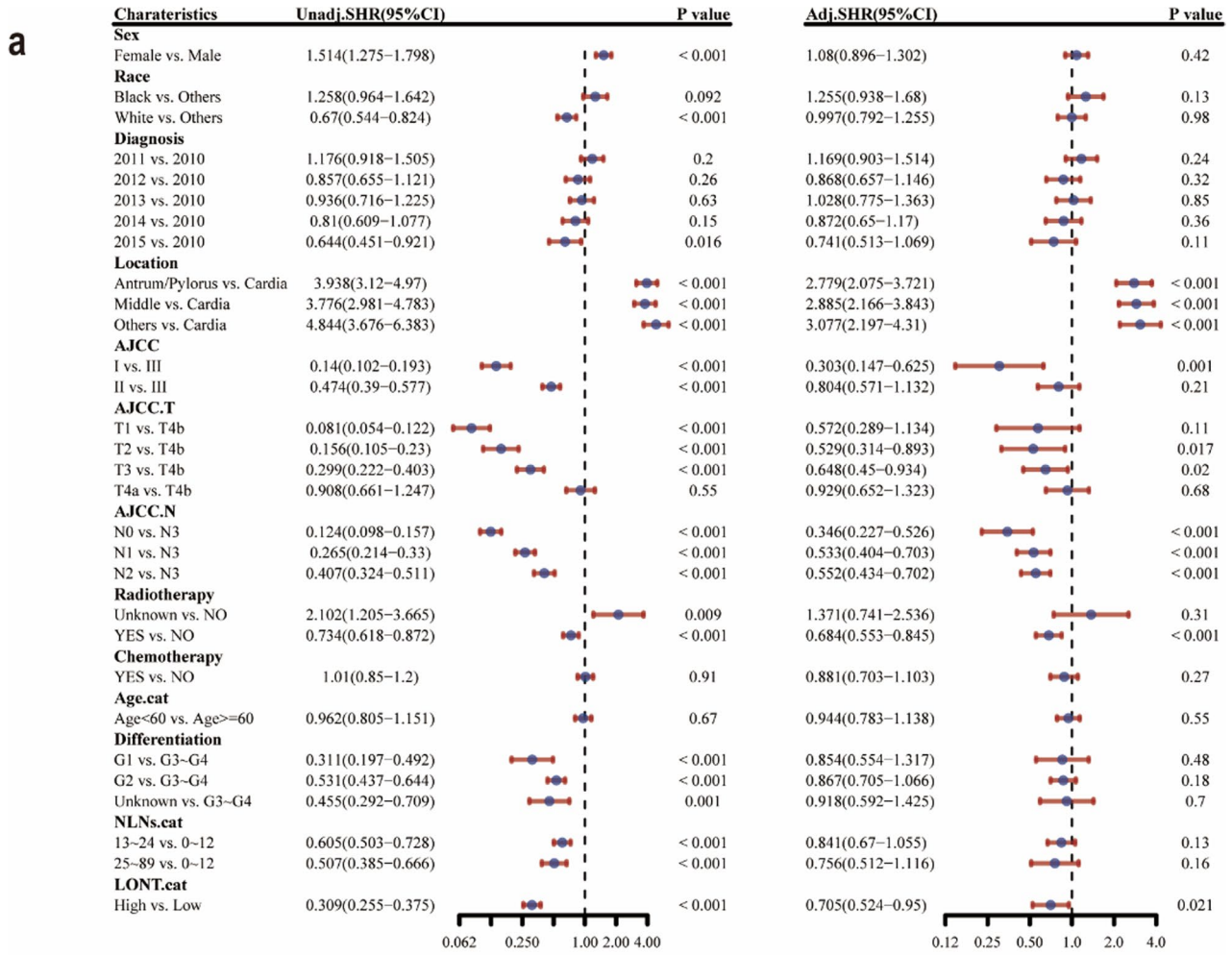


Fig. 3 **a** Univariate and multivariate analyses for evaluating the relationship between the LONT and cancer-specific mortality according to different clinicopathological factors in the training cohort. **b** The competing risk nomogram based on the minimum BIC to forecast the 12-month, 36-month, and 60-month cumulative cause-specific death probabilities in the training cohort

Statistical Analysis

Descriptive statistics were applied to check the characteristics of the baseline data in patients with GC. Continuous data are displayed as medians and interquartile ranges (if a normal distribution was not satisfied) or as means and standard deviations (if a normal distribution was satisfied). Categorical data are presented as counts and percentages. Continuous variables, such as age, were transformed into categorical variables. Pearson chi-square tests were used to analyze the classification variables. All *p*-values less than 0.05 were considered significant. All *p*-values were two-tailed.

Construction of the Competitive Risk Model

Data from eligible patients were collected, and the dataset was divided into a training set and an internal validation set at a ratio of 7:3 for cross-validation. First, restricted cubic spline (RCS) mapping of OS and DSS by the LONT was performed: according to RCS conditions, based on the best cutoff value for separating the groups, the LONT was converted to a binary variable: high vs. low. Second, definitions of the competing risk model, death caused by GC itself (gastric cancer-specific death) was an event of interest, death caused by other causes (nonspecific death) was a competition event, and loss of follow-up and survival by the end of follow-up were deletion events. Nelson–Aalen cumulative risk curves of the cumulative incidence of GC-specific death were also generated. By establishing the competing risk approach, cumulative risk curves of all variables were drawn using the R package *cmprsk* to assess the cumulative incidence function (CIF) for GC-specific death and death from other reasons at various points in time, and Gray's test was performed to recognize differences between two events in the CIF.

Univariate and multivariate analyses based on a competitive risk model of training concentration and Fine and Gray's competing risk regression were applied to screen variables affecting prognosis associated with GC-specific death. All statistical analyses were carried out using R statistical software (version 3.5.0, StataCorp LLC, College Station, Texas) and STATA (Version 13.0, StataCorp LLC) using the R package *cmprsk*. Second, the variables were screened based on the optimal subset method, and the included

variables were screened based on the minimum BIC. We constructed a Fine and Gray proportional subdistribution hazard model. According to the established Fine and Gray competing risk models in the training set, a nomogram was constructed based on the training set and drawn based on the *regplot* package in R language. This study drew and displayed nomograms of 12-month, 36-month, and 60-month cumulative mortality probabilities. Third, two models were constructed for comparison. The first was a model composed of all factors, and the second was a model of factors screened according to the BIC criteria.

Evaluation of Competitive Risk Models

The discrimination degree was evaluated with the time-dependent ROC curve and C-index (time AUC and C-index, respectively) to compare multiple indices: each single index vs. BIC criterion screening model vs. all factor models. We drew the calibration curve to evaluate the consistency of the model, that is, the difference between the predicted and actual values. The Brier score was used to calculate the relevant parameters at 12 months, 36 months, and 60 months. Model benefit decision curve analysis (DCA) was plotted at 12 months, 36 months, and 60 months. Then, based on the scores of the training set prediction model, the scores were divided into high/medium/low-risk groups by using the tertile method, and the cumulative risk curve was used to estimate the viability of the prediction model.

Model validation

Finally, in the validation set, we conducted univariate and multivariate analyses of the competitive risk model, and during the validation process, the optimal C-index, Brier score, DCA, and calibration curves were identified using the R package *pec* and *Desc Tools*.

Results

Baseline Characteristics

The data of 184,950 patients pathologically diagnosed with primary GC from 2010 to 2015 were collected from the SEER database. After screening according to the inclusion and exclusion criteria, a total of 3895 patients were included and randomly divided into two sets (Fig. 1a): the training set ($n = 2726$, 70%) (Supplementary Table 1) and the validation set ($n = 1169$, 30%) (Supplementary Table 2).

The latest follow-up date was December 2020. The median follow-up time was 27 months (range 15–47 months) in the training cohort and 27 months (range 16–48 months) in the validation cohort. There was no significant difference

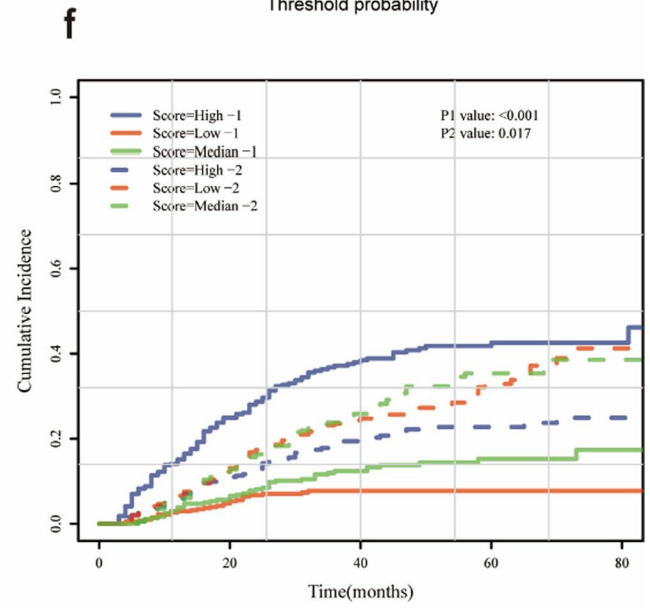
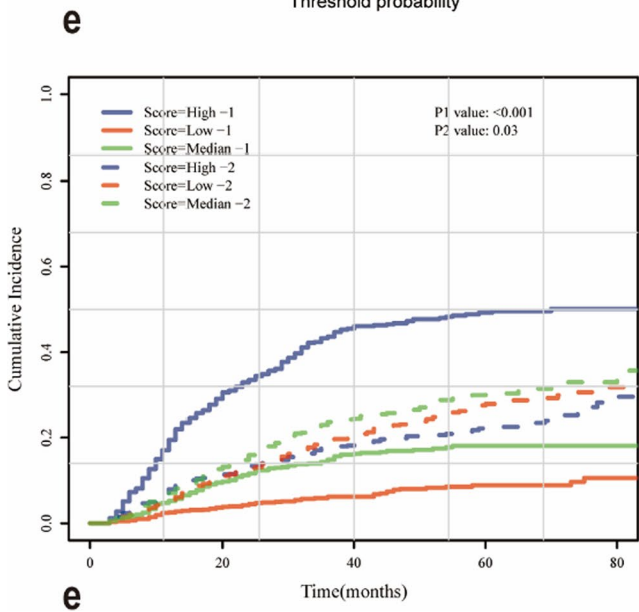
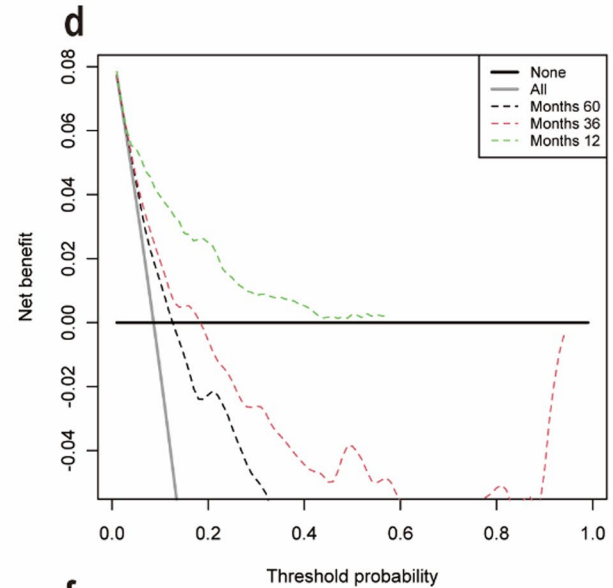
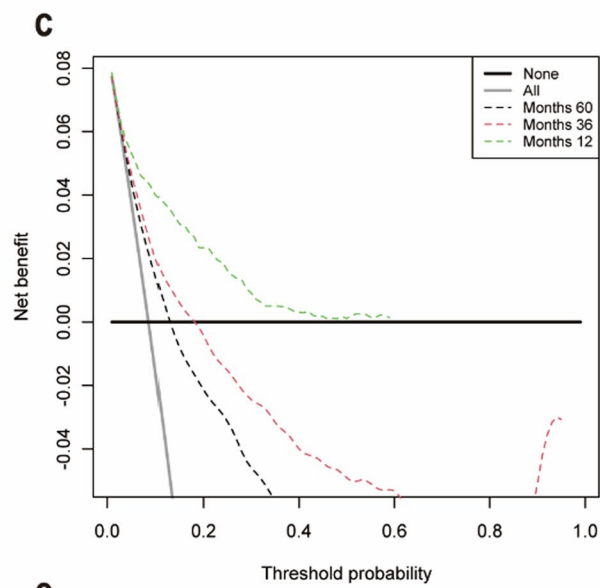
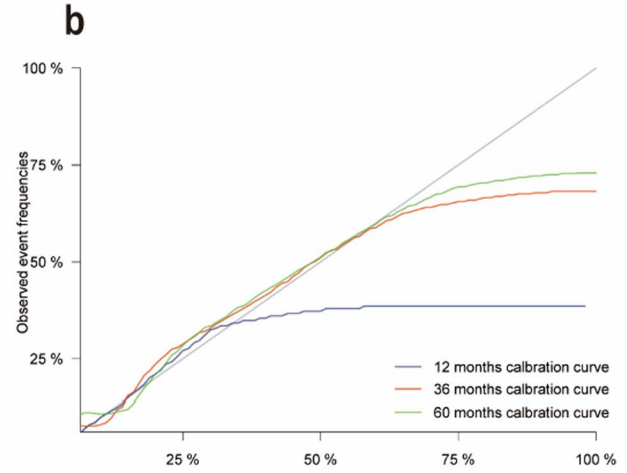
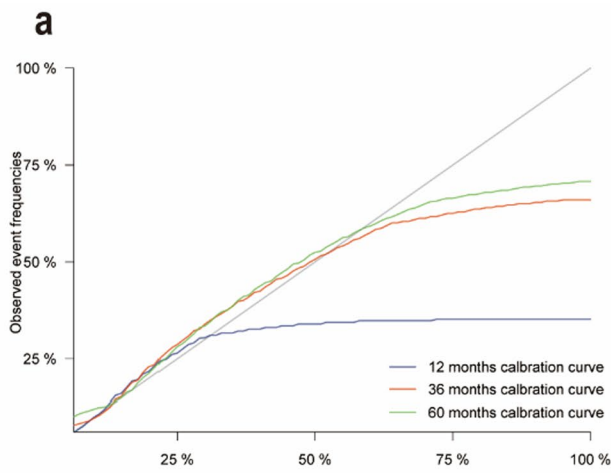


Fig. 4 Performance of the competing risk nomogram model. Calibration plots for the predicted and actual observed 12-, 36-, and 60-month cause-specific death probabilities in the all-factor model (a) and in the BIC model (b) in the training cohort. Decision curve analysis (DCA) for the predicted and actual observed 12-, 36-, and 60-month cause-specific death probabilities in the all-factor model (c) and in the BIC model (d) in the training cohort. Cumulative risk curve of the score based on the prediction model in the training (e) and validation (f) cohorts

in clinicopathological features between the two groups (all $p > 0.05$). The percentage of patients who were diagnosed with stomach cancer remained relatively constant between 2010 and 2015. According to restricted cubic spline (RCS) analysis, the best LONT cutoff value was determined to be 1.792 for distinguishing OS and DSS, and then the LONT was converted into a dichotomous variable: high vs. low, as shown in Fig. 1b, c.

In the training set, there were 1419 patients with a “low LONT,” 1307 patients with a “high LONT,” and there were 636 patients with a “low LONT” and 533 patients with a “high LONT” in the validation set. Male patients accounted for 71% in the two cohorts. Detailed information about the clinicopathological features is shown in Table 1.

Univariate Analysis by CIF

Twelve-month, 36-month, and 60-month cumulative incidences of cancer-specific mortality grouped by different variables are presented, and CIF curves for all variables are shown in Fig. 2, solid lines. Those patients with higher N staging, higher AJCC staging, higher AJCC T staging, higher differentiation grading, a low LONT.cat, a lower NLNs.cat, black race, no radiation therapy, and female sex were accompanied by high cumulative incidences of the study endpoint dying from GC.

There were higher cumulative incidences of dying from other causes in patients with advanced age ≥ 60 , higher N staging, higher AJCC staging, chemotherapy, a low LONT.cat, a lower NLNs.cat, white race, radiation therapy, and male sex, as shown in Fig. 2, dotted lines.

Multivariate Analysis with the Fine and Gray Model

On the basis of the consequences of Fine and Gray’s proportional subdistribution hazard model in the univariate and multivariate analysis, location, AJCC, AJCC.T, AJCC.N, radiotherapy, and LONT.cat could significantly and independently predict cancer-specific mortality in patients with gastric adenocarcinoma.

Cardia patients had a better prognosis (antrum/pylorus versus cardia: 2.779, 95% confidence interval [CI] 2.075–3.721, $p < 0.001$; middle versus cardia: 2.885, 95%

CI 2.166–3.843, $p < 0.001$; others versus cardia: 3.077, 95% CI 2.197–4.31, $p < 0.001$).

AJCC III patients had a worse prognosis except II (I versus III: 0.303, 95% CI 0.147–0.625, $p = 0.001$; II versus III: 0.804, 95% CI 0.571–1.132, $p = 0.210$).

Compared with T4b, other T statuses were associated with a better prognosis except T4a (T1 versus T4b: 0.572, 95% CI 0.289–1.134, $p = 0.11$; T2 versus T4b: 0.529, 95% CI 0.314–0.893, $p = 0.017$; T3 versus T4b: 0.648, 95% CI 0.45–0.934, $p = 0.02$, T4a versus T4b: 0.929, 95% CI 0.652–1.323, $p = 0.68$).

Compared with N3, other N statuses were associated with a better prognosis (N0 versus N3: 0.346, 95% CI 0.227–0.526, $p < 0.001$; N1 versus N3: 0.533, 95% CI 0.404–0.703, $p < 0.001$; N2 versus N3: 0.552, 95% CI 0.434–0.702, $p < 0.001$).

Patients who did not receive radiotherapy were more likely to die of GC (with versus without radiotherapy: 0.684, 95% CI 0.533–0.702, $p < 0.001$).

LONT.cat was a statistically significant independent predictor of patient outcome (high versus low: 0.705, 95% CI 0.524–0.95, $p = 0.021$), as shown in Fig. 3a.

Construction of the Competitive Risk Model

The selection of variables for the competitive risk model in the training set was based on the optimal subset method (the minimum BIC), and the variables included were as follows: location, AJCC, AJCC.T, AJCC.N, radiotherapy, LONT.cat, and chemotherapy. A competing risk nomogram based on the minimum BIC was constructed to forecast the 12-month, 36-month, and 60-month cumulative cause-specific death probabilities. The coefficients from the competing risk models fit very well (Supplementary Table 3). A total of 7 independent prognostic variables were incorporated into the graph. The LONT accounted for the largest proportion of risk scores. The nomogram is shown in Fig. 3b.

Evaluation of Competitive Risk Models

To evaluate the competitive risk model in the training set, two models were constructed and compared. The first model was composed of all factors, and the second model was composed of factors screened according to the BIC criterion.

The time-dependent ROC curve, C-index, and comprehensive evaluation index (time-AUC, C-index and Brier score, respectively) were used to compare multiple indicators: each individual indicator vs. the BIC criterion screening model vs. the all-factor model. In the model based on the BIC criterion, the time-AUCs at different time points (12 months, 36 months, and 60 months) were 0.811, 0.836,

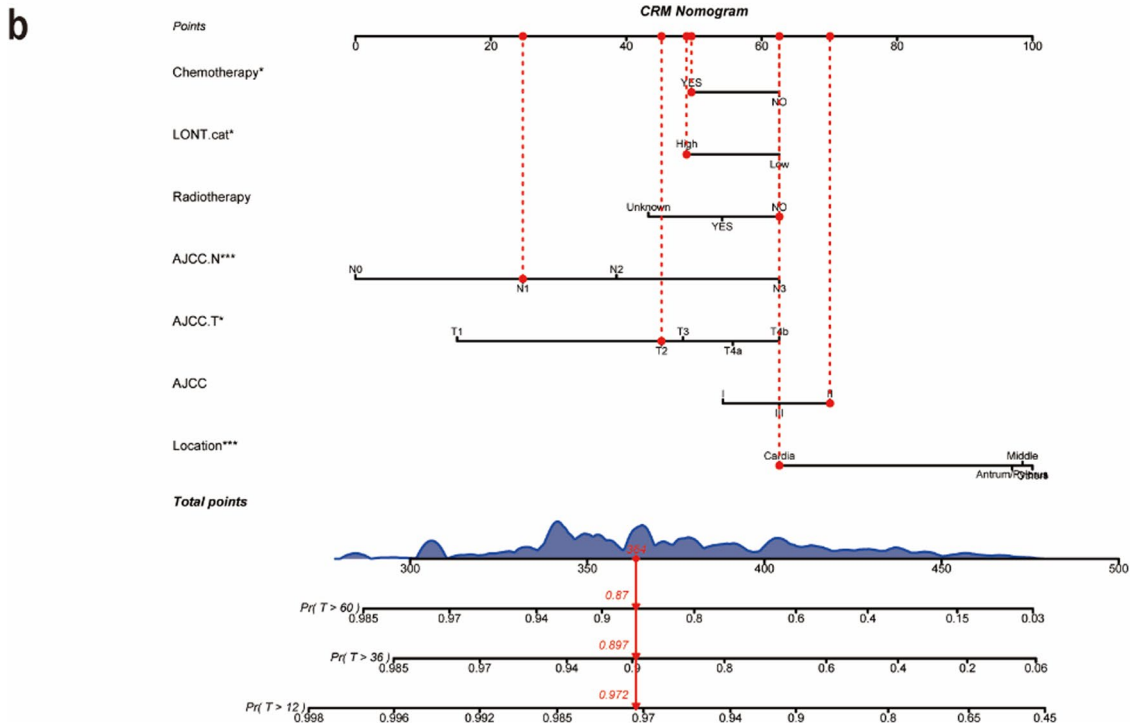
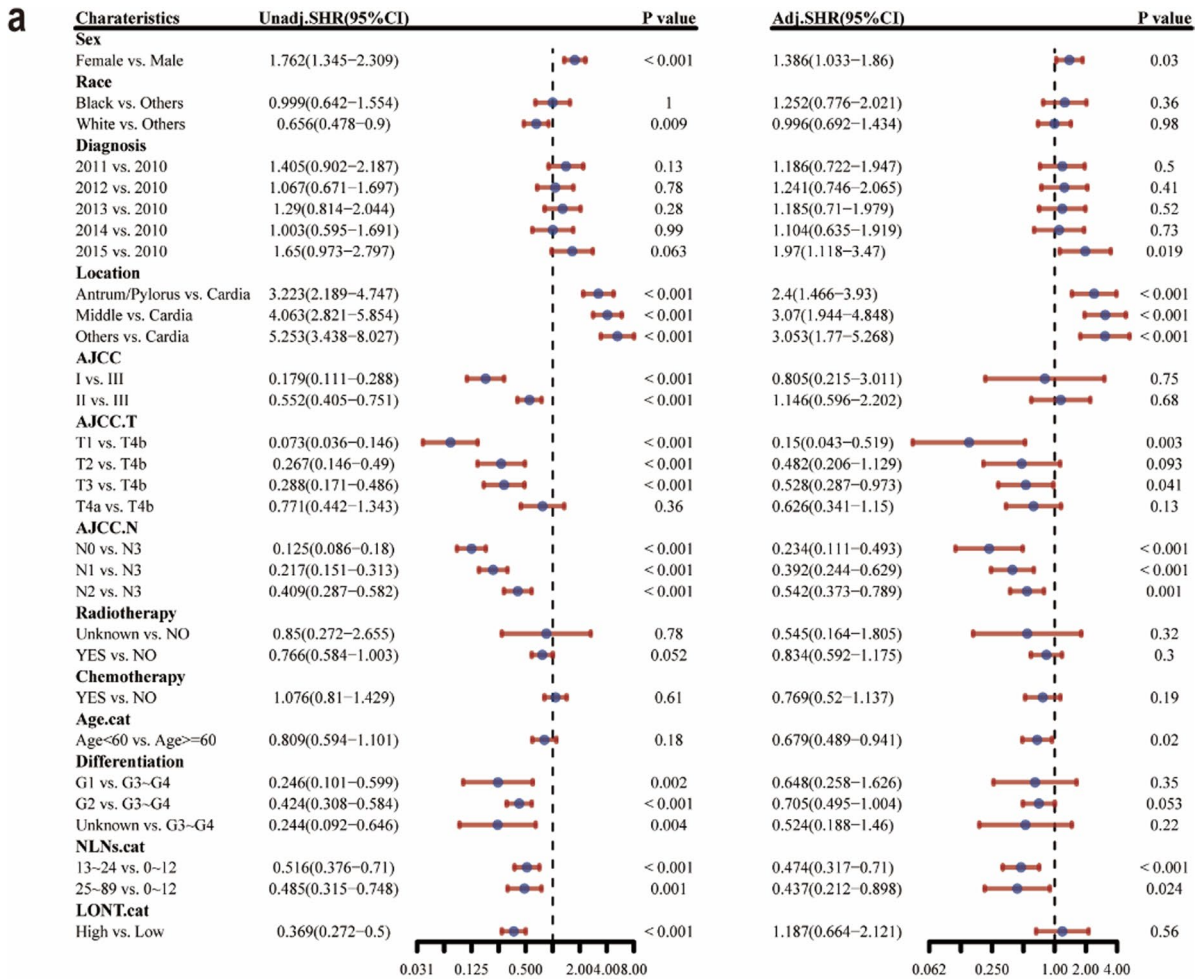


Fig. 5 a Univariate and multivariate analyses evaluating the relationship between the LONT and cancer-specific mortality according to different clinicopathological factors in the validation cohort. **b** The competing risk nomogram based on the minimum BIC to forecast the 12-month, 36-month, and 60-month cumulative cause-specific death probabilities in the validation cohort

and 0.832, respectively; the C-indexes were 0.812, 0.821, and 0.818, respectively; and the Brier scores were 0.067, 0.122, and 0.135, respectively. In the model composed of all factors, the time-AUCs at different time nodes (12-month, 36-month, and 60-month) were 0.819, 0.845, and

0.842, respectively; the C-indexes were 0.818, 0.821, and 0.814, respectively; and the Brier scores were 0.067, 0.123, and 0.138, respectively (Supplementary Tables 4 and 5).

Calibration plots revealed high consistency between the predicted and actual observed 12-, 36-, and 60-month cause-specific death probabilities for GC patients in the all-factor model in Fig. 4a and in the BIC model in Fig. 4b in the training cohort.

We applied DCA to evaluate the benefits of the model. DCA and calibration curves were plotted for 12, 36, and 60 months for the all-factor model (Fig. 4c) and for the BIC model (Fig. 4d) in the training cohort.

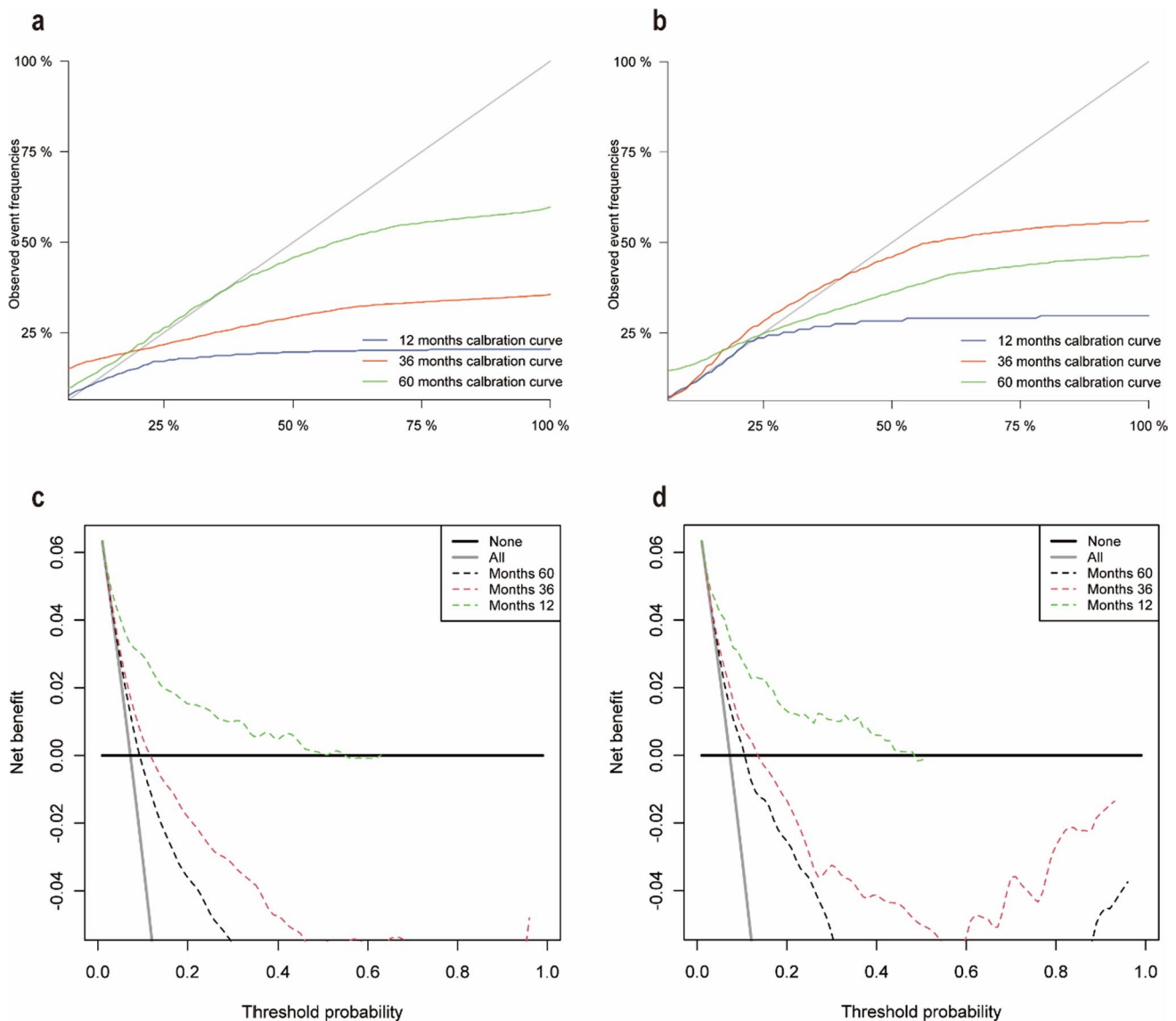


Fig. 6 Performance of the competing risk nomogram model. Calibration plots of the predicted and actual observed 12-, 36-, and 60-month cause-specific death probabilities in the all-factor model (a) and in the BIC model (b) in the validation cohort. Decision curve analysis

(DCA) of the predicted and actual observed 12-, 36-, and 60-month cause-specific death probabilities in the all-factor model (c) and in the BIC model (d) in the validation cohort

In the training set, scores based on the prediction model were divided into high-/medium-/low-risk groups by using the tertile method. A cumulative risk curve was used to evaluate the effectiveness of the prediction model and draw the cumulative risk curve of the score based on the prediction model. For cancer-specific mortality, the higher the score was, the higher the cumulative incidence, and the difference was statistically significant ($p < 0.01$), confirming that the nomogram risk score had excellent survival prediction ability for cancer-specific mortality (Fig. 4e) in the training cohort.

Model Validation

Univariate analysis and multivariate analysis of the competitive risk model were carried out in the validation set. Sex, age, location, AJCC.T, AJCC.N, and NLNs.cat could significantly and independently predict cancer-specific mortality in patients with gastric adenocarcinoma, as shown in Fig. 5a.

Variables were screened for inclusion according to the minimum BIC; the validation of the nomogram is shown in Fig. 5b.

To evaluate the competitive risk model in the validation set, two models were constructed and compared. The first model was composed of all factors, and the second model was composed of factors screened according to the BIC criterion.

In the model based on the BIC criterion, the time-AUCs at different time nodes (12-month, 36-month, and 60-month) were 0.819, 0.821, and 0.829, respectively; the C-indexes were 0.813, 0.812, and 0.817, respectively; and the Brier scores were 0.057, 0.118, and 0.126, respectively. In the model composed of all factors, the time-AUCs at different time nodes (12-month, 36-month, and 60-month) were 0.843, 0.845, and 0.845, respectively; the C-indexes were 0.836, 0.835, and 0.833, respectively; and the Brier scores were 0.056, 0.111, and 0.119, respectively (Supplementary Tables 6 and 7).

The calibration curves presented good consistency between the nomogram-predicted and actual observed 12-month, 36-month, and 60-month cancer-specific mortality values in the all-factor model (Fig. 6a) and in the BIC model (Fig. 6b) in the validation cohort.

We also applied DCA for the 12-month, 36-month, and 60-month data to evaluate the benefits of the two models in all-factor mode (Fig. 6c) and in the BIC model (Fig. 6d) in the validation cohort.

The cumulative risk curve of the validation set predicted scores was consistent with that of the training set predicted scores (Fig. 4f).

Discussion

In 2018, there were 1,033,000 cases of gastric cancer worldwide, including 783,000 deaths (2). Timely and effective assessment of clinical presentations and prognostic factors can help oncologists make early diagnosis and treatment decisions, but these aspects have not been fully explored. In this study, we screened gastric cancer patients from the SEER database to explore their clinical features. We included a new variable, the LONT, which not only represented the extent of lymph node dissection (LND) but also reflected the severity of the disease.¹²

It is well known that for resectable gastric cancer, prognostic factors include TNM staging,¹³ biological characteristics of the tumor (8,10), the extent of LND,^{14–16} and postoperative comprehensive treatment (1). Only ELNs and NLNs reflect the degree of lymph node dissection in radical gastrectomy for gastric cancer. In a previous study,¹⁷ the modified lymph node ratio (mLNR) was defined as $pLN + 0.5 / tLN + 0.5$, where pLN is the number of positive LNs and tLN is the total number of obtained LNs. That study established a novel nomogram based on the mLNR in 3339 breast cancer patients. In this study, the LONT was defined as the log of the ratio between the NLN counts plus one and the T stage; NLNs reflect the total level of LND, with T representing the severity of the disease. The NLN adjusted by T staging can be used to compare the relative levels of LND in different patients. One study showed that a high LONT was associated with improved survival in patients with gastric cancer and was an independent predictor of prognosis. Furthermore, the prognostic nomogram model based on the LONT can effectively predict DSS and OS in patients with resectable GC.¹²

A growing body of research suggests that nomograms may be an accurate, usable, and convenient tool for predicting patient outcomes using simple graphical representations.^{7,12,17} Some previous studies have found that variables related to long-term survival include age, tumor site, depth of invasion, and lymph node metastasis, and based on them, a predictive nomogram was established, which showed good predictive efficacy.^{18–21} However, previously published studies of nomograms included all prognostic variables.^{17,21} In our study, the selection of variables for the competitive risk model was based on the minimum BIC, considering the number of samples. When the number of samples was too large, model complexity resulting in precision that was too high could be effectively prevented, so the samples were filtered by location, AJCC, and AJCC.T, AJCC.N, radiotherapy, LONT.cat, and chemotherapy, and the results were more reliable. In addition, the current study built a nomogram for prognosis assessment that included the LONT. The results of the time-AUC and C-index analyses indicated that the predicted values of our nomograms had excellent

discriminant ability. Calibration plots revealed high consistency between the predicted and actual observed 12-, 36-, and 60-month cause-specific death probabilities. DCA proved that our nomogram provided great benefit. The Brier score reflected that the predicted values of our nomogram exhibited better prognostic performance. All the above results were verified in the validation set. All the above evaluation systems showed the high reliability of this study model, which could be used to guide individualized postoperative treatment strategies and follow-up plans.

Our study has some advantages. First, the study used the SEER database, and 3895 patients with gastric cancer were identified from among 183,950 patients. Second, few studies have quantified the relative extent of LND, and none of the studies applied competitive risk models. This study established a novel nomogram to predict the survival of gastric cancer patients after gastrectomy based on the LONT and clinicopathological factors.

The study still has some limitations. First, some important prognostic data from the SEER database were missing, such as the resection method, CEA, and CA19-9. Third, the histological type and genotyping of GC are also important biological characteristics affecting prognosis, but they are not included in our adjustment variable. In addition, some risk indicators, such as significant genetic or epigenetic characteristics verified by wet experiments, will be included in the comprehensive analysis and drawn into more rigorous nomogram. In addition, because the time span of data we mined was large, and SEER database cannot provide relevant data effectively. We could not mine these data about the types of chemotherapy and radiotherapy and the indication criteria from the SEER database, so all chemotherapy and radiation therapies, regardless of timing or regimen, were included in the nomogram. At last, this study is a retrospective study, and its argumentation level is lower than that of a prospective study.

Conclusion

In this study, we screened gastric cancer patients from the SEER database to explore their clinical features. We included a new variable, the LONT, which not only represented the extent of LND but also reflected the severity of the disease. As an independent clinicopathological feature, a low LONT was accompanied by a high CIF for GC-specific death. The competing risk nomogram based on the LONT exhibited better prognostic performance in predicting the 12-month, 36-month, and 60-month cumulative cause-specific death probabilities. This is a simple, accurate, and convenient way to predict cause-specific death, thus providing a basis for clinicians to formulate individual treatment strategies.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11605-022-05408-8>.

Acknowledgements The authors would like to thank Fujiao Duan, Ph.D. Department of Epidemiology and Biostatistics, College of Public Health, Zhengzhou University, for his guidance on the statistical methods used in this study.

Author Contribution Wei Yang designed the ideas for this paper. Wei Yang, Shuaibing Lu, and Fusheng Ge contributed to the writing of the manuscript. Wei Yang, Fusheng Ge, and Yawei Hua contributed to data collation and data analysis. Wei Yang, Shuaibing Lu, and Xiaobing Chen analyzed and interpreted the data.

All the authors contributed to this article and approved the submitted version.

All the authors critically revised the manuscripts. All the authors gave the final approval of the version to be published.

Funding This study was supported by the National Natural Science Foundation of China (No. 81472714) and the Medical Science and Technique Foundation of Henan Province (No. SB201901101).

Declarations

Informed Consent This was a retrospective study, and the public data did not involve patient-specific data, so according to Swedish law, this study was exempt from the need to obtain signed informed consent.

Conflict of Interest The authors declare no competing interests.

References

1. Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. *Lancet* 2020;396:635–48. [https://doi.org/10.1016/s0140-6736\(20\)31288-5](https://doi.org/10.1016/s0140-6736(20)31288-5)
2. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca Cancer J Clin* 2018 11;68(6):394–424. <https://doi.org/10.3322/caac.21492>
3. Smith DD, Schwarz RR, Schwarz RE. Impact of total lymph node count on staging and survival after gastrectomy. *J Clin Oncol* 2005 Oct 01;23(28):7114–24. <https://doi.org/10.1200/JCO.2005.14.621>
4. Son T, Hyung WJ, Lee JH, et al. Clinical implication of an insufficient number of examined lymph nodes after curative resection for gastric cancer. *Cancer* 2012 Oct 01;118(19):4687–93. <https://doi.org/10.1002/cncr.27426>
5. Kattan MW, Karpeh MS, Mazumdar M et al. Postoperative nomogram for disease-specific survival after an R0 resection for gastric carcinoma. *J Clin Oncol* 2003 Oct 01;21(19):3647–50. <https://doi.org/10.1200/JCO.2003.01.240>
6. Martinez-Ramos D, Calero A, Escrig-Sos J, et al. Prognosis for gastric carcinomas with an insufficient number of examined negative lymph nodes. *European journal of surgical oncology* 2014 Mar;40(3):358–65. <https://doi.org/10.1016/j.ejso.2013.08.027>
7. Wang ZX, Qiu MZ, Jiang YM, et al. Comparison of prognostic nomograms based on different nodal staging systems. *Journal of Cancer* 2017;8(6):950–958. <https://doi.org/10.7150/jca.17370>
8. Mao A, Zhou X, Liu YX, et al. KLF8 is associated with poor prognosis and regulates glycolysis by targeting GLUT4 in gastric cancer. *Journal of cellular and molecular medicine* 2019 08;23(8):5087–5097. <https://doi.org/10.1111/jcmm.14378>
9. Sun L, Chen G, Sun A, et al. BAG2 Promotes Proliferation and Metastasis of Gastric Cancer. *Frontiers in oncology* 2020;10:31. <https://doi.org/10.3389/fonc.2020.00031>

10. Wang WJ, Guo CA, Li R, et al. Long non-coding RNA CASC19 is associated with the progression and prognosis of advanced gastric cancer. *Aging* 2019 08 15;11(15):5829–5847. <https://doi.org/10.18632/aging.102190>
11. Zhu Y, Fang X, Wang L, et al. A Predictive Nomogram for Early Death of Metastatic Gastric Cancer: A Retrospective Study in the SEER Database and China. *Journal of Cancer* 2020;11(18):5527–5535. <https://doi.org/10.7150/jca.46563>
12. Xie J, Pang Y, Li X, et al. The log odds of negative lymph nodes/T stage. *Journal of cancer research and clinical oncology* 2021 Aug;147(8):2259–2269. <https://doi.org/10.1007/s00432-021-03654-y>
13. In H, Solsky I, Palis B, et al. Validation of the 8th Edition of the AJCC TNM Staging System for Gastric Cancer. *Annals of surgical oncology* 2017 Nov;24(12):3683–3691. <https://doi.org/10.1245/s10434-017-6078-x>
14. Degiuli M, Sasako M, Ponti A, et al. Survival results of a multicentre phase II study. *British journal of cancer* 2004 May 04;90(9):1727–32. <https://doi.org/10.1038/sj.bjc.6601761>
15. Enzinger PC, Benedetti JK, Meyerhardt JA, et al. Impact of hospital volume on recurrence and survival. *Annals of surgery* 2007 Mar;245(3):426–34. <https://doi.org/10.1097/01.sla.0000245469.35088.42>
16. Schwarz RE, Smith DD et al. Clinical impact of lymphadenectomy extent. *Annals of surgical oncology* 2007 Feb;14(2):317–28. <https://doi.org/10.1245/s10434-006-9218-2>
17. Wen J, Yang Y, Liu P, et al. Development and validation of a nomogram for predicting survival in breast cancer patients. *Breast* 2017 Jun;33:14–22. <https://doi.org/10.1016/j.breast.2017.01.017>
18. Strong VE, Song KY, Park CH, et al. Comparison of gastric cancer survival following R0 resection. *Annals of surgery* 2010 Apr;251(4):640–6. <https://doi.org/10.1097/SLA.0b013e3181d3d29b>
19. Marrelli D, Stefano AD, Manzoni GD, et al. Prediction of recurrence after radical surgery. *Annals of surgery* 2005 Feb;241(2):247–55. <https://doi.org/10.1097/01.sla.0000152019.14741.97>
20. Posteraro B, Persiani R, Dall'Armi V, et al. Prognostic factors and outcomes in Italian patients. *European journal of surgical oncology* 2014 Mar;40(3):345–51. <https://doi.org/10.1016/j.ejso.2013.11.002>
21. Kim Y, Spolverato G, Ejaz A, et al. A nomogram to predict overall survival. *Annals of surgical oncology* 2015;22(6):1828–35. <https://doi.org/10.1245/s10434-014-4230-4>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.