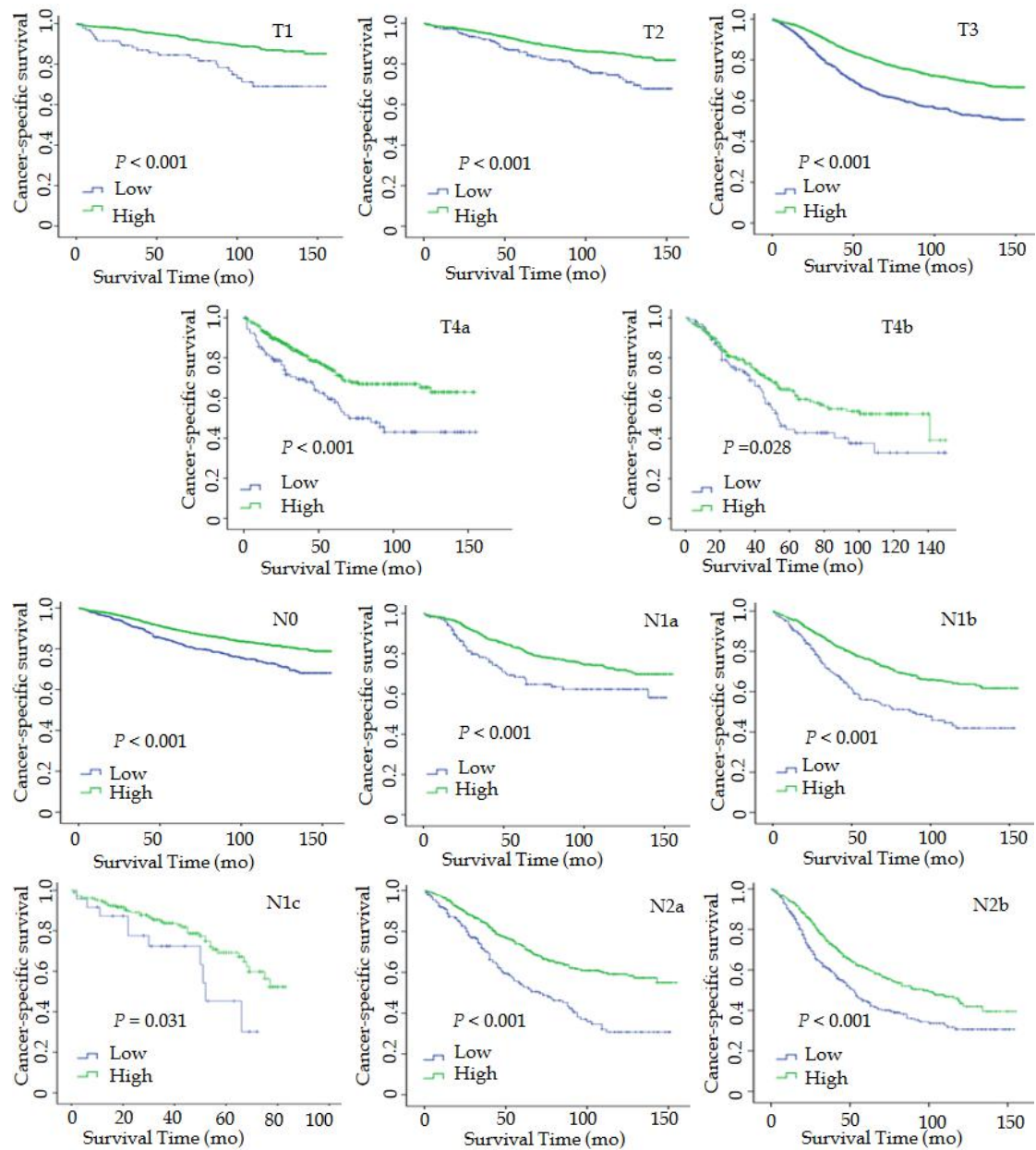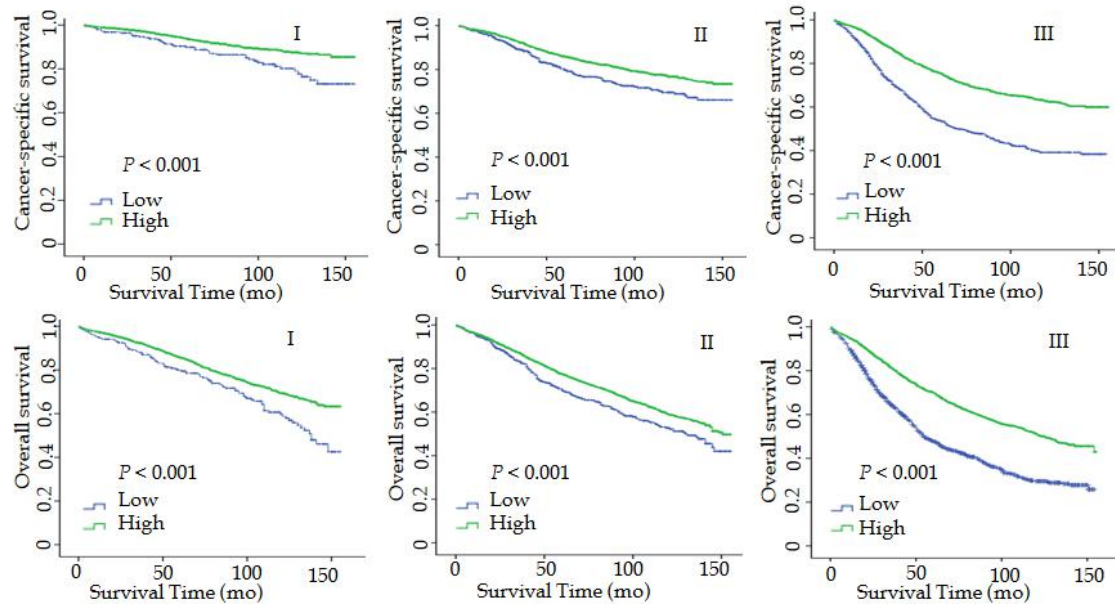**Supplementary Figure 1 Kaplan-Meier survival curves of T stage and N stage subgroup for overall survival in the training cohort.**

Supplementary Figure 2 Kaplan-Meier survival curves of T stage and N stage subgroup for cancer-specific survival in the validation cohort.

Supplementary Figure 3 Kaplan-Meier survival curves of tumor-node-metastasis stage subgroup for cancer-specific survival and overall survival in the validation cohort.

Supplementary Table 1 Univariate analysis of prognostic factors for overall survival

Variable	Training cohort (n = 371)			Validation cohort (n = 12325)		
	HR	95%CI	*P* value	HR	95%CI	*P* value
Age[1]	1.043	1.041-1.045	< 0.001	1.042	1.039-1.045	< 0.001
Tumor size[1]	1.003	1.003-1.004	< 0.001	1.003	1.002-1.003	< 0.001
NLNs/tumor size[1]	0.978	0.976-0.981	< 0.001	0.970	0.971-0.979	< 0.001
LONS[1]	0.423	0.400-0.447	< 0.001	0.420	0.386-0.456	< 0.001
Sex (ref = male)	0.876	0.839-0.915	< 0.001	0.866	0.811-0.925	< 0.001
TNM (ref = stage III)						
I	0.452	0.426-0.478	< 0.001	0.45	0.41-0.49	< 0.001
II	0.632	0.603-0.663	< 0.001	0.66	0.62-0.71	< 0.001
Differentiation (ref =G3 + G4)						
G1	0.604	0.547-0.666	< 0.001	0.604	0.521-0.699	< 0.001
G2	0.690	0.652-0.729	< 0.001	0.697	0.624-0.740	< 0.001
Unknown	0.652	0.582-0.732	< 0.001	0.591	0.497-0.705	< 0.001
Histology (ref = MAC and SRCC)	0.624	0.577-0.675	< 0.001	0.639	0.567-0.720	< 0.001
Preoperative RT (ref = Yes)	1.081	1.033-1.130	0.001	1.129	1.058-1.204	< 0.001
Chemotherapy (ref = Yes)	1.119	1.072-1.167	< 0.001	0.776	0.712-0.815	< 0.001
Location (ref = rectum)	0.946	0.906-0.989	0.014	0.931	0.870-0.996	0.038
Race (ref = others)						
white	1.114	1.063-1.230	< 0.001	1.154	1031-1.293	0.013
balck	1.460	1.326-1.608	< 0.001	1.537	1.326-1.781	< 0.001

CEA (ref = High)	0.605	0.573-0.639	< 0.001	0.553	0.509-0.600	< 0.001
Normal	0.781	0.742-0.822	< 0.001	0.741	0.686-0.801	< 0.001
Unknown						

[1]These variables were treated as continuous data. CI: Confidence interval; CEA: Carcinoembryonic antigen; HR: Hazard ratio; LONS: Log odds of negative lymph nodes/tumor size; MAC: Mucinous adenocarcinoma; NLNs: Negative lymph nodes; ref: Reference; TNM: Tumor-node-metastasis; RT: Radiotherapy; SRCC: Signet-ring cell carcinoma.

Supplementary Table 2 Multivariate analysis of prognostic factors overall survival

Variable	Training cohort (*n* = 28755)			Validation cohort (*n* = 12325)		
	HR	95%CI	P^2 value	HR	95%CI	P^2 value
Age (ref: ≤ 60 years old)	2.11	2.01-2.21	< 0.001	2.04	1.90-2.19	< 0.001
Tumor size[1]	1.00	1.00-1.00	0.999	1.00	1.00-1.00	0.867
LONS[1]	0.53	0.49-0.56	< 0.001	0.52	0.42-0.52	< 0.001
Sex (ref = male)	0.87	0.83-0.91	< 0.001	0.86	0.80-0.91	< 0.001
TNM1 (ref = stage III)						
I	0.40	0.37-0.43	< 0.001	0.42	0.38-0.46	< 0.001
II	0.59	0.56-0.62	< 0.001	0.64	0.59-0.69	< 0.001
Differentiation (ref = G3 + G4)						
G1	0.76	0.69-0.84	< 0.001	0.74	0.64-0.86	< 0.001
G2	0.79	0.75-0.84	< 0.001	0.80	0.73-0.87	< 0.001
Unknown	0.76	0.68-0.86	< 0.001	0.73	0.61-0.87	< 0.001
Histology (ref =MAC and SRCC)	0.80	0.74-0.87	< 0.001	0.78	0.69-0.86	< 0.001
Preoperative RT (ref = Yes)	0.90	0.85-0.96	< 0.001	0.96	0.88-1.05	0.411
Chemotherapy (ref = Yes)	1.67	1.58-1.77	< 0.001	1.60	1.47-1.74	< 0.001
Location (ref = rectum)	0.84	0.80-0.88	< 0.001	0.81	0.75-0.81	< 0.001
Race (ref = others)						
white	1.18	1.10-1.27	< 0.001	1.17	1.04-1.31	0.007
balck	1.51	1.37-1.66	< 0.001	1.58	1.37-1.83	< 0.001
CEA (ref = High)						
Normal	0.68	0.64-0.72	< 0.001	0.63	0.57-0.68	< 0.001
Unknown	0.83	0.78-0.87	< 0.001	0.78	0.72-0.84	< 0.001

¹These variables were treated as continuous data; [2]*P* values were adjusted for sex, age, race, tumor-node-metastasis stage, location, preoperative radiotherapy, carcinoembryonic antigen, differentiation and chemotherapy as covariates. HR: Hazard ratio; CI: Confidence interval; CEA: Carcinoembryonic antigen; LONS: Log odds of negative lymph nodes/tumor size; TNM: Tumor-node-metastasis; MAC: Mucinous adenocarcinoma; SRCC: Signet-ring cell carcinoma; RT: Radiotherapy.

Supplementary Table 3 Multivariate analyses for evaluating log odds of negative lymph nodes/tumor size effect on cancer-specific survival based on different clinicopathological factors[1]

Variable	Training cohort (*n* = 28755)			Validation cohort (*n* = 12325)		
	HR	95%CI	P^2 value	HR	95%CI	P^2 value
All	1.579	1.47-1.63	< 0.001	1.533	1.42-1.66	< 0.001
N stage						
N0	1.42	1.32-1.54	< 0.001	1.336	1.18-1.51	< 0.001
N1a	1.43	1.23-1.66	< 0.001	1.34	1.03-1.76	0.03
N1b	1.31	1.15-1.50	< 0.001	1.45	1.18-1.78	< 0.001
N1c	1.72	1.08-2.72	0.22	2.01	1.00-4.23	0.049
N2a	1.7	1.47-1.95	< 0.001	1.7	1.37-2.11	< 0.001
N2b	1.47	1.29-1.68	< 0.001	1.39	1.14-1.70	0.001
T stage						
T1	1.59	1.40-1.80	< 0.001	1.31	0.88-1.95	0.19
T2	1.48	1.29-1.69	< 0.001	1.51	1.22-1.86	< 0.001
T3	1.55	1.46-1.65	< 0.001	1.67	1.56-1.80	< 0.001
T4a	1.41	1.13-1.75	0.002	1.47	1.05-2.06	0.23
T4b	1.24	1.01-1.50	0.038	1.44	1.02-2.02	0.037
TNM stage						
I	1.59	1.48-1.71	< 0.001	1.483	1.19-1.86	0.001
II	1.42	1.30-1.56	< 0.001	1.27	1.11-1.47	0.001
III	1.63	1.52-1.74	< 0.001	1.64	1.48-1.82	< 0.001
Sex						
Male	1.583	1.49-1.69	< 0.001	1.504	1.36-1.66	< 0.001
Female	1.492	1.37-1.62	< 0.001	1.587	1.40-1.80	< 0.001
Age years						
≤ 60	1.802	1.65-1.97	< 0.001	1.867	1.63-2.14	< 0.001

	1			2		
> 60	1.447	1.36-1.54	< 0.001	1.408	1.28-1.55	< 0.001
Differentiation						
G1	1.349	1.08-1.69	0.008	1.422	1.05-1.92	0.022
G2	1.515	1.43-1.61	< 0.001	1.460	1.33-1.61	< 0.001
G3	1.8	1.61-2.0	< 0.001	1.784	1.51-2.11	< 0.001
Location						
Rectum	1.518	1.43-1.61	< 0.001	1.46	1.33-1.60	< 0.001
Rectosigmoid Junction	1.617	1.48-1.77	< 0.001	1.737	1.51-1.99	< 0.001
CEA						
High	1.584	1.45-1.74	< 0.001	1.6	1.40-1.83	< 0.001
Normal	1.616	1.47-1.77	< 0.001	1.429	1.23-1.66	< 0.001
Preoperative RT						
Yes	1.489	1.37-1.62	< 0.001	1.317	1.15-1.50	< 0.001
No	1.588	1.49-1.69	< 0.001	1.665	1.51-1.83	< 0.001
Chemotherapy						
Yes	1.582	1.49-1.68	< 0.001	1.488	1.35-1.64	< 0.001
No	1.48	1.36-1.61	< 0.001	1.591	1.40-1.80	< 0.001
ELN						
0-10	1.38	1.28-1.49	< 0.001	1.31	1.16-1.47	< 0.001
11-20	1.61	1.46-1.77	< 0.001	1.857	1.61-2.15	< 0.001
> 20	2.78	2.27-3.45	< 0.001	2.493	1.81-3.44	< 0.001

[1]Log odds of negative lymph nodes/tumor size was treated as categorical variable, and the high group as reference; [2]P values were adjusted for sex, age, race, tumor-node-metastasis stage, location, preoperative radiotherapy, carcinoembryonic antigen, differentiation and chemotherapy as covariates. CI:

Confidence interval; CEA: Carcinoembryonic antigen; HR: Hazard ratio; antigen; RT: radiotherapy; ELN: Eexamined lymph nodes; TNM: Tumor-node-metastasis.