

Table S1 Clinical features of the included early-stage NSCLC patients

Variables	Total (n=22403)	Training (n=14935)	Test (n=7468)	P value
Age (mean, SD)	66.7±10.3	66.7±10.3	66.6±10.3	0.36
Gender				0.81
Female	12,221	8,156	4,065	
Male	10,182	6,779	3,403	
Race				0.15
White	18,574	12,433	6,141	
Black	2,004	1,316	688	
Unknown	1,825	1,186	639	
Primary site				0.63
RUL	7,517	4,979	2,538	
RML	1,357	932	425	
RLL	3,978	2,644	1,334	
LUL	5,747	3,843	1,904	
LLL	3,303	2,199	1,104	
Unknown	501	338	163	
Pathological differentiation				0.62
Well; I	4,786	3,174	1612	
Moderately; II	10,173	6,827	3346	
Poorly; III	7,142	4,731	2411	
Undifferentiated; IV	302	203	99	
pT stage				0.53
T1	12,935	8,601	4,334	
T2a	9,468	6,334	3,134	
Lymph node status*			5,983	0.47
N0	17,934	11,915	804	
N1	2,343	1,539	640	
N2	2,007	42	25	
N3	67	36	16	
NX	52	36	16	
Metastasis status				0.25
M0	21,810	14,529	7,281	
M1a	201	145	56	
M1b	392	261	131	
TNM stage				0.47
I	17,600	11,722	5,878	
II	2,262	1,485	777	
III	1,905	1,291	614	
IV	593	406	187	
Unknown	43	31	12	

Table S1 (continued)

Table S1 (continued)

Variables	Total (n=22403)	Training (n=14935)	Test (n=7468)	P value
Tumor size				0.29
≤30mm	16,576	10,877	5,499	
30-50 mm	5,961	4,017	1,944	
Unknown	66	41	25	
Histological type				0.79
Adenocarcinoma	13,471	8,972	4,499	
Squamous	5,376	3,569	1,807	
Large cell	398	265	133	
NSCLC (NOS)	3,158	2,129	1,029	
Pleural invasion				0.91
No invasion	15,480	10,322	5158	
Invasion	4,181	2,778	1403	
Unknown	2,742	1,835	907	
Surgery type				0.86
Wedge resection	4,338	2,891	1,447	
Lobectomy/Bilobectomy	17,560	11,713	5,847	
Pneumonectomy	505	331	174	
Distant metastasis				1
Negative	22,011	14,674	7,337	
Positive	392	261	131	
Bone metastasis				0.65
Negative	22,270	14,844	7,426	
Positive	95	63	32	
Unknown	38	28	10	
Brain metastasis				0.5
Negative	22,154	14,770	7,384	
Positive	210	136	74	
Unknown	39	29	10	
Liver metastasis				0.31
Negative	22,334	14,883	7,451	
Positive	29	22	7	
Unknown	40	30	10	
Lung metastasis				0.35
Negative	22,282	14,847	7,435	
Positive	75	54	21	
Unknown	46	34	12	
Overall survival				0.41
Alive	18,257	12,148	6,109	
Death	4,146	2,787	1,359	

*, Nx means that there is lymph node metastasis, but unknown for which stage. SD, standard deviation; RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL, left upper lobe; LLL, left lower lobe; pT, pathologic T; NSCLC (NOS), Non-small cell lung cancer, Not otherwise specified.

Table S2 Point distributions and predictive scores for every parameter in both nomograms

Variables	Nomogram score	
	LNM	Distant metastasis
Age		
≤65	12	23
>65	0	0
Gender		
Female	0	0
Male	5	13
Primary site		
RUL	0	–
RML	13	–
RLL	13	–
LUL	14	–
LLL	15	–
Pathological differentiation		
Well; I	0	0
Moderately; II	51	40
Poorly; III	64	59
Undifferentiated; IV	67	70
pT stage		
T1	0	–
T2a	25	–
Lymph node status		
N0	–	0
N1	–	25
N2	–	42
N3	–	100

Table S2 (continued)

Table S2 (continued)

Variables	Nomogram score	
	LNM	Distant metastasis
Tumor size		
≤30mm	0	–
30-50 mm	16	–
Histological type		
Adenocarcinoma	23	30
Squamous	0	0
Large cell	9	42
NSCLC (NOS)	14	26
Pleural invasion		
No invasion	–	0
Invasion	–	22
Surgery type		
Wedge resection	0	39
Lobectomy/bilobectomy	28	1
Pneumonectomy	100	0
Distant metastasis		
Negative	0	–
Positive	52	–

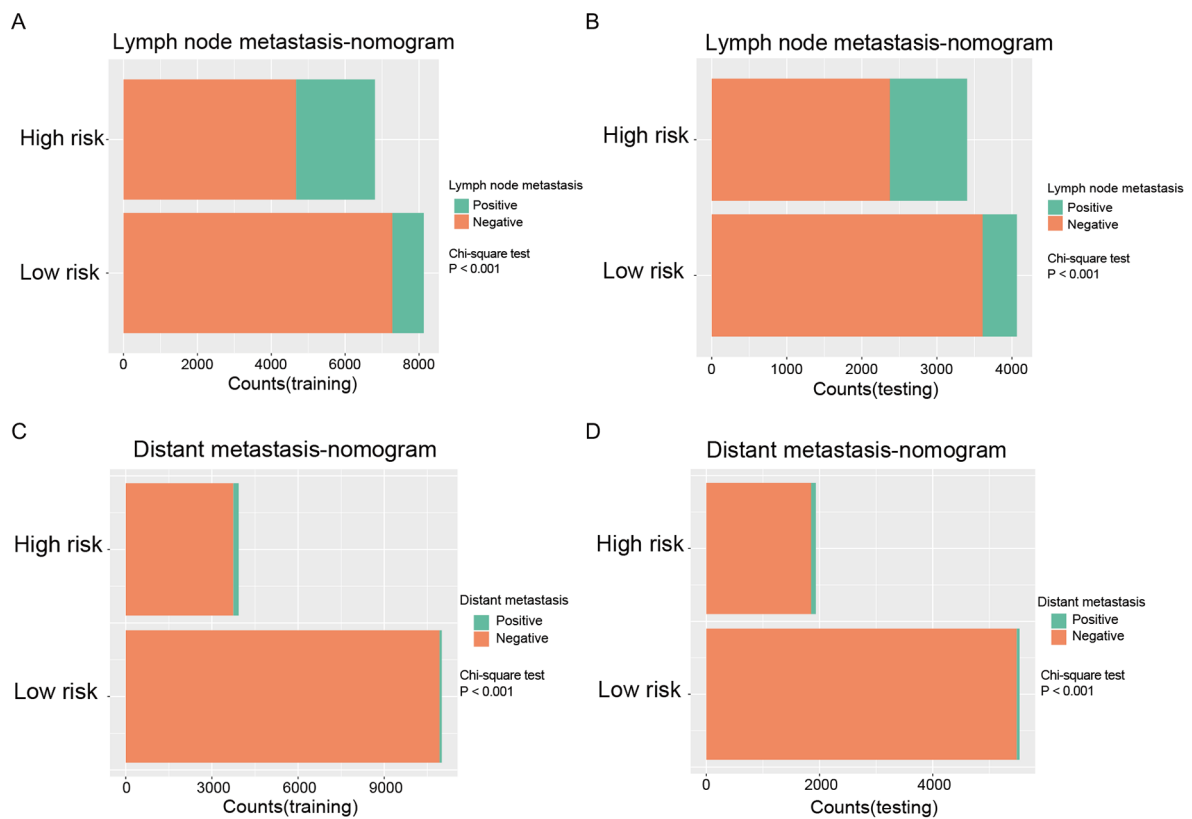


Figure S1 Discriminatory efficiency of the predictive nomogram models for LNM and distant metastasis. Risk stratification for the nomogram models was performed through the maximum Youden index, and their performance in differentiating LNM and distant metastasis in (A,C) the training cohort and (B,D) the testing cohort were plotted.

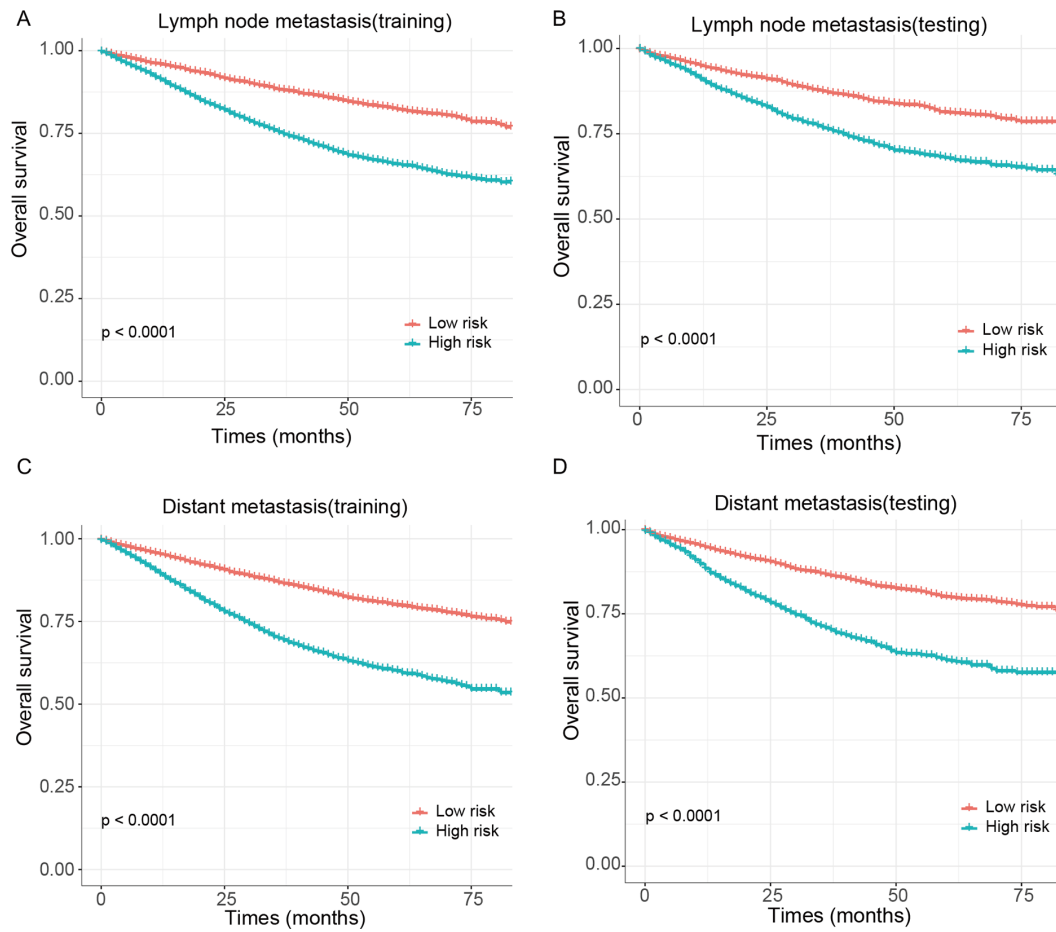


Figure S2 Survival analysis of patients with LNM and distant metastasis in (A,C) the training cohort and (B,D) the testing cohort, respectively. OS, overall survival.