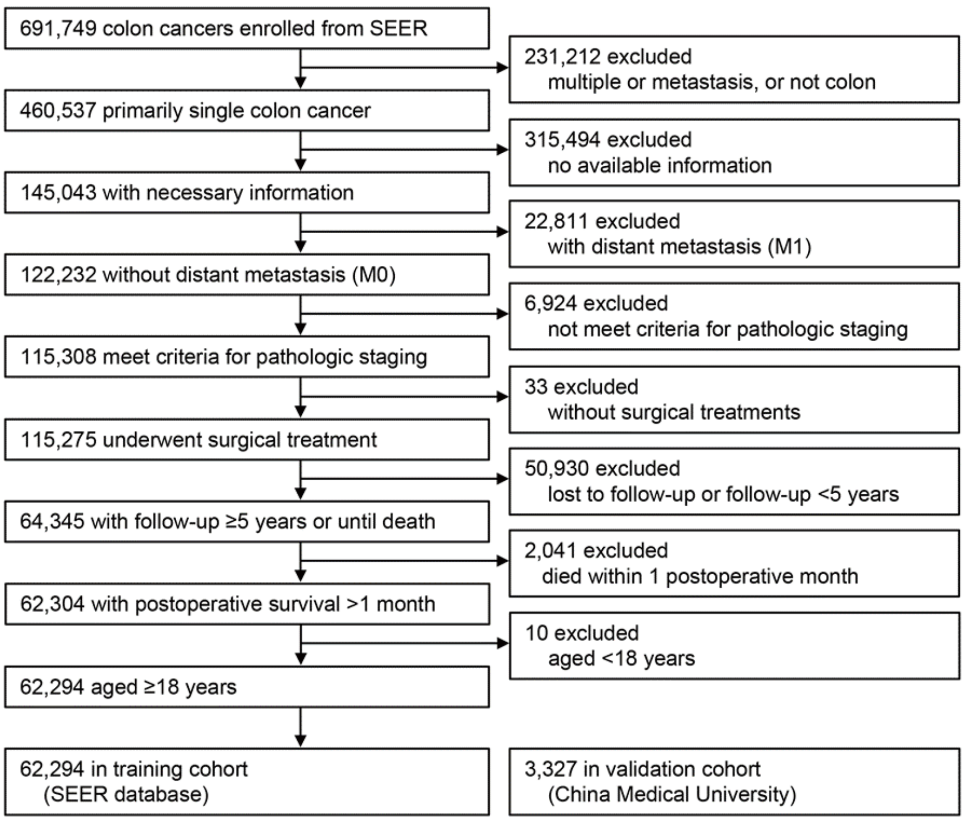

Figure S1 Data screening diagram.

https://dx.doi.org/10.21037/atm-21-3170

Table S1 Characteristics of baselines in the training and validation cohorts

Variables	Training cohort, N = 62,294 (%)	Validation cohort, N = 3,327 (%)
Gender		
Female	33,192 (53.3)	1,452 (43.6)
Male	29,102 (46.7)	1,875 (56.4)
Age, years		
Mean (SD)	68.1 (13.8)	59.9 (11.6)
Tumor size, cm		
Median (SD)	4.5 (2.2)	5.5 (2.1)
Histological grade		
Grade I	5,523 (8.9)	99 (3.0)
Grade II	44,452 (71.4)	2,610 (78.4)
Grade III	11,245 (18.1)	615 (18.5)
Grade IV	1,074 (1.7)	3 (0.1)
AJCC 8th T stage		
T1	5,879 (9.4)	22 (0.7)
T2	10,729 (17.2)	68 (2.0)
T3	38,359 (61.6)	1,789 (53.8)
T4a	4,352 (7.0)	1,015 (30.5)
T4b	2,975 (4.8)	433 (13.0)
AJCC 8th N stage		
N0	37,998 (61.0)	1,298 (39.0)
N1a	7,694 (12.4)	723 (21.7)
N1b/1c	7,705 (12.4)	709 (21.3)
N2a	4,988 (8.0)	345 (10.4)
N2b	3,909 (6.3)	252 (7.6)
Retrieved lymph nodes		
<12	16,674 (26.8)	1,052 (31.6)
≥12	45,620 (73.2)	2,275 (68.4)
Mean (SD)	17.2 (9.6)	16.7 (10.0)
Adjuvant chemotherapy		
Yes	–	1,582 (47.6)
No	–	1,739 (52.3)
Unknown	–	6 (0.2)

AJCC, American Joint Committee on Cancer; SD, standard deviation.

https://dx.doi.org/10.21037/atm-21-3170

Table S2 Baseline characteristics of the training and validation cohorts in relation to the number of retrieved lymph nodes

Variables	Training cohort, N=62,294 (%)			Validation cohort, N=3,327 (%)		
	Retrieved LNs < 12	Retrieved LNs ≥ 12	P value	Retrieved LNs < 12	Retrieved LNs ≥ 12	P value
Age, year			<0.001			0.006
≤60 years	4,028 (24.2)	14,330 (31.4)		469 (44.6)	1,131 (49.7)	
>60 years	12,646 (75.8)	31,290 (68.6)		583 (55.4)	1,144 (50.3)	
Gender			<0.001			0.910
Female	8,631 (51.8)	24,561 (53.8)		461 (43.8)	991(43.6)	
Male	8,043 (48.2)	21,059 (46.2)		591 (56.2)	1,284 (56.4)	
Race			0.002			–
White	13,064 (78.3)	36,341 (79.7)		–	–	
Black	2,110 (12.7)	5,457 (12.0)		–	–	
Other	1,500 (9.0)	3,822 (8.4)		–	–	
Tumor size, cm			<0.001			<0.001
≤4.5 cm	11,908 (71.4)	24,647 (54.0)		453 (43.1)	672 (29.5)	
>4.5 cm	4,766 (28.6)	20,973 (46.0)		599 (56.9)	1,603 (70.5)	
Histological grade			<0.001			0.034
Grade I	1,847 (11.1)	3,676 (8.1)		34 (3.2)	65 (2.9)	
Grade II	12,190 (73.1)	32,262 (70.7)		853 (81.1)	1,757 (77.2)	
Grade III	2,455 (14.7)	8,790 (19.3)		164 (15.6)	451 (19.8)	
Grade IV	182 (1.1)	892 (2.0)		1 (0.1)	2 (0.1)	
AJCC 8th pT stage			<0.001			<0.001
pT1	2,420 (14.5)	3,459 (7.6)		13 (1.2)	9 (0.4)	
pT2	3,317 (19.9)	7,412 (16.2)		32 (3.0)	36 (1.6)	
pT3	9,167 (55.0)	29,192 (64)		481 (45.7)	1,308 (57.5)	
pT4a	1,065 (6.4)	3,287 (7.2)		408 (38.8)	607 (26.7)	
pT4b	705 (4.2)	2,270 (5.0)		118 (11.2)	315 (13.8)	
AJCC 8th pN stage			<0.001			<0.001
pN0	11,022 (66.1)	26,976 (59.1)		451 (42.9)	847 (37.2)	
pN1a	2,103 (12.6)	5,591 (12.3)		243 (23.1)	480 (21.1)	
pN1b	2,038 (12.2)	5,667 (12.4)		234 (22.2)	475 (20.9)	
pN2a	1,166 (7.0)	3,822 (8.4)		93 (8.8)	252 (11.1)	
pN2b	345 (2.1)	3,564 (7.8)		31 (2.9)	221 (9.7)	
Adjuvant chemotherapy			–			0.129
Yes	–	–		521 (49.5)	1,061 (46.6)	
No	–	–		530 (50.4)	1,209 (53.1)	
Unknown	–	–		1 (0.1)	5 (0.3)	

Abbreviations: AJCC, American Joint Committee on Cancer; LNs, lymph nodes.

https://dx.doi.org/10.21037/atm-21-3170

Table S3 Univariate and multivariable analyses of prognostic factors in the training and validation cohorts in relation to the number of retrieved lymph nodes

Variables	Training cohort, rLNs <12			Training cohort, rLNs ≥12			Validation cohort, rLNs <12			Validation cohort, rLNs ≥12		
	UA		MA	UA		MA	UA		MA	UA		MA
	5-Y OS	P	P	5-Y OS	P	P	5-Y OS	P	P	5-Y OS	P	P
Age, year		<0.001	<0.001		<0.001	<0.001		<0.001	<0.001		<0.001	<0.001
≤60 years	79.4%			82.9%			79.6%			82.0%		
>60 years	57.6%			64.1%			67.6%			73.3%		
Gender		0.886	–		0.092	–		0.509	–		0.045	0.036
Female	62.7%			70.4%			74.7%			79.5%		
Male	62.9%			69.4%			71.6%			76.5%		
Race		<0.001	<0.001		<0.001	<0.001		–	–		–	–
White	62.4%			69.6%			–			–		
Black	60.0%			67.9%			–			–		
Other	70.5%			76.3%			–			–		
Size, cm		<0.001	<0.001		<0.001	<0.001		<0.001	0.021		0.013	0.242
≤4.5 cm	66.9%			72.6%			79.2%			81.8%		
>4.5 cm	52.5%			66.8%			68.6%			76.2%		
Histological grade		<0.001	<0.001		<0.001	<0.001		0.005	0.174		<0.001	0.164
Grade I	70.8%			76.8%			100%			70.0%		
Grade II	64.3%			72.0%			73.8%			79.4%		
Grade III	50.3%			60.8%			63.6%			73.2%		
Grade IV	53.8%			56.3%			100%			100%		
AJCC 8th pT stage		<0.001	<0.001		<0.001	<0.001		<0.001	<0.001		0.011	<0.001
pT1	80.6%			84.6%			88.9%			100%		
pT2	73.2%			80.3%			92.6%			96.6%		
pT3	58.9%			69.6%			78.8%			77.9%		
pT4a	43.0%			51.5%			70.5%			79.4%		
pT4b	33.2%			44.9%			56.9%			71.9%		
AJCC 8th pN stage		<0.001	<0.001		<0.001	<0.001		<0.001	<0.001		<0.001	<0.001
pN0	69.0%			76.9%			76.3%			83.7%		
pN1a	58.1%			70.0%			77.8%			80.5%		
pN1b	52.0%			64.3%			67.3%			77.2%		
pN2a	41.0%			56.2%			66.6%			73.3%		
pN2b	31.1%			40.7%			49.7%			55.5%		
Adjuvant chemotherapy		–	–		–	–		<0.001	<0.001		0.009	<0.001
Yes	–			–			78.8%			80.3%		
No	–			–			67.5%			77.1%		
Unknown	–			–			100%			100%		

Abbreviations: AJCC, American Joint Committee on Cancer; P, p value; rLNs, retrieved lymph nodes. MA, multivariate analysis; UA, univariate analysis.

https://dx.doi.org/10.21037/atm-21-3170

Table S4 The proposed LNR stage in the training cohort

Values	5-Y OS, % (95% CI)	HRs (95% CI)[†]	Log-rank (Mantel-Cox)[‡]	
			χ^2 value	P value
LNR1	74.5 (74.1-75.0)	--	--	--
0.00 (n=37,998)	74.6 (74.2-75.0)	1.00 (reference)	--	--
(0.00, 0.05] (n=2,578)	73.7 (71.9-75.3)	0.95 (0.89-1.02)	1.83	0.176
LNR2	62.4 (61.6-63.1)	--	--	--
(0.05, 0.10] (n=5,507)	67.3 (66.0-68.5)	1.26 (1.20-1.31)	50.5	<0.001
(0.10, 0.15] (n=3,229)	62.7 (61.0-64.3)	1.40 (1.33-1.48)	11.2	0.001
(0.15, 0.20] (n=2,877)	59.6 (57.7-61.3)	1.58 (1.50-1.67)	11.1	0.001
(0.20, 0.25] (n=1,846)	57.5 (55.2-59.7)	1.66 (1.55-1.77)	1.39	0.238
(0.25, 0.30] (n=1,266)	53.8 (51.0-56.5)	1.79 (1.66-1.94)	2.25	0.134
LNR3	48.6 (47.0-50.1)	--	--	--
(0.30, 0.35] (n=1,182)	49.8 (46.9-52.7)	2.02 (1.87-2.18)	4.52	0.033
(0.35, 0.40] (n=1,133)	50.0 (47.1-52.9)	2.09 (1.93-2.26)	0.509	0.476
(0.40, 0.45] (n=672)	47.0 (43.1-50.7)	2.21 (2.00-2.43)	0.821	0.365
(0.45, 0.50] (n=952)	46.4 (43.2-49.5)	2.22 (2.04-2.40)	0.004	0.95
LNR4	40.6 (38.1-43.0)	--	--	--
(0.50, 0.55] (n=277)	39.0 (33.2-44.7)	2.67 (2.31-3.09)	4.50	0.034
(0.55, 0.60] (n=570)	41.8 (37.7-45.8)	2.52 (2.28-2.79)	0.226	0.634
(0.60, 0.65] (n=278)	39.4 (33.7-45.1)	2.69 (2.33-3.10)	0.365	0.546
(0.65, 0.70] (n=391)	40.7 (35.8-45.6)	2.60 (2.31-2.94)	0.04	0.841
LNR5	26.9 (24.7-29.2)	--	--	--
(0.70, 0.75] (n=329)	32.1 (27.0-37.2)	3.45 (3.05-3.91)	8.62	0.003
(0.75, 0.80] (n=255)	27.5 (22.1-33.1)	3.71 (3.22-4.26)	0.682	0.409
(0.80, 0.85] (n=174)	26.6 (20.3-33.4)	4.09 (3.45-4.85)	0.84	0.359
(0.85, 0.90] (n=207)	26.3 (20.5-32.5)	4.13 (3.54-4.81)	0	0.987
(0.90, 0.95] (n=135)	21.3 (14.8-28.6)	4.84 (4.02-5.83)	1.20	0.273
(0.95, 1.00] (n=438)	25.0 (21.0-29.1)	4.17 (3.75-4.63)	1.05	0.305

Abbreviations: 5Y-OS, 5-year overall survival; CI, confidence interval; HR, hazard ratio; LNR, lymph node ratio; No., number. [†]Log-rank tests were conducted between two sequential stages and twenty-one χ^2 values were generated. All stages were compared with LNR = 0 as reference by values of HRs of Cox proportional hazards. [‡]Log-rank tests were conducted between two sequential stages.

https://dx.doi.org/10.21037/atm-21-3170

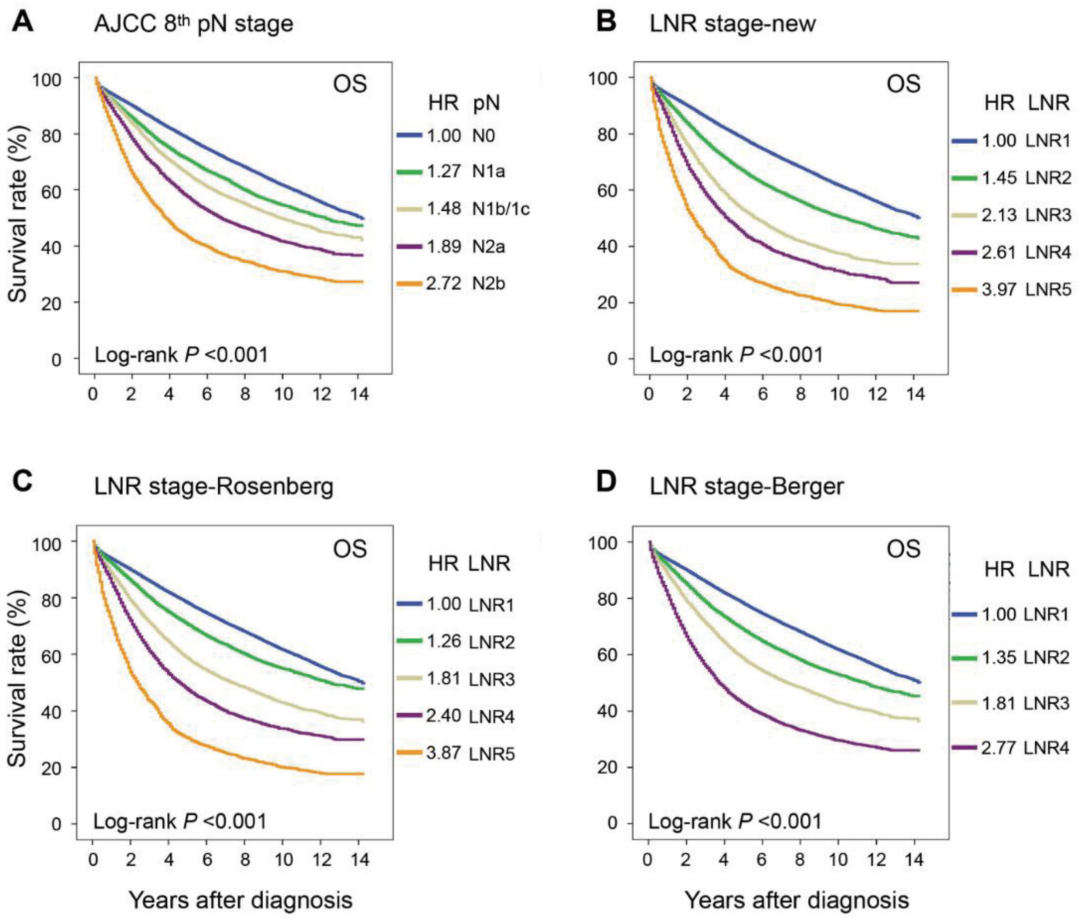

Figure S2 Kaplan-Meier estimates of the AJCC 8th pN stage and three LNR stages in the training cohort. (A) AJCC 8th pN stage. (B) LNR stage-new. (C) LNR stage-Rosenberg. (D) LNR stage-Berger. AJCC, American Joint Committee on Cancer; LNR, lymph node ratio.

https://dx.doi.org/10.21037/atm-21-3170

Table S5 Comparisons of LNR stages and the AJCC 8th pN stage in the training cohort

Training cohort	AIC†	AUC (95% CI)‡	P value*
Overall patients (N=62,294)			
AJCC 8th pN stage	563,291	0.584 (0.581-0.588)	reference
LNR stage -Pei§	562,703	0.589 (0.585-0.593)	<0.001
LNR stage-Rosenberg¶	562,771	0.588 (0.584-0.592)	<0.001
LNR stage-Berger&	562,863	0.590 (0.586-0.594)	<0.001
Patients with lymph nodes < 12 (n=16,674)			
AJCC 8th pN stage	132,884	0.584 (0.577-0.592)	reference
LNR stage-Pei§	132,751	0.589 (0.581-0.596)	0.017
LNR stage-Rosenberg¶	132,765	0.587 (0.579-0.594)	0.018
LNR stage-Berger&	132,795	0.589 (0.581-0.596)	0.007
Patients with lymph nodes ≥ 12 (n=45,620)			
AJCC 8th pN stage	399,397	0.584 (0.580-0.589)	reference
LNR stage-Pei§	398,942	0.589 (0.585-0.593)	<0.001
LNR stage-Rosenberg¶	398,993	0.589 (0.584-0.593)	<0.001
LNR stage-Berger&	399,057	0.591 (0.586-0.595)	<0.001

Abbreviations: AIC, Akaike's information criterion; AJCC, American Joint Committee on Cancer; AUC, Areas under the receiver-operating characteristic curve; CI, confidence interval; DFS, disease-free survival; LNR, lymph node ratio. †A lower AIC indicates superior model-fitting. ‡A higher AUC indicates better discrimination. §Cutoff values of LNR-Pei: 0.05, 0.30, 0.50, 0.70. ¶Cutoff values of LNR-Rosenberg: 0, 0.17, 0.41, 0.69. &Cutoff values of LNR-Berger: 0.05, 0.19, 0.39. *P value of Hanley & McNeil test of AUCs.

https://dx.doi.org/10.21037/atm-21-3170

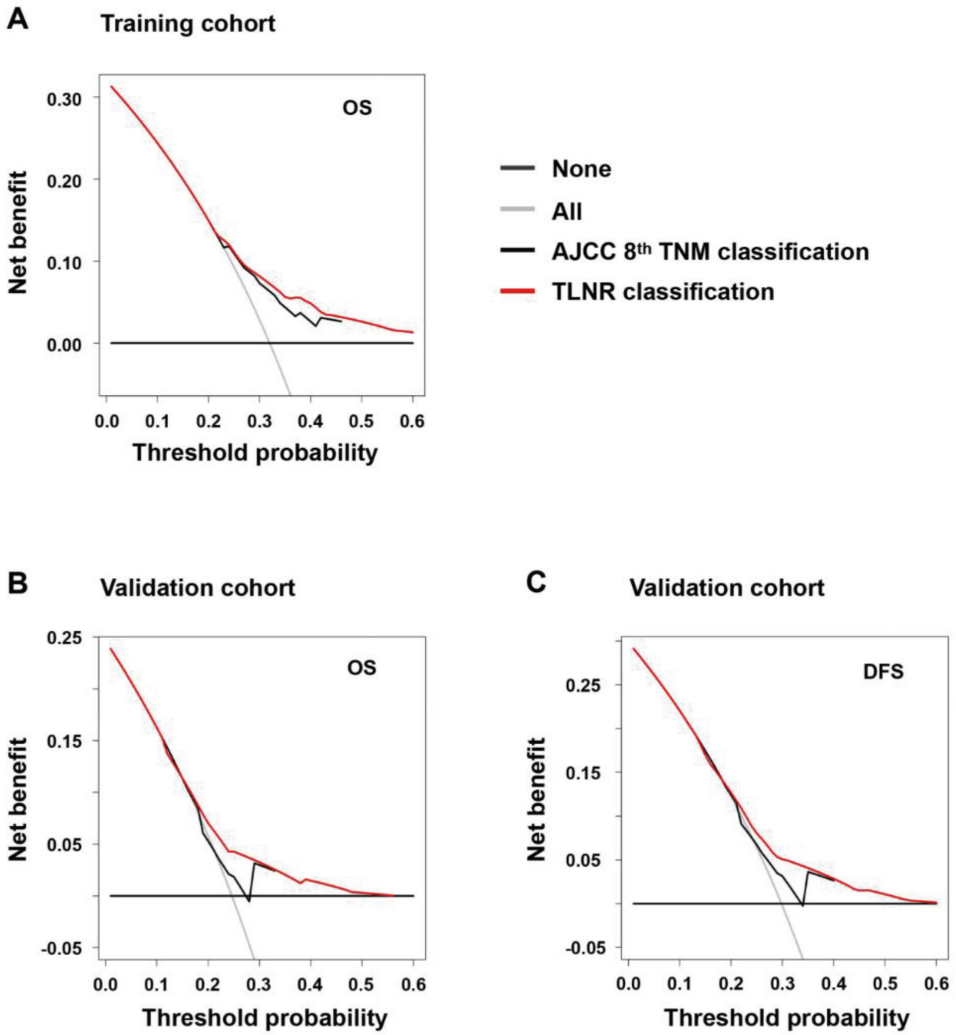

Figure S3 Decision curve analyses between the AJCC 8th TNM and TLNR classifications in the training and validation cohorts. (A) OS in the training cohort. (B) OS in the validation cohort. (C) DFS in the validation cohort. In decision curve analyses, the horizontal solid black line is the net benefit of treating no patients assuming that no patients would die, and the thin grey line is the net benefit of treating all patients similarly regardless of their stages assuming all patients would die. The dotted red line and the dotted black line is the net benefit of treating patients based on the AJCC 8th and TLNR classifications. AJCC, American Joint Committee on Cancer; TNM, tumor/node/metastasis; TLNR, T stage-Lymph Node Ratio classification; LNR, lymph node ratio; OS, overall survival; DFS, disease-free survival.

https://dx.doi.org/10.21037/atm-21-3170