

Figure S1 Stepwise extraction process from the Surveillance, Epidemiology, and End Results database.

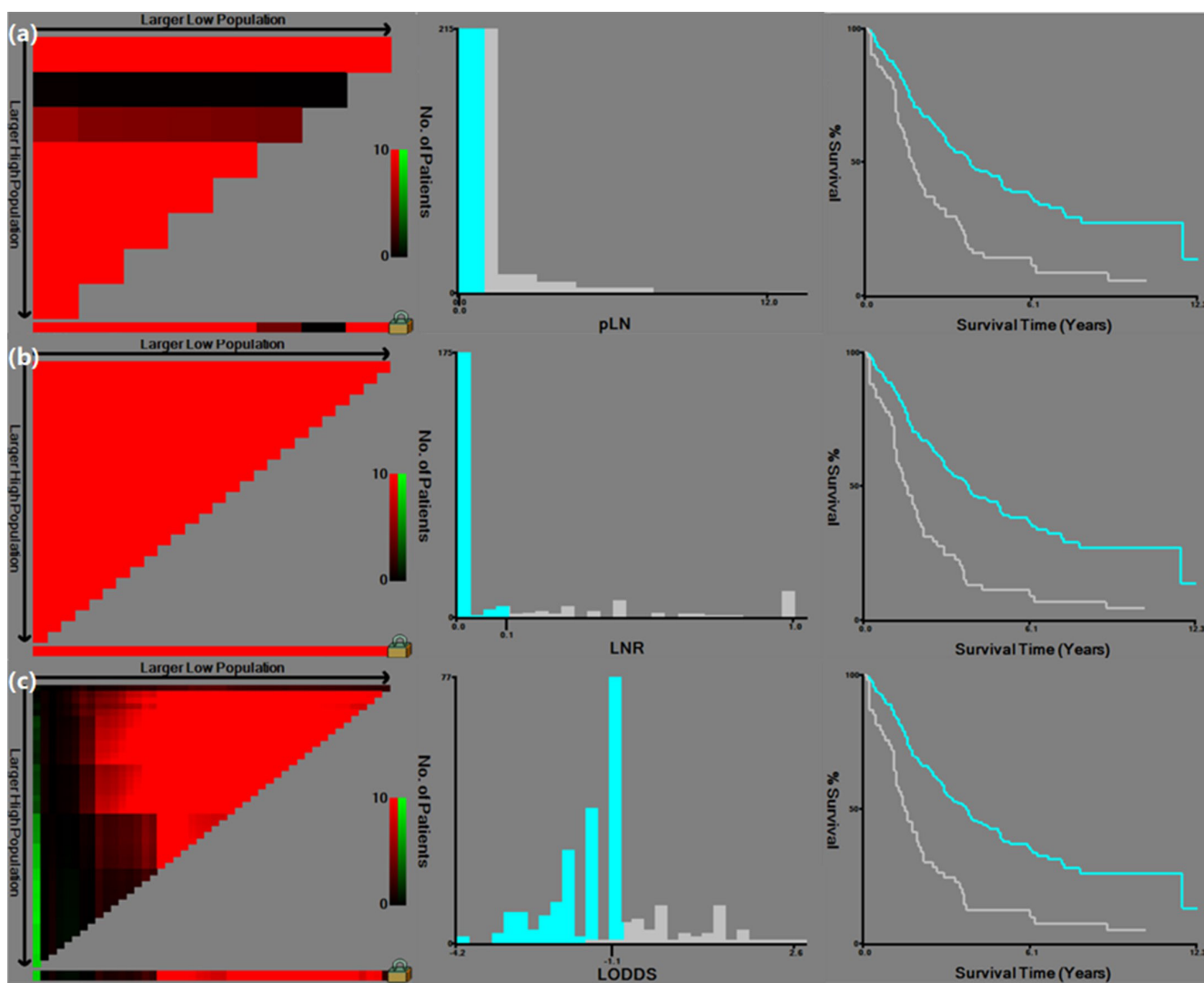


Figure S2 X-tile plots for determining optimal cut-off values of different staging schemes according to cancer-specific survival. (A) Number of positive lymph nodes (pLN); (B) ratio of positive lymph nodes (LNR); (C) Log odds of positive lymph nodes (LODDS).

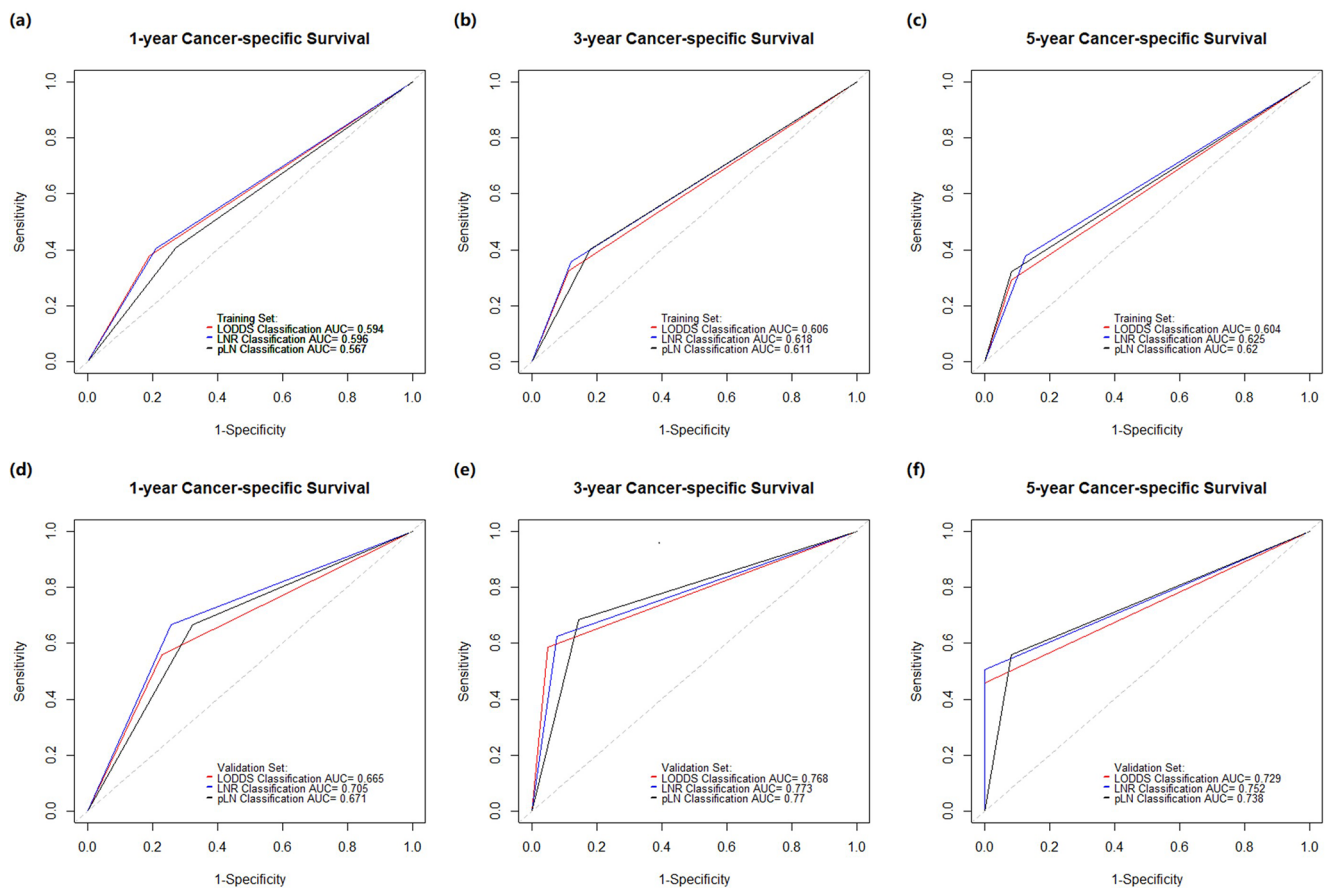


Figure S3 Receiver operative characteristics (ROC) analysis of the pLN, LNR and LODDS classifications in the prediction of prognosis of patients with LND at 1-, 3-, 5- year point for cancer-specific survival (CSS). (A,B,C) Training set; (D,E,F) Validation set. pLN, Number of positive lymph nodes; LNR, Positive lymph node ratio; LODDS, Log odds of positive lymph nodes; AUC, Area under the curve; LND, Lymph node dissection.

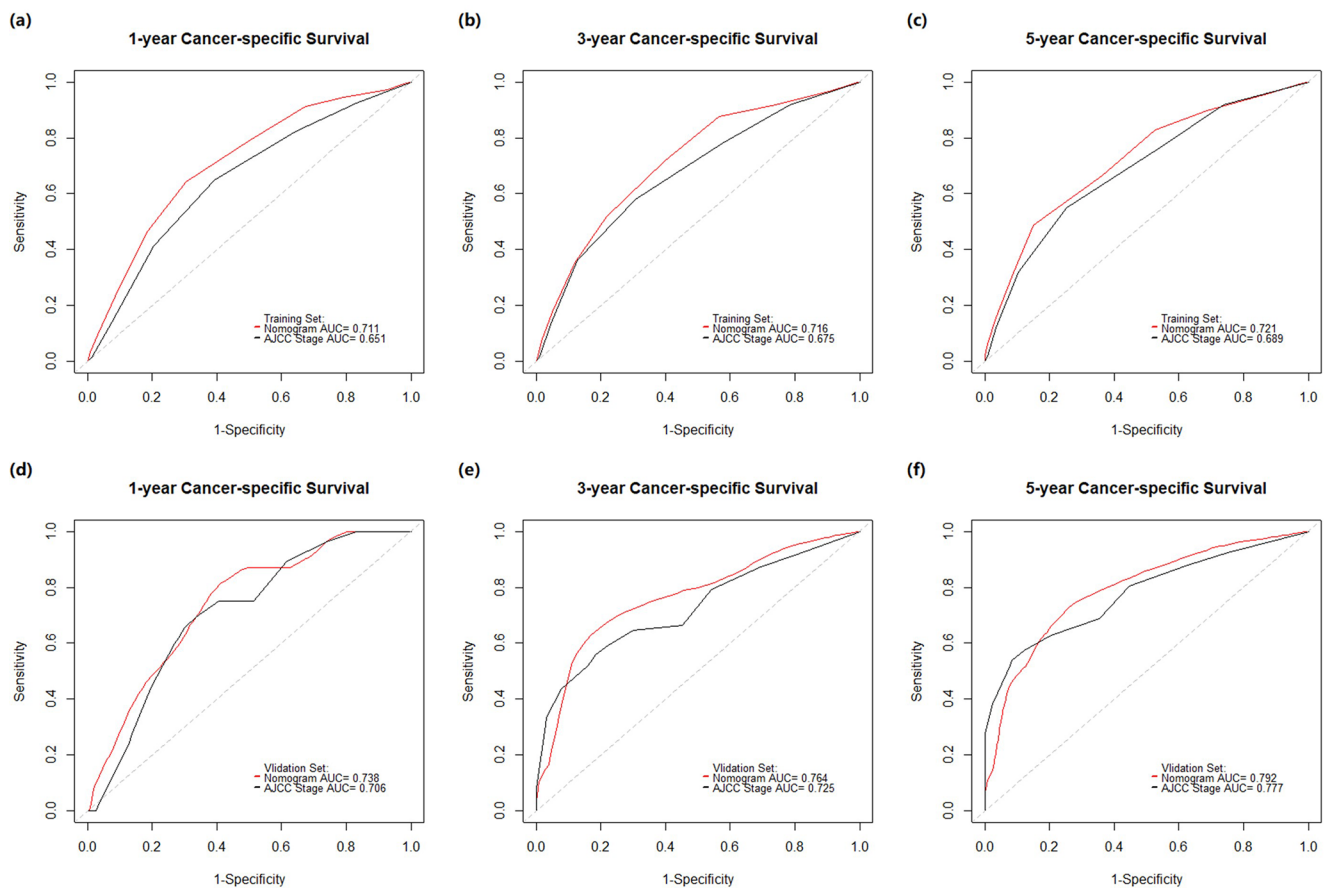


Figure S4 Receiver operative characteristics (ROC) analysis of the nomogram and the current AJCC staging system (8th edition) in the prediction of prognosis of patients at 1-, 3-, 5-year point for cancer-specific survival (CSS). (A,B,C) Training set; (D,E,F) Validation set. AJCC, American Joint Committee on Cancer; AUC, Area under the curve.

Table S1 Comparison between LND and non-LND before and after PSM in all patients

Factors	Before PSM			After PSM		
	LND (n=331)	non-LND (n=313)	P	LND (n=251)	non-LND (n=251)	P
Age			<0.001			0.464
≤60	162 (48.9)	106 (33.9)		102 (40.6)	94 (37.5)	
>60	169 (51.1)	207 (66.1)		149 (59.4)	157 (62.5)	
Sex			0.423			0.473
Female	187 (56.5)	167 (53.4)		143 (57.0)	135 (53.8)	
Male	144 (43.5)	146 (46.6)		108 (43.0)	116 (46.2)	
Race			0.001			0.621
White	273 (82.4)	226 (72.2)		202 (80.5)	195 (77.7)	
Asia-Pacific	29 (8.8)	59 (18.8)		26 (10.4)	33 (13.1)	
Other	29 (8.8)	28 (8.9)		23 (9.2)	23 (9.2)	
Tumor numbers			0.050			0.380
Single	306 (92.4)	275 (87.9)		228 (90.8)	222 (88.4)	
Multiple	25 (7.6)	38 (12.1)		23 (9.2)	29 (11.6)	
Neoadjuvant therapy			0.729			0.715
Yes	20 (6.0)	21 (6.7)		15 (6.0)	17 (6.8)	
No	311 (94.0)	292 (93.3)		236 (94.0)	234 (93.2)	
Radiotherapy			0.003			0.780
Yes	60 (18.1)	31 (9.9)		28 (11.2)	30 (12.0)	
No/unknown	271 (81.9)	282 (90.1)		223 (88.8)	221 (88.0)	
Chemotherapy			<0.001			0.855
Yes	159 (48.0)	107 (34.2)		99 (39.4)	97 (38.6)	
No/unknown	172 (52.0)	206 (65.8)		152 (60.6)	154 (61.4)	
AFP			0.232			0.774
Negative	133 (40.2)	145 (46.3)		108 (43.0)	107 (42.6)	
Positive	38 (11.5)	37 (11.8)		26 (10.4)	31 (12.4)	
Borderline/unknown	160 (48.3)	131 (41.9)		117 (46.6)	113 (45.0)	
Fibrosis score			0.073			0.359
0-4	51 (15.4)	47 (15.0)		40 (15.9)	36 (14.3)	
5-6	14 (4.2)	27 (8.6)		10 (4.0)	17 (6.8)	
Unknown	266 (80.4)	239 (76.4)		201 (80.1)	198 (78.9)	
Tumor size (cm)			0.019			0.200
≤5	142 (42.9)	153 (48.9)		111 (44.2)	111 (44.2)	
5-10	137 (41.4)	133 (42.5)		106 (42.2)	118 (47.0)	
>10	52 (15.7)	27 (8.6)		34 (13.5)	22 (8.8)	
T stage			<0.001			0.950
T1a	68 (20.5)	100 (39.5)		66 (26.3)	68 (27.1)	
T1b	68 (20.5)	64 (25.3)		56 (22.3)	53 (21.1)	
T2	87 (26.3)	27 (10.7)		75 (29.9)	71 (28.3)	
T3	90 (27.2)	55 (21.7)		50 (19.9)	53 (21.1)	
T4	18 (5.4)	7 (2.8)		4 (1.6)	6 (2.4)	
Grade [†]			0.206			0.386
G1	30 (9.1)	42 (13.4)		24 (9.6)	32 (12.7)	
G2	192 (58.0)	176 (56.2)		143 (57.0)	130 (51.8)	
G3-4	109 (32.9)	95 (30.4)		84 (33.5)	89 (35.5)	
Median survival time (months)	39.0	55.0	0.001	42.0	49.0	0.095

[†]G1, Well differentiated; G2, Moderately differentiated; G3-4, Poorly differentiated/Undifferentiated. LND, Lymph node dissection; PSM, Propensity score matching; AFP, Alpha Fetoprotein.

Table S2 Comparison between LNM and non-LNM before and after PSM in patients with LND

Factors	Before PSM			After PSM		
	LNM (n=03)	non-LNM (n=228)	P	LNM (n=96)	non-LNM (n=96)	P
Age			0.706			0.564
≤60	52 (50.5)	110 (48.2)		50 (52.1)	45 (46.9)	
>60	51 (49.5)	118 (51.8)		46 (47.9)	51 (53.1)	
Sex			0.964			0.461
Female	58 (56.3)	129 (56.6)		55 (52.3)	61 (63.5)	
Male	45 (43.7)	99 (43.4)		41 (42.7)	35 (36.5)	
Race			0.707			0.655
White	83 (80.6)	190 (83.3)		78 (81.2)	82 (85.4)	
Asia-Pacific	11 (10.7)	18 (7.9)		10 (10.4)	9 (9.4)	
Other	9 (8.7)	20 (8.8)		8 (8.3)	5 (5.2)	
Tumor numbers			0.089			1.000
Single	99 (96.1)	207 (90.8)		92 (95.8)	92 (95.8)	
Multiple	4 (3.9)	21 (9.2)		4 (4.2)	4 (4.2)	
Neoadjuvant therapy			0.911			0.495
Yes	6 (5.8)	14 (6.1)		6 (6.2)	3 (3.1)	
No	97 (94.2)	214 (93.9)		90 (93.8)	93 (96.9)	
Radiotherapy			0.004			1.000
Yes	28 (27.2)	32 (14.0)		22 (22.9)	22 (22.9)	
No/unknown	75 (72.8)	196 (86.0)		74 (77.1)	74 (77.1)	
Chemotherapy			0.003			1.000
Yes	62 (60.2)	97 (42.5)		56 (58.3)	56 (58.3)	
No/unknown	41 (39.8)	131 (57.5)		40 (41.7)	40 (41.7)	
AFP			0.278			0.283
Negative	35 (34.0)	98 (43.0)		31 (34.4)	41 (42.7)	
Positive	12 (11.7)	26 (11.4)		11 (11.5)	14 (14.6)	
Borderline/unknown	56 (54.3)	104 (45.6)		52 (54.2)	41 (42.7)	
Fibrosis score			0.613			0.669
0-4	15 (14.6)	36 (15.8)		14 (14.6)	17 (17.7)	
5-6	6 (5.8)	8 (3.5)		5 (5.2)	3 (3.1)	
Unknown	82 (79.6)	184 (80.7)		77 (80.2)	76 (79.2)	
Tumor size (cm)			0.654			0.936
≤5	43 (41.7)	99 (43.4)		39 (40.6)	38 (39.6)	
5-10	41 (39.8)	96 (42.1)		39 (40.6)	38 (39.6)	
>10	19 (18.4)	33 (14.5)		18 (18.8)	20 (20.8)	
T stage			0.017			0.964
T1a	14 (13.6)	54 (23.7)		14 (14.6)	14 (14.6)	
T1b	15 (14.6)	53 (23.2)		15 (15.6)	14 (14.6)	
T2	36 (35.0)	51 (22.4)		30 (31.2)	29 (30.2)	
T3	30 (29.1)	60 (26.3)		29 (30.2)	33 (34.4)	
T4	8 (7.8)	10 (4.4)		8 (8.3)	6 (6.2)	
Grade [†]			0.015			0.705
G1	3 (2.9)	27 (11.8)		3 (3.1)	5 (5.2)	
G2	59 (57.3)	133 (58.3)		57 (59.4)	53 (55.2)	
G3-4	41 (39.8)	68 (29.8)		36 (37.5)	38 (39.6)	
Median survival time (months)	19.0	54.0	<0.001	19.0	45.0	<0.001

[†]G1, Well differentiated; G2, Moderately differentiated; G3-4, Poorly differentiated/Undifferentiated. LNM, Lymph node metastasis; LND, Lymph node dissection; PSM, Propensity score matching; AFP, Alpha Fetoprotein.

Table S3 Multivariable analyses for preoperative risk factors for LNM in patients with LND

Factors	OR	95% CI	P
T stage			
T1a	Reference		
T1b	1.092	0.480-2.481	0.834
T2	2.723	1.317-5.629	0.007
T3	1.929	0.926-4.015	0.079
T4	3.086	1.027-9.269	0.045
Tumor numbers			
Single			0.119
Multiple			

LNM, Lymph node metastasis; LND, Lymph node dissection; OR, Odds ratio; CI, Confidence interval.

Table S4 Comparison between AT and non-AT before and after PSM in non-LNM patients

Factors	Before PSM			After PSM		
	AT (n=100)	non-AT (n=128)	P	AT (n=54)	non-AT (n=54)	P
Age			0.160			1.000
≤60	54 (54.0)	56 (43.8)		28 (51.9)	28 (51.9)	
>60	46 (46.0)	72 (56.2)		26 (48.1)	26 (48.1)	
Sex			0.111			0.314
Female	63 (63.0)	66 (51.6)		32 (59.3)	38 (70.4)	
Male	37 (37.0)	62 (48.4)		22 (40.7)	16 (29.6)	
Race			0.582			0.941
White	86 (86.0)	104 (81.2)		46 (85.2)	45 (83.3)	
Asia-Pacific	6 (6.0)	12 (9.4)		4 (7.4)	5 (9.3)	
Other	8 (8.0)	12 (9.4)		4 (7.4)	4 (7.4)	
Tumor numbers			0.743			0.713
Single	92 (92.0)	115 (89.8)		49 (90.7)	51 (94.4)	
Multiple	8 (8.0)	13 (10.2)		5 (9.3)	3 (5.6)	
Neoadjuvant therapy			0.015			1.000
Yes	11 (11.0)	3 (2.3)		3 (5.6)	2 (3.7)	
No	89 (89.0)	125 (97.7)		51 (94.4)	52 (96.3)	
AFP			0.021			0.176
Negative	33 (33.0)	65 (50.8)		21 (38.9)	30 (55.6)	
Positive	15 (15.0)	11 (8.6)		7 (13.0)	7 (13.0)	
Borderline/unknown	52 (52.0)	52 (40.6)		26 (48.1)	17 (31.5)	
Fibrosis score			0.283			0.839
0-4	13 (13.0)	23 (18.0)		10 (18.5)	10 (18.5)	
5-6	2 (2.0)	6 (4.7)		1 (1.9)	2 (3.7)	
Unknown	85 (85.0)	99 (77.3)		43 (79.6)	42 (77.8)	
TLNC [M (IQR)]	2 (1,4)	2 (1,3.5)	0.436	2 (1,4)	2 (1,4)	0.750
Tumor size (cm)			0.025			0.276
≤5	44 (44.0)	55 (43.0)		23 (42.6)	19 (35.2)	
5-10	35 (35.0)	61 (47.7)		19 (35.2)	27 (50.0)	
>10	21 (21.0)	12 (9.4)		12 (22.2)	8 (14.8)	
T stage			0.015			0.850
T1a	20 (20.0)	34 (26.6)		11 (20.4)	10 (18.5)	
T1b	16 (16.0)	37 (28.9)		9 (16.7)	11 (20.4)	
T2	25 (25.0)	26 (20.3)		16 (29.6)	16 (29.6)	
T3	31 (31.0)	29 (22.7)		18 (33.3)	16 (29.6)	
T4	8 (8.0)	2 (1.6)		0 (0)	1 (1.9)	
Grade [†]			0.088			0.635
G1	9 (9.0)	18 (14.1)		4 (7.4)	6 (11.1)	
G2	54 (54.0)	79 (61.7)		30 (55.6)	32 (59.3)	
G3-4	37 (37.0)	31 (24.2)		20 (37.0)	16 (29.6)	
Median survival time (months)	45.0	74.0	0.010	45.0	72.0	0.140

[†]G1, Well differentiated; G2, Moderately differentiated; G3-4, Poorly differentiated/Undifferentiated. AT, Adjuvant therapy (chemotherapy and/or radiotherapy); LNM, Lymph node metastasis; PSM, Propensity score matching; TLNC, Total lymph node count; IQR, Interquartile range; AFP, Alpha Fetoprotein.

Table S5 Comparison between AT and non-AT in LNM patients

Factors	AT (n=65)	non-AT (n=38)	P
Age			0.273
≤60	36 (55.4)	16 (42.1)	
>60	29 (44.6)	22 (57.9)	
Sex			0.387
Female	34 (52.3)	24 (70.6)	
Male	31 (47.7)	14 (36.8)	
Race			0.476
White	54 (83.1)	29 (76.3)	
Asia-Pacific	7 (10.8)	4 (10.5)	
Other	4 (6.2)	5 (13.2)	
Tumor numbers			0.279
Single	64 (98.5)	35 (92.1)	
Multiple	1 (1.5)	3 (7.9)	
Neoadjuvant therapy			0.135
Yes	6 (9.2)	0 (0)	
No	59 (90.8)	38 (100.0)	
AFP			0.885
Negative	21 (32.3)	14 (36.8)	
Positive	8 (12.3)	4 (10.5)	
Borderline/unknown	36 (55.4)	20 (52.6)	
Fibrosis score			0.236
0-4	11 (16.9)	4 (10.5)	
5-6	2 (3.1)	4 (10.5)	
Unknown	52 (80.0)	30 (79.0)	
TLNC [M (IQR)]	5 (2,8)	4 (2,9)	0.514
Tumor size (cm)			0.571
≤5	26 (40.0)	17 (44.7)	
5-10	25 (38.5)	16 (42.1)	
>10	14 (21.5)	5 (13.2)	
T stage			0.249
T1a	7 (10.8)	7 (18.4)	
T1b	8 (12.3)	7 (18.4)	
T2	21 (32.3)	15 (39.5)	
T3	22 (33.8)	8 (21.1)	
T4	7 (10.8)	1 (2.6)	
Grade [†]			0.422
G1	1 (1.5)	2 (5.3)	
G2	36 (55.4)	23 (60.5)	
G3-4	28 (43.1)	13 (34.2)	
Median survival time (months)	23.0	13.0	0.018

[†]G1, Well differentiated; G2, Moderately differentiated; G3-4, Poorly differentiated/Undifferentiated. AT, Adjuvant therapy (chemotherapy and/or radiotherapy); LNM, Lymph node metastasis; TLNC, Total lymph node count; IQR, Interquartile range; AFP, Alpha Fetoprotein.

Table S6 Baseline characteristics data of patients in the training and validation set

Factors	Training Set (n=483)	Validation Set (n=161)
Age		
≤60	203 (42.0)	65 (40.4)
>60	280 (58.0)	96 (59.6)
Sex		
Female	254 (52.6)	100 (62.1)
Male	229 (47.4)	61 (37.9)
Race		
White	372 (77.0)	127 (78.9)
Asia-Pacific	68 (14.1)	20 (12.4)
Other	43 (8.9)	14 (8.7)
Tumor numbers		
Single	436 (90.3)	145 (90.1)
Multiple	47 (9.7)	16 (9.9)
Neoadjuvant therapy		
Yes	31 (6.4)	10 (6.2)
No	452 (93.6)	151 (93.8)
AFP		
Negative	213 (44.1)	65 (40.4)
Positive	64 (13.3)	11 (6.8)
Borderline/unknown	206 (42.7)	85 (52.8)
Fibrosis score		
0-4	69 (14.3)	29 (18.0)
5-6	31 (6.4)	10 (6.2)
Unknown	383 (79.3)	122 (75.8)
TLNC [M (IQR)]	1 (0,3)	1 (0,2)
Radiotherapy		
Yes	65 (13.5)	26 (16.1)
No/unknown	418 (86.5)	135 (83.9)
Chemotherapy		
Yes	201 (41.6)	65 (40.4)
No/unknown	282 (58.4)	96 (59.6)
Tumor size (cm)		
≤2	32 (6.6)	12 (7.5)
2-5	188 (38.9)	63 (39.1)
5-10	205 (42.4)	65 (40.4)
>10	58 (12.0)	21 (13.0)
T stage		
T1a	124 (25.7)	44 (27.3)
T1b	94 (19.5)	38 (23.6)
T2	129 (26.7)	45 (28.0)
T3	117 (24.2)	28 (17.4)
T4	19 (3.9)	6 (3.7)
Grade [†]		
G1	54 (11.2)	18 (11.2)
G2	283 (58.6)	85 (52.8)
G3-4	146 (30.2)	58 (36.0)
pN stage		
N0	175 (36.2)	53 (32.9)
N1	73 (15.1)	30 (18.6)
Nx	235 (48.7)	78 (48.4)
Median survival time (months)	43.0	50.0

[†]G1, Well differentiated; G2, Moderately differentiated; G3-4, Poorly differentiated/Undifferentiated. TLNC, Total lymph node count; IQR, Interquartile range; AFP, Alpha Fetoprotein.

Table S7 Survival analysis of patients in the training set

Factors	No. of Patients (n=483)	CSS			Univariable	Multivariate	
		1yr	3yr	5yr	P	HR (95% CI)	P
Age					0.150		
≤60	203 (42.0)	89.4	57.0	43.2		Reference	
>60	280 (58.0)	82.4	51.3	36.8		1.375 (1.076-1.756)	0.011
Sex					0.448		
Female	254 (52.6)	83.8	51.4	37.3			
Male	229 (47.4)	87.1	56.3	42.3			
Race					0.289		
White	372 (77.0)	85.4	53.4	37.5			
Asia-Pacific	68 (14.1)	90.8	57.0	49.5			
Other	43 (8.9)	75.8	51.2	40.7			
Tumor Numbers					<0.001		
Single	436 (90.3)	100.0	88.8	62.7		Reference	
Multiple	47 (9.7)	83.7	49.9	37.1		2.023 (1.259-3.252)	0.004
Neoadjuvant Therapy					0.396		
Yes	31 (6.4)	87.1	45.0	25.0			
No	452 (93.6)	85.2	54.3	40.6			
TLNC [M (IQR)]	1 (0,2)	-	-	-	0.006		
Radiotherapy					0.120		
Yes	65 (13.5)	89.2	49.0	28.6			
No/Unknown	418 (86.5)	84.7	54.5	41.5			
Chemotherapy					0.061		
Yes	201 (41.6)	89.0	55.2	45.0			
No/Unknown	282 (58.4)	82.6	51.8	32.8			
AFP					0.032		
Negative	213 (44.1)	85.2	56.8	46.0			
Positive	64 (13.3)	82.5	40.5	22.8			
Borderline/Unknown	206 (42.7)	86.3	54.7	38.1			
Fibrosis Score					0.901		
0-4	69 (14.3)	89.6	55.0	32.4			
5-6	31 (6.4)	76.7	51.9	44.5			
Unknown	383 (79.3)	85.3	53.6	40.5			
Tumor Size (cm)					<0.001		
≤2	32 (6.6)	86.9	65.7	54.9		Reference	
2-5	188 (38.9)	91.4	65.1	52.3		1.213 (0.679-2.168)	0.514
5-10	205 (42.4)	83.8	48.5	33.7		1.485 (0.814-2.709)	0.198
>10	58 (12.0)	70.2	28.8	12.1		2.461 (1.296-4.673)	0.006
T Stage					<0.001		
T1a	124 (25.7)	94.2	75.3	61.5		Reference	
T1b	94 (19.5)	87.6	57.7	40.4		1.217 (0.726-2.043)	0.456
T2	129 (26.7)	85.2	52.6	39.8		1.477 (0.989-2.207)	0.057
T3	117 (24.2)	76.2	32.3	19.4		1.975 (1.243-3.137)	0.004
T4	19 (3.9)	72.7	28.3	10.6		2.030 (1.081-3.815)	0.028
Grade [†]					<0.001		
G1	54 (11.2)	94.2	78.3	59.4		Reference	
G2	283 (58.6)	87.3	57.4	42.1		1.319 (0.863-2.017)	0.201
G3-4	146 (30.2)	78.4	37.9	28.4		1.815 (1.158-2.843)	0.009

Table S7 (continued)

Table S7 (continued)

Factors	No. of Patients (n=483)	CSS			Univariable	Multivariate	
		1yr	3yr	5yr	P	HR (95% CI)	P
pLN Classification					<0.001		
Unknown	235 (48.7)	87.1	59.4	46.9			
pLN=0	175 (36.2)	86.2	56.3	40.8			
pLN≥1	73 (15.1)	77.4	29.6	14.1			
LNR Classification					<0.001		
Unknown	235 (48.7)	87.1	59.4	46.9		Reference	
LNR≤0.15	188 (38.9)	87.1	56.3	40.2		1.112 (0.852-1.451)	0.434
LNR>0.15	60 (12.4)	72.7	24.2	11.2		2.171 (1.521-3.098)	<0.001
LODDS Classification					<0.001		
Unknown	235 (48.7)	87.1	59.4	46.9			
LODDS≤-0.85	193 (40.0)	86.9	55.3	39.0			
LODDS>-0.85	55 (11.3)	72.0	24.6	12.4			

[†]G1, Well differentiated; G2, Moderately differentiated; G3-4, Poorly differentiated/Undifferentiated. CSS, Cancer-specific survival; HR, Hazard ratio; CI, Confidence interval; TLNC, Total lymph node count; IQR, Interquartile range; pLN, Number of positive lymph nodes; LNR, Positive lymph node ratio; LODDS, Log odds of positive lymph nodes; AFP, Alpha Fetoprotein.