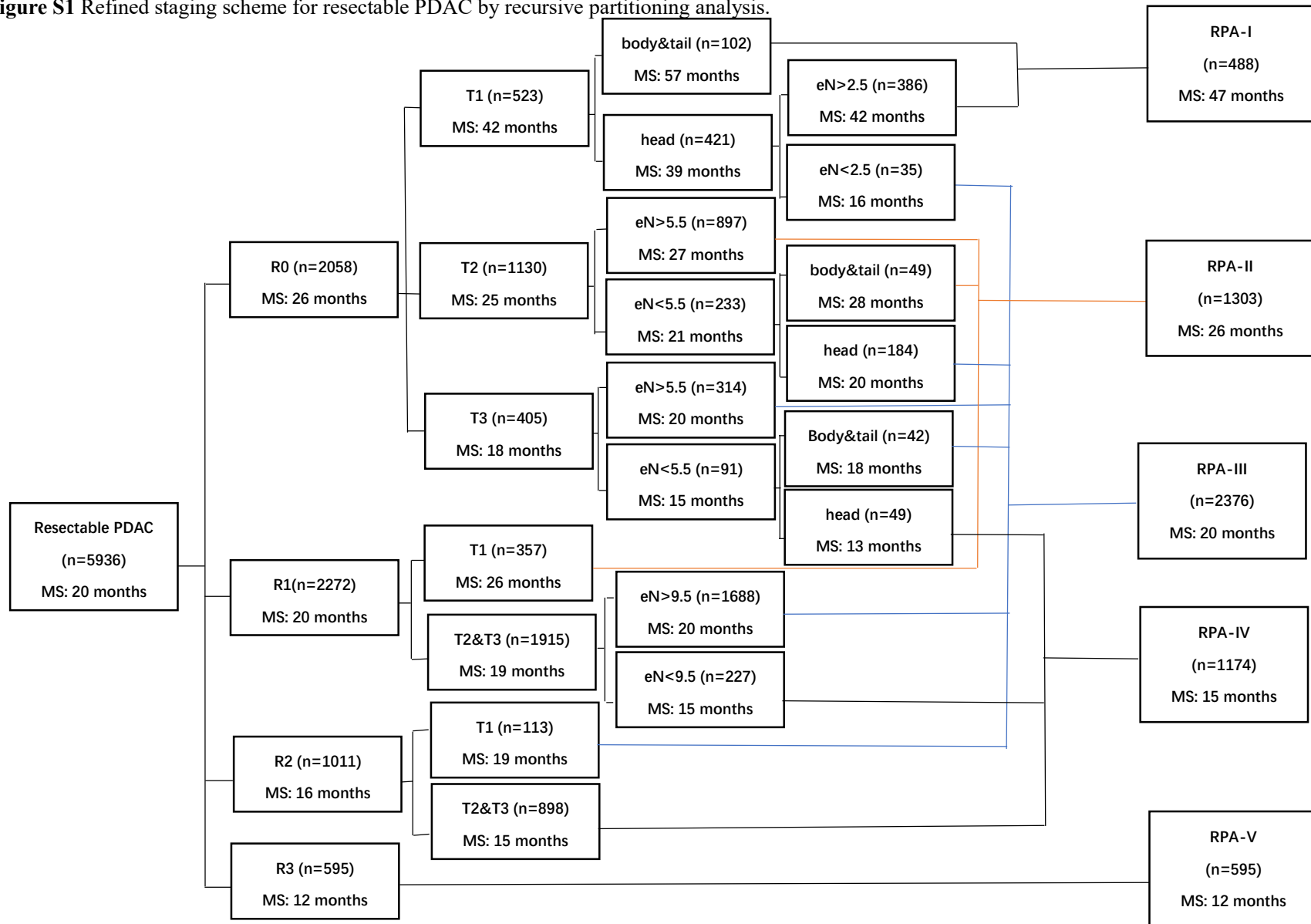


398 **Figure S1** Refined staging scheme for resectable PDAC by recursive partitioning analysis.

399

400



401 **Table S1** Characteristics of patients from SEER database.

Variable	Median(IQR)/N%
Age(years)	66(58-73)
Race	
White	6935 (81.8%)
Black	858 (10.1%)
Other	687 (8.1%)
Sex	
Male	4271 (50.4%)
Female	4209 (49.6%)
Marital status	
Married	5299 (62.5%)
Never married	962 (11.3%)
Other	1960 (23.1%)
Unknown	259 (3.1%)
Date of diagnosis	
2004-2006	1816 (21.4%)
2007-2010	3129 (36.9%)
2011-2014	3535 (41.7%)
Tumor site	
head	7167(84.5%)
body or tail	1313(15.5%)
Tumor grade	
I	801(9.4%)
II	4116(48.5%)
III	2951(34.8%)
IV	74(0.9%)
not determined	538(6.4%)
Tumor size	
≤2cm	1465(17.2%)
>2cm and ≤4cm	5050(59.6%)
>4cm	1965(23.2%)
number of positive lymph node	1(0-3)
0 positive lymph node	2887(34.0%)
1-3 positive lymph nodes	3528(41.6%)
≥4 positive lymph nodes	2065(24.4%)
number of examined lymph nodes	14(9-20)

402 Abbreviations: IQR = interquartile range

403

404

Table S2 Characteristics of patients with positive or negative nodal status.

Variable	Median(IQR)/N%		p-value
	N0	N+	
Age(years)	66(59-74)	65(58-73)	p=0.009
Race			p=0.005
White	2311(80.0%)	4624(82.7%)	
Black	307(10.7%)	551(9.9%)	
Other	269(9.3%)	418(7.4%)	
Sex			p=0.098
Male	1418(49.1%)	2853(51.0%)	
Female	1469(50.9%)	2740(49.0%)	
Marital status			p=0.222
Married	1788(61.9%)	3511(62.8%)	
Never married	313(10.8%)	649(11.6%)	
Other	686(23.9%)	1274(22.8%)	
Unknown	100(3.5%)	159(2.8%)	
Date of diagnosis			p=0.704
2004-2006	627(21.7%)	1189(21.3%)	
2007-2010	1048(36.3%)	2081(37.2%)	
2011-2014	1212(50.0%)	2323(41.5%)	
Tumor site			p<0.001
head	2265(78.5%)	4902(87.6%)	
body or tail	622(21.5%)	691(12.4%)	
Tumor grade			p<0.001
I	359(12.4%)	442(7.9%)	
II	1388(48.1%)	2728(48.8%)	
III	818(28.3%)	2133(38.1%)	
IV	27(0.9%)	47(0.8%)	
not determined	295(10.2%)	243(4.3%)	
Tumor size	30(21-40) mm	33(25-40) mm	p<0.001
≤2cm	706(24.5%)	759(13.6%)	p<0.001
>2cm and ≤4cm	1612(55.8%)	3438(61.5%)	
>4cm	569(19.7%)	1396(24.9%)	
Number of positive lymph nodes			
0 positive lymph node	2887(100%)	0	
1-3 positive lymph nodes	0	3528(63.1%)	
≥4 positive lymph nodes	0	2065(36.9%)	
Number of examined lymph nodes	12(6-18)	15(10-22)	p<0.001

Abbreviations: IQR = interquartile range; N0 = negative nodal status; N+ = positive nodal status.

Table S3 Characteristics of patients with PDAC in head or body&tail of pancreas in SEER.

Variable	Median(IQR)/N%		
Site of tumor	head	body&tail	p-value
Age(years)	65(58-73)	67(59-75)	p=0.001
Race			
White	5920(82.6%)	1015(77.3%)	p<0.001
Black	698(9.7%)	160(12.2%)	
Other	549(7.7%)	138(10.5%)	
Sex			p=0.034
Male	3645(50.9%)	626(47.7%)	
Female	3522(49.1%)	687(52.3%)	
Marital status			p=0.644
Married	4467(62.3%)	832(63.4%)	
Never married	824(11.5%)	138(10.5%)	
Other	1661(23.2%)	299(22.8%)	
Unknown	215(3.0%)	44(3.4%)	
Date of diagnosis			p<0.001
2004-2006	1591(22.2%)	225(17.1%)	
2007-2010	2656(37.1%)	473(36.1%)	
2011-2014	2920(40.7%)	615(46.8%)	
Tumor grade			p=0.097
I+II	4126(57.6%)	791(60.2%)	
III+IV	2591(36.1%)	434(33.1%)	
not determined	450(6.3%)	88(6.7%)	
Tumor size	30(24-40) mm	40(26-52) mm	p<0.001
≤2cm	1282(17.9%)	183(13.9%)	p<0.001
>2cm and ≤4cm	4487(62.6%)	563(42.9%)	
>4cm	1398(19.5%)	567(4.3%)	
Number of positive lymph nodes	2(0-4)	1(0-2)	p<0.001
0 positive lymph node	2265(31.6%)	622(47.4%)	p<0.001
1-3 positive lymph nodes	3026(42.2%)	502(38.2%)	
≥4 positive lymph nodes	1876(26.2%)	189(14.4%)	
Number of examined lymph nodes	14(9-21)	12(7-19)	p<0.001

409

Abbreviations: IQR = interquartile range; head = head of pancreas; body&tail = body and tail of pancreas.

410

411

Table S4 Characteristics of patients in training cohort and internal validation cohort.

Variable	Median(IQR)/N%		p-value
	train cohort	internal validation cohort	
Age(years)	66(58-73)	66(58-74)	p=0.159
Race			p=0.307
White	4866(82.0%)	2069(81.3%)	
Black	582(9.8%)	276(10.9%)	
Other	488(8.2%)	199(7.8%)	
Sex			p=0.270
Male	3013(50.8%)	1258(49.4%)	
Female	2923(49.2%)	1286(50.6%)	
Marital status			p=0.072
Married	3715(62.6%)	1584(62.3%)	
Never married	689(11.6%)	273(10.7%)	
Other	1338(22.5%)	622(24.4%)	
Unknown	194(3.3%)	65(2.6%)	
Date of diagnosis			p=0.741
2004-2006	1258(21.2%)	558(21.9%)	
2007-2010	2194(37.0%)	935(36.8%)	
2011-2014	2484(41.8%)	1051(41.3%)	
Tumor site			p=0.642
head	5024(84.6%)	2143(84.2%)	
body or tail	912(15.4%)	401(15.8%)	
Tumor grade			p=0.604
I+II	3421(57.6%)	1496(58.8%)	
III+IV	2135(36.0%)	890(35.0%)	
not determined	380(6.4%)	158(6.2%)	
Tumor size			p=0.357
≤2cm	1047(17.6%)	418(16.4%)	
>2cm and ≤4cm	3511(59.1%)	1539(60.5%)	
>4cm	1378(23.3%)	587(23.1%)	
number of positive lymph nodes			p=0.168
0 positive lymph node	2058(34.7%)	829(32.6%)	
1-3 positive lymph nodes	2452(41.3%)	1076(42.3%)	
≥4 positive lymph nodes	1426(24.0%)	639(25.1%)	
number of examined lymph nodes	14(9-21)	14(9-20)	p=0.987

413

Abbreviations: IQR=interquartile range

414

Table S5 Prognostic homogeneity comparison between 8th AJCC staging system and our staging system in internal validation set.

	8th AJCC-IA		8th AJCC-IB		8th AJCC-IIA		8th AJCC-IIB		8th AJCC-III		p-value
	n	MS	n	MS	n	MS	n	MS	n	MS	
RPA-I	173	40									-
RPA-II			415	27			129	30	22	18	p=0.372
RPA-III	10	38	67	23	147	20	631	19	169	18	p=0.147
RPA-IV					17	31	244	16	259	14	p=0.058
RPA-V							72	12	189	12	p=0.294
p-value	p=0.9051		p=0.0434		p=0.5469		p<0.0001		p<0.0001		

Abbreviations: AJCC= American Joint Committee on Cancer; MS=median survival; RPA=recursive partitioning analysis

Table S6 Lymph node metastasis and T stage of pancreatic ductal carcinoma.

	T1	T2	T3
All	1465	5050	1965
N+	759	3438	1396
Positive rate	51.8%	68.1%	71.0%

Abbreviations: N+ = positive nodal status

Table S7 Characteristics of patients of R0 with number of examined lymph nodes (eNs)<5.5 and lymph node-positive patients with positive lymph nodes ratios (LNR) <0.078.

	Median(IQR)/N%		p-value
	0<LNR<0.078	R0 with eNs<5.5	
Age(years)	64 (57-72)	67 (59-74)	p<0.001
Race			p=0.689
White	746(80.4%)	463(79.3%)	
Black	105(11.3%)	65(11.1%)	
Other	77(8.3%)	56(9.6%)	
Sex			p=0.053
Male	450(48.5%)	313(53.6%)	
Female	478(51.5%)	271(46.4%)	
Marital status			p=0.214
Married	572(61.6%)	356(61.0%)	
Never married	92(9.9%)	60(10.3%)	
Other	228(24.6%)	156(26.7%)	
Unknown	36(3.9%)	12(2.0%)	
Date of diagnosis			p<0.001
2004-2006	153(16.5%)	192(32.9%)	
2007-2010	316(34.1%)	231(39.6%)	
2011-2014	459(49.4%)	161(27.5%)	
Tumor grade			p<0.001
I+II	537(57.9%)	362(62.0%)	
III+IV	350(37.7%)	152(26.0%)	
not determined	41(4.4%)	70(12.0%)	
Tumor size	31(25-40)mm	30(21-40)mm	p=0.349
Number of examined lymph nodes	21(16-28)	3(2-4)	p<0.001

Abbreviations: IQR = interquartile range; LNR = positive lymph nodes ratios; R0 = negative nodal status; eNs: examined lymph nodes.