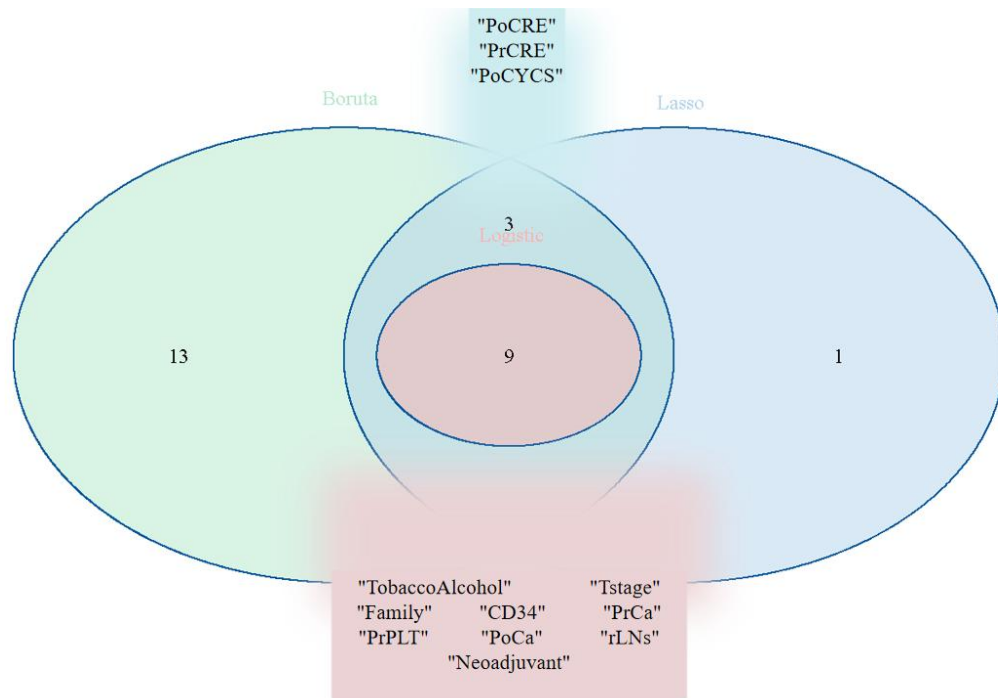




Supplementary Figure 1 Logistic regression, Boruta regression, and Lasso regression for feature selection using Venn diagrams. PoCa: Postoperative calcium ion concentration; PoCRE: Preoperative serum creatinine concentration; PrCRE: Preoperative serum creatinine concentration; PoCYCS: Postoperative C-reactive protein; rLNs: Positive lymph node count; PrPLT: Preoperative platelet concentration.

Table 1 Training set baseline data

	NAL, <i>n</i> = 1212	AL, <i>n</i> = 61	<i>P</i> value
Sex = Male (%)	732 (60.4)	44 (72.1)	0.089
TobaccoAlcohol (%)			0.01
No	878 (72.4)	34 (55.7)	
Tobacco	182 (15.0)	11 (18.0)	
Alcohol	15 (1.2)	1 (1.6)	
All	137 (11.3)	15 (24.6)	
Tstage (%)			<0.001
T1	136 (11.2)	2 (3.3)	
T2	329 (27.1)	13 (21.3)	

T3	685 (56.5)	32 (52.5)	
T4	62 (5.1)	14 (23.0)	
Nstage (%)			0.145
N0	206 (17.0)	9 (14.8)	
N1	883 (72.9)	41 (67.2)	
N2	123 (10.1)	11 (18.0)	
TISSUE (%)			<0.001
Adenocarcinoma	30 (2.5)	0 (0.0)	
Adenosquamouscarcinoma	0 (0.0)	4 (6.6)	
Others	1161 (95.8)	57 (93.4)	
Multi	1 (0.1)	0 (0.0)	
Neuroendocrine	10 (0.8)	0 (0.0)	
Mesenchyme	10 (0.8)	0 (0.0)	
Family = Yes (%)	28 (2.3)	8 (13.1)	<0.001
CD34 = Yes (%)	458 (37.8)	48 (78.7)	<0.001
FOBT = M1 (%)	909 (75.0)	47 (77.0)	0.834
Location = Others (%)	51 (4.2)	0 (0.0)	0.193
Age = High (%)	599 (49.4)	23 (37.7)	0.098
Neoadjuvant = Yes (%)	104 (8.6)	41 (67.2)	<0.001

NAL: Non-anastomotic leakage; AL: Anastomotic leakage.

Table 2 Validation set baseline data

	NAL, <i>n</i> = 520	AL, <i>n</i> = 25	<i>P</i> value
Sex = Male (%)	323 (62.1)	20 (80.0)	0.11
TobaccoAlcohol (%)			0.029
No	378 (72.7)	12 (48.0)	
Tobacco	76 (14.6)	7 (28.0)	
Alcohol	10 (1.9)	0 (0.0)	
All	56 (10.8)	6 (24.0)	

Tstage (%)			<0.001
T1	56 (10.8)	0 (0.0)	
T2	124 (23.8)	4 (16.0)	
T3	310 (59.6)	11 (44.0)	
T4	30 (5.8)	10 (40.0)	
Nstage (%)			0.144
N0	85 (16.3)	3 (12.0)	
N1	377 (72.5)	16 (64.0)	
N2	58 (11.2)	6 (24.0)	
TISSUE (%)			<0.001
Adenocarcinoma	6 (1.2)	0 (0.0)	
Adenosquamouscarcinoma	0 (0.0)	2 (8.0)	
Others	502 (96.5)	23 (92.0)	
Neuroendocrine	7 (1.3)	0 (0.0)	
Mesenchyme	5 (1.0)	0 (0.0)	
Family = Yes (%)	4 (0.8)	2 (8.0)	0.016
CD34 = Yes (%)	189 (36.3)	19 (76.0)	<0.001
FOBT = M1 (%)	367 (70.6)	17 (68.0)	0.959
Location = Others (%)	25 (4.8)	2 (8.0)	0.805
Age = High (%)	272 (52.3)	16 (64.0)	0.348
Neoadjuvant = Yes (%)	38 (7.3)	16 (64.0)	<0.001

NAL: Non-anastomotic leakage; AL: Anastomotic leakage.

Table 3 External validation baseline data

	NAL, <i>n</i> = 40	AL, <i>n</i> = 20	<i>P</i> value
Sex = Male (%)	16 (40.0)	8 (40.0)	1
TobaccoAlcohol (%)			0.037
No	15 (37.5)	5 (25.0)	

Tobacco	22 (55.0)	8 (40.0)	
Alcohol	1 (2.5)	5 (25.0)	
All	2 (5.0)	2 (10.0)	
Tstage (%)			0.007
T1	1 (2.5)	0 (0.0)	
T2	23 (57.5)	3 (15.0)	
T3	15 (37.5)	14 (70.0)	
T4	1 (2.5)	3 (15.0)	
Nstage (%)			0.649
N0	10 (25.0)	5 (25.0)	
N1	25 (62.5)	14 (70.0)	
N2	5 (12.5)	1 (5.0)	
TISSUE = Adenocarcinoma (%)	27 (67.5)	0 (0.0)	<0.001
Family = Yes (%)	10 (25.0)	14 (70.0)	0.002
CD34 = Yes (%)	26 (65.0)	14 (70.0)	0.923
FOBT = Yes (%)	27 (67.5)	13 (65.0)	1
Location = Others (%)	40 (100.0)	20 (100.0)	NA
Age = Yes (%)	29 (72.5)	9 (45.0)	0.072
Neoadjuvant = Yes (%)	2 (5.0)	6 (30.0)	0.022

NAL: Non-anastomotic leakage; AL: Anastomotic leakage.