

S5 Table. Clinicopathologic correlation between ground truth and prediction

Characteristic	True labels					Predicted labels				
	Luminal A	Luminal B	HER-2	TNBC	p-value	Luminal A	Luminal B	HER-2	TNBC	p-value
No. (% of total)	24 (26)	56 (61)	2 (2)	10 (11)		15 (16)	64 (70)	8 (9)	5 (5)	
Mean age (yr)	59.1	55.1	66.0	55.0	0.192	56.7	56.8	55.3	51.6	0.778
Mean tumor size (cm)	1.8	2.0	1.0	2.2	0.249	1.6	1.9	2.1	2.6	0.062 ^{a)}
Histology										
Ductal	14 (58)	46 (82)	1 (50)	7 (70)	0.088 ^{a)}	9 (60)	48 (75)	7 (88)	4 (80)	0.532
Nuclear grade										
High	0	21 (38)	0	6 (60)	< 0.001 ^{b)}	2 (13)	17 (27)	4 (50)	4 (80)	0.015 ^{b)}
Nodal involvement										
Yes	5 (21)	15 (27)	0	2 (20)	0.253	4 (27)	14 (22)	3 (38)	1 (20)	0.662
Multifocality										
Yes	6 (25)	9 (16)	0	1 (10)	0.723	1 (7)	13 (20)	0	2 (40)	0.172
LVI										
Yes	3 (13)	10 (18)	0	2 (20)	0.865	0	11 (17)	3 (38)	1 (20)	0.1
EIC										
Yes	7 (29)	15 (27)	1 (50)	4 (40)	0.39	5 (33)	19 (30)	3 (38)	0	0.491

Values are presented as number (%). The chi-square test was employed for binary variables, while the Kruskal-Wallis test was utilized for continuous variables to compare the proportion of characteristics among the molecular subtypes. The value in the bracket corresponds to the percentage within each label. AI, artificial intelligence; EIC, extensive intraductal component; HER-2, human epidermal growth factor receptor 2; LVI, lymphovascular invasion; TNBC, triple-negative breast cancer. ^{a)}p-values that are not statistically significant but lower than 0.1, ^{b)}Statistically significant p-values.