

Supplementary Material: Prognostic Model for Predicting Recurrence in Breast Cancer Patients in Saudi Arabia

Supplementary Table 1. Distribution of the dependent variable

Class	Frequency	Proportion (%)
0 (cancer did not recur)	319	78.2%
1 (cancer did recur)	89	21.8%
total	408	100%

Supplementary Table 2. Medical description of the nominal variables in the dataset. LIQ and LUQ stands for lower inner quadrant and lower outer quadrant respectively, while UIQ and UOQ stands for upper inner quadrant and upper outer quadrant respectively. NOS stands for not otherwise specified.

Name	Definition	Value	Count (Prop.(%))
Lymph Node	This checks if the tumour has spread to the lymph node or not.	Yes	250 (61.3)
		No	158 (38.7)
Surgery	This determines if the patient has had surgery or not.	Yes	398 (97.5)
		No	10 (2.5)
Radiotherapy	This checks if the patient had undergone radiotherapy or not.	Radiotherapy	153 (37.5)
		No Radiotherapy	255 (62.5)
Chemotherapy	This establishes whether or not the patient has had chemotherapy.	Chemotherapy	265 (65.0)
		No chemotherapy	143 (35.1)
Hormonal therapy	This determines if the patient has had hormone therapy or not.	Hormonal therapy	123 (30.2)
		No therapy	285 (69.9)
Stage	The stage is determined by the size and extent of the cancer tumor.	In situ	20 (4.9)
		Localized	122 (29.9)
		Regional	229 (56.1)
		Distant	37 (9.1)
Menopause status	Checks the menopausal status of the patient.	Premenopausal	237 (58.1)
		Postmenopausal	171 (41.9)
Site (Breast)	Location of the cancer tumor in the breast.	LIQ	18 (4.4)
		LOQ	36 (8.8)
		UIQ	36 (8.8)
		UOQ	116 (28.4)
		Central	22 (5.4)
		Overlapping lesion	90 (22.1)
		NOS	90 (22.1)

Supplementary Table 3. Medical description of the numerical variables in the dataset, including their distributions.

Name	Definition	Mean	Range	Median	(Q1, Q3)
Age at diagnosis	Age of the patient in years at the time of disease diagnosis	47.53	22-90	46	(39, 56)
Tumor size	The size of the cancer tumor in <i>mm</i>	35.48	0-160	30	(19.75, 46)

Supplementary Table 4. Comparison of patients' characteristics between the discovery and validation cohort.

Predictors	Discovery (N=285)	Validation (N=123)	Total (N=408)	p value
response				1
NO	223 (78.2%)	96 (78.0%)	319 (78.2%)	
YES	62 (21.8%)	27 (22.0%)	89 (21.8%)	
Age				0.141
Median	45.0	48.0	46.0	
(Range)	(24.0, 90.0)	(22.0, 90.0)	(22.0, 90.0)	
BS				0.498
CENTRAL	14 (4.9%)	8 (6.5%)	22 (5.4%)	
LIQ	13 (4.6%)	5 (4.1%)	18 (4.4%)	
LOQ	27 (9.5%)	9 (7.3%)	36 (8.8%)	
NOS	69 (24.2%)	21 (17.1%)	90 (22.1%)	
Overlapping lesion	62 (21.8%)	28 (22.8%)	90 (22.1%)	
UIQ	21 (7.4%)	15 (12.2%)	36 (8.8%)	
UOQ	79 (27.7%)	37 (30.1%)	116 (28.4%)	
TS				0.561
Median	30.0	30.0	30.0	
(Range)	(0.0, 160.0)	(0.0, 160.0)	(0.0, 160.0)	
Stage				0.191
DISTANT	21 (7.4%)	16 (13.0%)	37 (9.1%)	
IN SITU	14 (4.9%)	6 (4.9%)	20 (4.9%)	
LOCALIZED	82 (28.8%)	40 (32.5%)	122 (29.9%)	
REGIONAL	168 (58.9%)	61 (49.6%)	229 (56.1%)	
LN				0.525
No	107 (37.5%)	51 (41.5%)	158 (38.7%)	
Yes	178 (62.5%)	72 (58.5%)	250 (61.3%)	
MS				0.05
Postmenopausal	110 (38.6%)	61 (49.6%)	171 (41.9%)	
Premenopausal	175 (61.4%)	62 (50.4%)	237 (58.1%)	
Surgery				0.735
No	6 (2.1%)	4 (3.3%)	10 (2.5%)	
Yes	279 (97.9%)	119 (96.7%)	398 (97.5%)	
Radiotherapy				0.717
No	176 (61.8%)	79 (64.2%)	255 (62.5%)	
Yes	109 (38.2%)	44 (35.8%)	153 (37.5%)	
HT				0.299
No	204 (71.6%)	81 (65.9%)	285 (69.9%)	
Yes	81 (28.4%)	42 (34.1%)	123 (30.1%)	
Chemotherapy				0.716
No	102 (35.8%)	41 (33.3%)	143 (35.0%)	
Yes	183 (64.2%)	82 (66.7%)	265 (65.0%)	

Supplementary Table 5. Factors significantly associated with recurrence in multivariable logistic regression analysis.

Variables	Hazard Ratio (95% CI)	P-value
Age	0.97 (0.93, 1.00)	0.043
BS		
CENTRAL	1.0 (reference)	
Lower inner quadrant (LIQ)	1.00 (0.06, 16.78)	0.999
Lower outer quadrant (LOQ)	1.33 (0.15, 11.58)	0.794
NOS	1.83 (0.24, 14.21)	0.564
Overlapping lesion	1.39 (0.17, 11.34)	0.756
Upper inner quadrant (UIQ)	1.34 (0.13, 13.71)	0.803
Upper outer quadrant (UOQ)	1.71 (0.22, 13.51)	0.609
TS	1.01 (1.00, 1.02)	0.027
Stage		
DISTANT	1.0 (reference)	
IN SITU	0.05 (0.01, 0.54)	0.013
LOCALIZED	0.22 (0.07, 0.72)	0.012
REGIONAL	0.78 (0.25, 2.38)	0.662
LN		
No	1.0 (reference)	
Yes	0.55 (0.16, 1.93)	0.354
MS		
Postmenopausal	1.0 (reference)	
Premenopausal	0.75 (0.36, 1.57)	0.444
Surgery		
No	1.0 (reference)	
Yes	0.20 (0.07, 0.60)	0.004
Radiotherapy		
No	1.0 (reference)	
Yes	0.77 (0.42, 1.41)	0.396
HT		
No	1.0 (reference)	
Yes	10.01 (2.88, 34.73)	< 0.0001
Chemotherapy		
No	1.0 (reference)	
Yes	0.05 (0.02, 0.18)	< 0.0001