

Figure S1 The flowchart described the process of conducting the study.

Table S1 Hyperparametric search space and results

Parameters	Range	Result
RSF		
n tree	[200, 500]	200
m try	[3, 5]	3
nodesize	[15, 21]	21
CoxPH	None	None
DT		
alpha	[0.01, 0.1]	0.05
minbucket	[5, 25]	25
XGBoost		
n rounds	[100, 500]	482
max depth	[1, 5]	1
eta	[0.0001, 1]	0.5631

Table S2 Baseline characteristics of the included MBC cohorts

Characteristics	Total population (N=1,458)	Training cohort (N=1,035) [1]	Internal validation cohort (N=350) [2]	External validation cohort (N=73) [3]	P _(1 vs. 2)	P _(1 vs. 3)
Age					0.672	0.012
≤60	1,009 (69.2%)	706 (68.2%)	243 (69.4%)	60 (82.2%)		
>60	449 (30.8%)	329 (31.8%)	107 (30.6%)	13 (17.8%)		
Primary site					0.349	0.001
Others	1037 (71.1%)	739 (71.4%)	259 (74.0%)	39 (53.4%)		
Upper outer	421 (28.9%)	296 (28.6%)	91 (26.0%)	34 (46.6%)		
Histologic grade					0.646	<0.001
I/II	507 (34.8%)	348 (33.6%)	113 (32.3%)	46 (63.0%)		
III/IV	951 (65.2%)	687 (66.4%)	237 (67.7%)	27 (37.0%)		
T stage					0.482	<0.001
T1	174 (11.9%)	106 (10.2%)	36 (10.3%)	32 (43.8%)		
T2	587 (40.3%)	421 (40.7%)	130 (37.1%)	36 (49.3%)		
T3/T4	697 (47.8%)	508 (49.1%)	184 (52.6%)	5 (6.85%)		
N stage					0.843	<0.001
0	228 (15.6%)	144 (13.9%)	48 (13.7%)	36 (49.3%)		
1–3	611 (41.9%)	451 (43.6%)	144 (41.1%)	16 (21.9%)		
4–9	308 (21.1%)	216 (20.9%)	79 (22.6%)	13 (17.8%)		
≥10	311 (21.3%)	224 (21.6%)	79 (22.6%)	8 (11.0%)		

Table S2 (continued)

Table S2 (continued)

Characteristics	Total population (N=1,458)	Training cohort (N=1,035) [1]	Internal validation cohort (N=350) [2]	External validation cohort (N=73) [3]	P _(1 vs. 2)	P _(1 vs. 3)
Radiotherapy					0.530	0.026
No	759 (52.1%)	527 (50.9%)	185 (52.9%)	47 (64.4%)		
Yes	699 (47.9%)	508 (49.1%)	165 (47.1%)	26 (35.6%)		
Bone metastasis					0.295	0.017
No	538 (36.9%)	367 (35.5%)	135 (38.6%)	36 (49.3%)		
Yes	920 (63.1%)	668 (64.5%)	215 (61.4%)	37 (50.7%)		
Brain metastasis					0.311	<0.001
No	1379 (94.6%)	987 (95.4%)	329 (94.0%)	63 (86.3%)		
Yes	79 (5.42%)	48 (4.64%)	21 (6.00%)	10 (13.7%)		
Liver metastasis					0.855	0.006
No	1025 (70.3%)	721 (69.7%)	242 (69.1%)	62 (84.9%)		
Yes	433 (29.7%)	314 (30.3%)	108 (30.9%)	11 (15.1%)		
Lung metastasis					0.554	0.002
No	1010 (69.3%)	730 (70.5%)	241 (68.9%)	39 (53.4%)		
Yes	433 (29.7%)	305 (29.5%)	109 (31.1%)	34 (46.6%)		
Molecular subtypes					0.740	0.029
Luminal	629 (43.1%)	444 (42.9%)	143 (40.9%)	42 (57.5%)		
HER2 positive	544 (37.3%)	388 (37.5%)	139 (39.7%)	17 (23.3%)		
Triple negative	285 (19.5%)	203 (19.6%)	68 (19.4%)	14 (19.2%)		
Over survival event					0.203	0.584
No	568 (39.0%)	392 (37.9%)	146 (41.7%)	30 (41.1%)		
Yes	890 (61.0%)	643 (62.1%)	204 (58.3%)	43 (58.9%)		

Table S3 Characteristics of various metastatic sites based on molecular subtypes

	Total population (N=1,035)	Luminal (N=444)	HER2 positive (N=388)	TNBC (N=203)	P
Training cohort					
Bone metastasis					<0.001
No	367 (35.5%)	104 (23.4%)	158 (40.7%)	105 (51.7%)	
Yes	668 (64.5%)	340 (76.6%)	230 (59.3%)	98 (48.3%)	
Brain metastasis					0.017
No	987 (95.4%)	426 (95.9%)	375 (96.6%)	186 (91.6%)	
Yes	48 (4.64%)	18 (4.05%)	13 (3.35%)	17 (8.37%)	
Liver metastasis					<0.001
No	721 (69.7%)	348 (78.4%)	231 (59.5%)	142 (70.0%)	
Yes	314 (30.3%)	96 (21.6%)	157 (40.5%)	61 (30.0%)	
Lung metastasis					<0.001
No	730 (70.5%)	343 (77.3%)	278 (71.6%)	109 (53.7%)	
Yes	305 (29.5%)	101 (22.7%)	110 (28.4%)	94 (46.3%)	
Internal validation cohort					
Bone metastasis					<0.001
No	135 (38.6%)	29 (20.3%)	66 (47.5%)	40 (58.8%)	
Yes	215 (61.4%)	114 (79.7%)	73 (52.5%)	28 (41.2%)	
Brain metastasis					0.107
No	329 (94.0%)	139 (97.2%)	128 (92.1%)	62 (91.2%)	
Yes	21 (6.00%)	4 (2.80%)	11 (7.91%)	6 (8.82%)	
Liver metastasis					0.016
No	242 (69.1%)	110 (76.9%)	85 (61.2%)	47 (69.1%)	
Yes	108 (30.9%)	33 (23.1%)	54 (38.8%)	21 (30.9%)	
Lung metastasis					0.080
No	241 (68.9%)	106 (74.1%)	95 (68.3%)	40 (58.8%)	
Yes	109 (31.1%)	37 (25.9%)	44 (31.7%)	28 (41.2%)	
External validation cohort					
Bone metastasis					0.010
No	36 (49.3%)	17 (40.5%)	7 (41.2%)	12 (85.7%)	
Yes	37 (50.7%)	25 (59.5%)	10 (58.8%)	2 (14.3%)	
Brain metastasis					0.460
No	63 (86.3%)	38 (90.5%)	14 (82.4%)	11 (78.6%)	
Yes	10 (13.7%)	4 (9.52%)	3 (17.6%)	3 (21.4%)	
Liver metastasis					0.213
No	62 (84.9%)	34 (81.0%)	14 (82.4%)	14 (100%)	
Yes	11 (15.1%)	8 (19.0%)	3 (17.6%)	0 (0.00%)	
Lung metastasis					0.096
No	39 (53.4%)	26 (61.9%)	9 (52.9%)	4 (28.6%)	
Yes	34 (46.6%)	16 (38.1%)	8 (47.1%)	10 (71.4%)	

Table S4 Univariate and multivariate Cox-regression analyses of factors associated with overall survival in the training cohort

Characteristics	Total (N=1,035)	Univariable		Multivariable	
		HR (95% CI)	P value	HR (95% CI)	P value
Age, years					
≤60	706 (68.2%)	Reference		Reference	
>60	329 (31.8%)	1.38 (1.18–1.63)	<0.001	1.19 (1.01–1.41)	0.036
Primary site					
Other	739 (71.4%)	Reference			
Upper outer	296 (28.6%)	1.12 (0.95–1.33)	0.191		
Histologic grade					
I/II	348 (33.6%)	Reference		Reference	
III/IV	687 (66.4%)	1.66 (1.40–1.97)	<0.001	1.38 (1.15–1.66)	0.001
T stage					
T1	106 (10.2%)	Reference		Reference	
T2	421 (40.7%)	1.59 (1.16–2.17)	0.004	1.29 (0.94–1.77)	0.110
T3/T4	508 (49.1%)	2.10 (1.55–2.85)	<0.001	1.67 (1.22–2.28)	0.001
N stage					
0	144 (13.9%)	Reference		Reference	
1–3	451 (43.6%)	0.99 (0.78–1.27)	0.957	0.94 (0.74–1.21)	0.639
4–9	216(20.9%)	1.10 (0.84–1.44)	0.489	1.07 (0.82–1.41)	0.611
≥10	224(21.6%)	1.36 (1.05–1.77)	0.022	1.35 (1.03–1.77)	0.029
Radiotherapy					
No	527 (50.9%)	Reference		Reference	
Yes	508 (49.1%)	0.83 (0.71–0.96)	0.016	0.88 (0.75–1.03)	0.114
Bone metastasis					
No	367(35.5%)	Reference		Reference	
Yes	668 (64.5%)	0.79 (0.67–0.93)	0.004	1.23 (1.02–1.49)	0.031
Brain metastasis					
No	987 (95.4%)	Reference		Reference	
Yes	48 (4.6%)	2.51 (1.83–3.54)	<0.001	2.34 (1.68–3.25)	<0.001
Liver metastasis					
No	712 (69.7%)	Reference		Reference	
Yes	314 (30.3%)	1.33 (1.13–1.57)	<0.001	1.70 (1.42–2.04)	<0.001
Lung metastasis					
No	730 (70.5%)	Reference		Reference	
Yes	305 (29.5%)	1.57 (1.34–1.85)	<0.001	1.39 (1.15–1.68)	0.001
Molecular subtypes					
Luminal	444 (42.9%)	Reference		Reference	
HER2 positive	388 (37.5%)	0.67 (0.56–0.81)	<0.001	0.56 (0.46–0.68)	<0.001
TNBC	203 (19.6%)	2.89 (2.38–3.50)	<0.001	2.30 (1.87–2.84)	<0.001

Table S5 Prediction performance of the ML models

Subgroup	Model	C-Index	AUC (95%CI)			Brier score (95% CI)		
			1-year	3-year	5-year	1-year	3-year	5-year
Training cohort	RSF	0.723 (0.704–0.740)	0.826 (0.791–0.860)	0.805 (0.777–0.832)	0.775 (0.745–0.805)	0.093 (0.080–0.105)	0.174 (0.163–0.185)	0.193 (0.184–0.203)
	CoxPH	0.707 (0.687–0.726)	0.763 (0.719–0.808)	0.779 (0.750–0.808)	0.742 (0.710–0.773)	0.096 (0.083–0.109)	0.183 (0.173–0.194)	0.202 (0.191–0.213)
	XGBoost	0.707 (0.686–0.726)	0.763 (0.719–0.807)	0.779 (0.750–0.808)	0.741 (0.709–0.773)	0.096 (0.083–0.109)	0.183 (0.173–0.194)	0.202 (0.191–0.213)
	DT	0.680 (0.657–0.700)	0.731 (0.688–0.774)	0.748 (0.717–0.779)	0.730 (0.698–0.763)	0.102 (0.089–0.116)	0.191 (0.178–0.203)	0.205 (0.195–0.216)
Internal validation cohort	RSF	0.727 (0.693–0.761)	0.737 (0.654–0.820)	0.798 (0.751–0.846)	0.783 (0.731–0.835)	0.099 (0.077–0.122)	0.181 (0.162–0.199)	0.196 (0.180–0.212)
	CoxPH	0.723 (0.685–0.760)	0.723 (0.632–0.814)	0.792 (0.742–0.842)	0.781 (0.730–0.833)	0.099 (0.076–0.122)	0.181 (0.162–0.200)	0.193 (0.173–0.212)
	XGBoost	0.722 (0.684–0.760)	0.723 (0.632–0.814)	0.791 (0.741–0.841)	0.780 (0.729–0.832)	0.099 (0.076–0.122)	0.181 (0.162–0.200)	0.192 (0.173–0.212)
External validation cohort	DT	0.697 (0.661–0.736)	0.703 (0.619–0.787)	0.769 (0.718–0.820)	0.738 (0.683–0.793)	0.103 (0.079–0.126)	0.187 (0.166–0.207)	0.205 (0.187–0.223)
	RSF	0.685 (0.606–0.758)	0.857 (0.775–0.940)	0.824 (0.730–0.919)	0.798 (0.662–0.934)	0.029 (0.011–0.047)	0.159 (0.116–0.203)	0.227 (0.182–0.272)

Partial dependence survival profiles

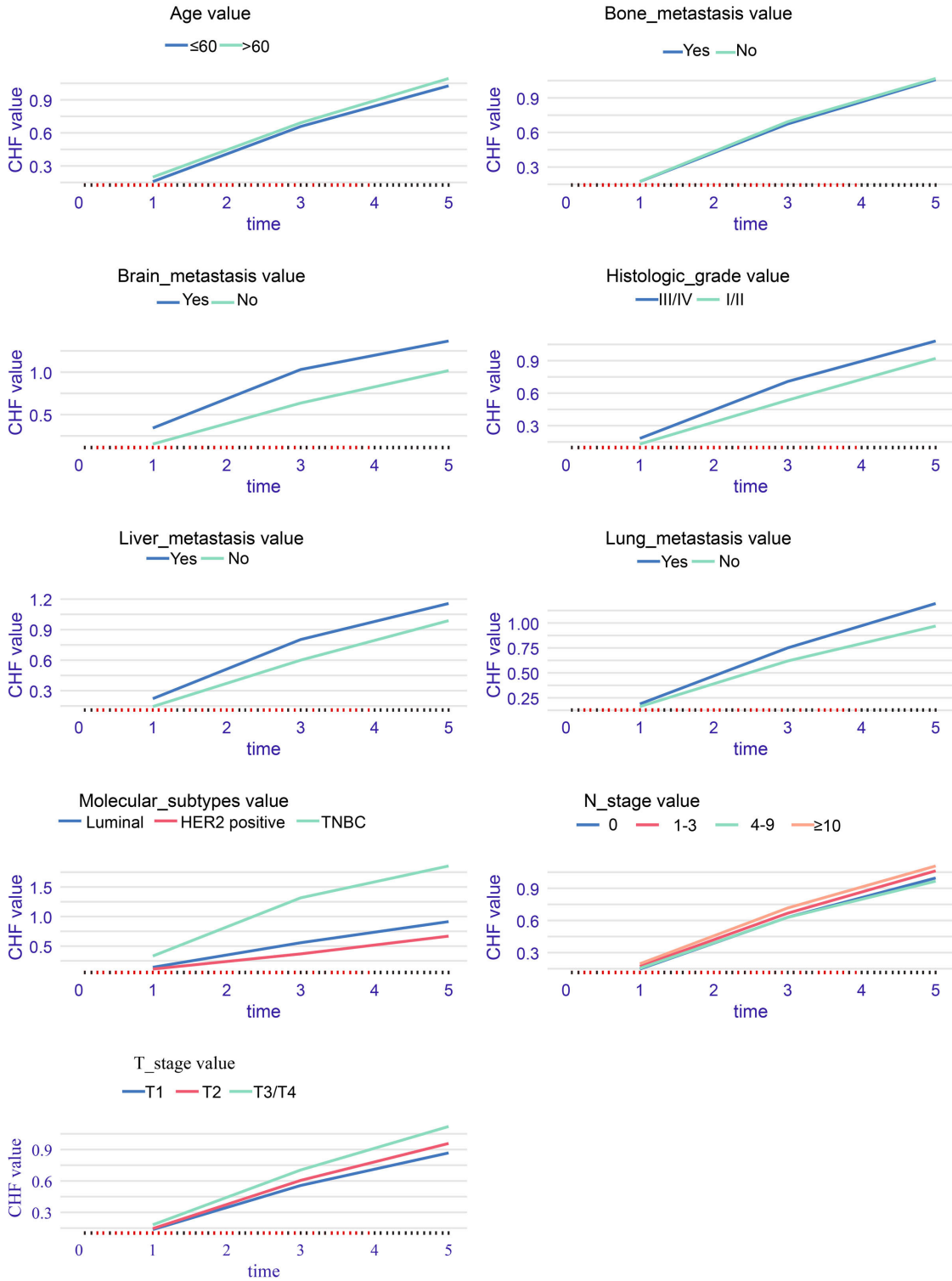


Figure S2 Partial dependence plot of SHAP model.