

Supplementary Table S7: Comparison of patient characteristics from dataset pool #1 and pool #2

Parameter		total		pool #1		pool #2		Odds Ratio	P-Value
Lymph node status	LNN	1,267	70.7%	751	85.7%	516	56.3%	4.67	< 0.001
	N+	526	29.3%	125	14.3%	401	43.7%		
Age	Age > 50	969	69.2%	508	68.0%	461	70.6%	0.89	n.s.
	Age ≤ 50	431	30.8%	239	32.0%	192	29.4%		
Tumor size	≤ 2 cm	160	30.4%	121	39.8%	39	17.5%	3.12	< 0.001
	> 2 cm	367	69.6%	183	60.2%	184	82.5%		
Grade	G3	423	30.4%	226	31.8%	197	28.8%	1.15	n.s.
	G1 & G2	970	69.6%	484	68.2%	486	71.2%		
PgR status	positive	1,396	71.9%	655	74.3%	741	70.0%	1.24	0.038
	negative	545	28.1%	227	25.7%	318	30.0%		
HER2 status	positive	124	6.4%	58	6.6%	66	6.2%	1.06	n.s.
	negative	1,817	93.6%	824	93.4%	993	93.8%		
Subtype (Hugh et al.)	Luminal A	1,052	54.2%	504	57.1%	548	51.7%	1.24	0.020
	Luminal B	889	45.8%	378	42.9%	511	48.3%		
Systemic treatment	Untreated	686	39.7%	504	57.1%	182	21.5%	4.87	< 0.001
	Endocrine Only	913	52.8%	359	40.7%	554	65.4%		
	Chemotherapy	130	7.5%	19	2.2%	111	13.1%		
Outcome variable	RFS	892	47.4%	323	38.9%	569	54.0%	0.54	< 0.001
	DMFS	991	52.6%	507	61.1%	484	46.0%		
OPG	low	1,134	58.4%	464	52.6%	670	63.3%	0.64	< 0.001
	high	807	41.6%	418	47.4%	389	36.7%		
Platform	U133A	1,517	78.2%	745	84.5%	772	72.9%	2.02	< 0.001
	U133Plus2.0	424	21.8%	137	15.5%	287	27.1%		