

Supplementary Table 1. Summary of public breast cancer microarray datasets used in the survival analysis

Accession No.	GSE1456	GSE2034	GSE4922	GSE7390	GSE10885	GSE222226	GSE24450	GSE25066	GSE53031	GSE58812	NKI
No. of patients	159	286	249	198	244	150	183	508	167	107	295
Assessable cases	159	253	213	158	242	149	183	198	167	107	295
Date of study	1994–1996	1980–1995	1987–1989/2000–2002	1980–1998	N/A	N/A	1997–2004	N/A	1996–2010	1998–2007	1984–1995
Platforms*	GPL96, GPL97	GPL96	GPL96, GPL97	GPL96	GPL885, GPL887	GPL1708, GPL4133	GPL6947	GLP96	GPL13667	GPL570	Agilent 25K Chip
Country	Sweden	USA	Singapore/Sweden	France/Sweden/UK	USA	USA	Finland	USA	Italy	France	Netherlands
Age (yr) [†]	N/A	N/A	NA	46 (24–60)	53 (24–88)	48 (31–65)	N/A	48 (24–72)	36 (29–47)	57 (28–84)	44 (26–53)
Tumor size	N/A	N/A	Y	Y	Y	Y	N/A	N/A	Y	N/A	Y
Histologic grade	Y	N/A	Y	Y	Y	Y	N/A	Y	Y	N/A	Y
Lymph node status	N/A	Y	Y	Y	N/A	N/A	N/A	N/A	NA	N/A	Y
ER status	Y	Y	Y	Y	Y	Y	N/A	Y	Y	Y	Y
Chemotherapy	N/A	N/A	N/A	N/A	N/A	Y	N/A	N/A	N/A	N/A	Y
OS range (mo)	2.2–101.9	N/A	N/A	4.9–303.6	1–96.0	9.5–82.0	22.4–120.0	N/A	N/A	1–171.7	1.0–220.1
DFS range (mo)	2.2–101.9	2.0–171.0	0–153.0	4–231.4	1–96.0	6.0–84.0	0.8–60.0	1.7–88.3	2.0–156.0	1–171.7	1.0–220.1

NA = not applicable; Y = yes; ER = estrogen receptor; OS = overall survival; DFS = disease-free survival.

*Platforms: GPL96: [HG-U133A] Affymetrix Human Genome U133A Array; GPL97: [HG-U133B] Affymetrix Human Genome U133B Array; GPL570: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array; GPL885: Agilent-011521 Human 1A Microarray G4110A; GPL887: Agilent-012097 Human 1A Microarray (V2) G4110B; GPL1708: Agilent-012391 Whole Human Genome Oligo Microarray G4112A; GPL4133: Agilent-014850 Whole Human Genome Microarray 4x44K G4112F; GPL6947: Illumina HumanHT-12 V3.0 Expression Beadchip; GPL8269: Agilent UNC PerouLab 244K Custom Human Array version 5; GPL13667: [HG-U219] Affymetrix Human Genome U219 Array; †Median (range).