

	GSE1456	GSE1992	GSE4335	Vijver	GSE3143	Merged-ComBat*
GSE1456	NA	2.01(0.95-4.26) p=0.06	2.50(1.16-5.37) p= <b>0.017</b>	6.55(2.5-17.19) p= <b>1e-04</b>	3.33(1.47-7.53) p= <b>0.003</b>	7.12(2.71-18.66) p= <b>6.5e-05</b>
GSE1992	3.06(1.22-7.66) p= <b>0.01</b>	NA	3.44(1.49-7.96) p= <b>0.0032</b>	4.82(1.81-12.81) p= <b>0.001</b>	5.08(1.9-13.57) p= <b>0.001</b>	4.22(1.69-10.54) p= <b>0.002</b>
GSE4335	1.78(0.86-3.66) p=0.11	1.59(0.7-3.62) p=0.26	NA	2.84(1.3-6.21) p= <b>p=0.008</b>	2.27(1.07-4.8) p= <b>0.03</b>	2.53(1.19-5.35) p= <b>0.015</b>
Vijver	5.96(3.14-11.32) p= <b>4.8e-08</b>	3.99(2.37-6.74) p= <b>2.02e-07</b>	3.38(2.05-5.57) p= <b>1.7e-06</b>	NA	3.43(2.07-5.68) p= <b>1.83e-06</b>	6.20(3.40-11.30) p= <b>2.5e-09</b>
GSE3143	p $\geq$ 0.05	p $\geq$ 0.05	p $\geq$ 0.05	1.99(1.08-3.64) p= <b>0.02</b>	NA	2.13(1.17-3.87) p= <b>0.012</b>

**Table S1: Cross-data set performance of breast cancer predictors trained on the individual and combined data sets (adjusted by ComBat) with respect to OS.** Significant HR ( $p < 0.05$ ) are shown in bold. The training sets are listed in the column header and the testing sets are indicated in the row header of the table. \* indicates that the predictor was trained from all data sets except the testing set. NA stands for Not Available.