

**Supplementary data 1. Clinical characteristics of the datasets included in the survival analysis(19).**

Dataset	Platform	Age	Sex (% male)	% of never smokers	Histology (% A/S/L)	Stage (% 1/2/3/4)
<a href="#">GSE4573</a> (28)	<a href="#">GPL96</a>	67±9.8	63%	3.10%	0/100/0	56/26/18/0
<a href="#">GSE14814</a> (24)	<a href="#">GPL96</a>	62±8.5	74%	-	31/58/11	50/50/0/0
<a href="#">GSE8894</a> (27)	<a href="#">GPL570</a>	62±10	75%	-	46/54/0	-
<a href="#">GSE19188</a> (25)	<a href="#">GPL570</a>	-	75%	-	49/30/21	-
<a href="#">GSE3141</a> (29)	<a href="#">GPL570</a>	-	-	-	52/48/0	-
<a href="#">GSE31210</a> (21)	<a href="#">GPL570</a>	60±8.1	47%	50%	100/0/0	74/26/0/0
CaArray (26)	<a href="#">GPL96</a>	64±10	51%	14%	-	-
TCGA(22)	<a href="#">GPL3921</a>	66±9.3	67%	7.50%	0/100/0	-
<a href="#">GSE29013</a> (23)	<a href="#">GPL570</a>	64±8.7	69%	3.60%	55/45/0	44/25/31/0
<a href="#">GSE37745</a> (20)	<a href="#">GPL570</a>	64±9.2	55%	-	54/37/12	66/18/14/2
<u>Entire Data Sets</u>		64±10	58%	17.80%	50/45/5	63/27/10/1