Supplementary table 1 Parameter setting range of each model used in this model

Learner	Parameter	Parameter type	Parameter range
	mtry	Integer	1,15
	nodesize	Discrete	3, 5, 8, 10,15,18,20
	ntree	Discrete	500, 1000, 1500, 2000
Random forest	nodedepth	Integer	5, 20
	mstop	Integer	1e2, 1e3
glmboost	nu	Discrete	0.05, 0.1, 0.3, 0.5, 0.8,1
coxboost	stepno	Integer	lower=50, upper = 200
	alpha	Numeric	lower = 0, upper = 1
Elastic net	s	Numeric	lower=0.001, upper=30
Ridge	s	Numeric	lower = 0, upper = 20
Lasso	s	Numeric	lower = 0, upper = 20

Supplementary table 2 Univariate Cox proportional hazards regression model on the overall survival of patients with breast cancer in METABRIC cohort

Genes	Coefficient	HR	LCI	UCI	P Value
ABAT	(0.298)	0.742	0.666	0.827	<0.001
ADHFE1	(0.294)	0.745	0.666	0.833	<0.001
ADRA2A	(0.400)	0.670	0.575	0.781	<0.001
AK3	(0.438)	0.645	0.561	0.742	<0.001
ANGPT2	0.350	1.419	1.257	1.603	<0.001
ANKAR	(0.650)	0.522	0.410	0.665	<0.001
APLN	0.406	1.500	1.284	1.754	<0.001
ARHGEF12	(0.457)	0.633	0.536	0.749	<0.001
ARL6IP5	(0.360)	0.698	0.611	0.797	<0.001
ASPM	0.196	1.217	1.131	1.309	<0.001
AURKA	0.241	1.273	1.184	1.369	<0.001
AURKB	0.228	1.256	1.163	1.356	<0.001
BCL2	(0.190)	0.827	0.776	0.882	<0.001
BIRC5	0.169	1.184	1.109	1.264	<0.001
BUB1	0.246	1.279	1.173	1.395	<0.001
C5	(0.337)	0.714	0.629	0.811	<0.001
CA9	0.213	1.237	1.147	1.335	<0.001
CATSPERB	0.296	1.345	1.202	1.505	<0.001
CBX2	0.135	1.144	1.089	1.202	<0.001
CBX7	(0.327)	0.721	0.639	0.814	<0.001
CCDC25	(0.349)	0.705	0.620	0.802	<0.001
CCDC74A	(0.139)	0.870	0.826	0.916	<0.001
CCNA2	0.263	1.300	1.188	1.423	<0.001
CCNB2	0.209	1.232	1.152	1.318	<0.001

CCNE2	0.255	1.291	1.182	1.410	<0.001
CCNYL1	0.418	1.519	1.296	1.781	<0.001
CD1C	(0.357)	0.700	0.609	0.804	<0.001
CDC20	0.169	1.184	1.121	1.251	<0.001
CDCA3	0.209	1.232	1.137	1.335	<0.001
CDCA5	0.240	1.271	1.190	1.358	<0.001
CDCA8	0.236	1.267	1.158	1.385	<0.001
CDKN3	0.235	1.265	1.162	1.378	<0.001
CENPA	0.226	1.254	1.154	1.361	<0.001
CENPE	0.317	1.374	1.239	1.523	<0.001
CENPF	0.212	1.237	1.143	1.338	<0.001
CENPI	0.754	2.126	1.599	2.826	<0.001
CENPL	0.437	1.548	1.338	1.791	<0.001
CENPO	0.606	1.832	1.510	2.224	<0.001
CEP55	0.231	1.260	1.157	1.372	<0.001
CFL1	0.559	1.749	1.485	2.060	<0.001
CKAP2L	0.317	1.373	1.244	1.516	<0.001
CLIC6	(0.106)	0.900	0.872	0.928	<0.001
COL14A1	(0.289)	0.749	0.671	0.837	<0.001
COL17A1	(0.175)	0.839	0.786	0.896	<0.001
COL4A1	0.304	1.355	1.221	1.505	<0.001
CPT1A	0.445	1.561	1.365	1.785	<0.001
CS	0.409	1.505	1.287	1.759	<0.001
CYP4F22	(0.176)	0.839	0.788	0.892	<0.001
DHRS12	(0.529)	0.589	0.483	0.718	<0.001
DYNLRB2	(0.245)	0.782	0.720	0.850	<0.001
EIF2S2	0.401	1.494	1.278	1.747	<0.001
ELF3	0.224	1.251	1.147	1.364	<0.001
ELMO2	0.468	1.597	1.334	1.913	<0.001
ENC1	0.319	1.376	1.241	1.525	<0.001
ERCC6L	0.596	1.814	1.453	2.266	<0.001
ESPL1	0.551	1.735	1.447	2.081	<0.001
EXO1	0.236	1.267	1.166	1.376	<0.001
EXOSC2	0.531	1.700	1.401	2.063	<0.001
FAM83D	0.272	1.313	1.211	1.424	<0.001
FAM91A1	0.409	1.505	1.298	1.744	<0.001
FANCD2	0.341	1.407	1.253	1.579	<0.001
FCER1A	(0.208)	0.812	0.762	0.865	<0.001
FEN1	0.292	1.339	1.220	1.471	<0.001
FGD3	(0.170)	0.844	0.798	0.892	<0.001
FLT3	(0.472)	0.624	0.525	0.742	<0.001
FOXM1	0.220	1.247	1.150	1.351	<0.001
GAL3ST4	0.542	1.720	1.403	2.107	<0.001
GLA	(0.207)	0.813	0.752	0.879	<0.001

GPI	0.335	1.398	1.248	1.566	<0.001
GPRC5A	0.205	1.228	1.133	1.330	<0.001
GRB7	0.131	1.140	1.085	1.198	<0.001
GREB1	(0.267)	0.766	0.691	0.850	<0.001
GSDMB	0.149	1.161	1.096	1.229	<0.001
GSK3B	0.359	1.432	1.278	1.604	<0.001
GSTK1	(0.394)	0.674	0.590	0.770	<0.001
GSTM2	(0.114)	0.892	0.853	0.932	<0.001
GSTM3	(0.188)	0.829	0.778	0.882	<0.001
GSTM4	(0.463)	0.629	0.532	0.745	<0.001
GTSE1	0.381	1.464	1.298	1.651	<0.001
HJURP	0.263	1.300	1.188	1.424	<0.001
HS3ST1	0.473	1.605	1.348	1.910	<0.001
HSP90AB1	0.370	1.448	1.270	1.651	<0.001
IFT57	(0.517)	0.596	0.489	0.728	<0.001
IGFBP4	(0.193)	0.825	0.770	0.883	<0.001
IQUB	(0.856)	0.425	0.310	0.582	<0.001
IRS1	(0.236)	0.790	0.726	0.859	<0.001
KCNN2	0.724	2.063	1.575	2.702	<0.001
KIF20A	0.279	1.321	1.219	1.432	<0.001
KIF23	0.324	1.382	1.238	1.543	<0.001
KIF2C	0.239	1.270	1.169	1.380	<0.001
KIF4A	0.380	1.462	1.304	1.639	<0.001
KIFC1	0.230	1.258	1.157	1.368	<0.001
LAMA3	(0.188)	0.828	0.770	0.891	<0.001
LARP1	0.613	1.846	1.539	2.216	<0.001
LDHA	0.324	1.382	1.226	1.558	<0.001
LMNB1	0.505	1.657	1.361	2.017	<0.001
LSG1	0.526	1.692	1.416	2.023	<0.001
LSR	0.396	1.485	1.301	1.696	<0.001
ME3	(0.247)	0.781	0.714	0.856	<0.001
MELK	0.216	1.241	1.159	1.328	<0.001
MFAP4	(0.142)	0.868	0.823	0.916	<0.001
MYL5	(0.308)	0.735	0.662	0.817	<0.001
N4BP2L1	(0.441)	0.643	0.546	0.757	<0.001
NACC2	0.410	1.507	1.304	1.742	<0.001
NCAPG	0.239	1.270	1.169	1.379	<0.001
NCBP2	0.405	1.500	1.290	1.744	<0.001
NDRG1	0.192	1.211	1.134	1.294	<0.001
NME5	(0.213)	0.808	0.747	0.874	<0.001
NUSAP1	0.250	1.284	1.184	1.392	<0.001
OGN	(0.208)	0.812	0.756	0.874	<0.001
OIP5	0.269	1.309	1.188	1.441	<0.001
OMD	(0.239)	0.787	0.727	0.853	<0.001

OPRL1	(0.395)	0.674	0.581	0.782	<0.001
PA2G4	0.374	1.454	1.265	1.671	<0.001
PARP3	(0.348)	0.706	0.633	0.787	<0.001
PCDH17	0.233	1.262	1.156	1.378	<0.001
PFKL	0.398	1.488	1.285	1.724	<0.001
PGR	(0.206)	0.814	0.758	0.873	<0.001
PHC3	0.654	1.922	1.496	2.470	<0.001
PHYHD1	(0.252)	0.777	0.707	0.854	<0.001
PIGV	(0.480)	0.619	0.535	0.715	<0.001
PKMYT1	0.381	1.464	1.313	1.631	<0.001
PLK1	0.531	1.701	1.445	2.003	<0.001
PLK4	0.332	1.394	1.242	1.564	<0.001
POLQ	0.284	1.328	1.202	1.467	<0.001
POLR3GL	(0.370)	0.690	0.605	0.788	<0.001
PPIL3	(0.500)	0.606	0.521	0.706	<0.001
PRC1	0.238	1.268	1.179	1.365	<0.001
PREX1	(0.265)	0.767	0.696	0.845	<0.001
PRR11	0.325	1.384	1.229	1.557	<0.001
PSMD3	0.252	1.286	1.180	1.402	<0.001
PTTG1	0.230	1.259	1.175	1.349	<0.001
PTTG3P	0.217	1.242	1.155	1.335	<0.001
RACGAP1	0.376	1.456	1.316	1.610	<0.001
RAI2	(0.323)	0.724	0.645	0.813	<0.001
RALGAPB	0.482	1.619	1.377	1.904	<0.001
RBBP8	(0.245)	0.782	0.723	0.847	<0.001
RLN2	(0.398)	0.672	0.586	0.770	<0.001
RPA2	(0.483)	0.617	0.517	0.735	<0.001
RRM2	0.386	1.472	1.291	1.678	<0.001
RSBN1	(0.482)	0.618	0.520	0.733	<0.001
S100P	0.086	1.090	1.059	1.122	<0.001
SENP5	0.491	1.634	1.354	1.972	<0.001
SERPINE1	0.248	1.281	1.172	1.401	<0.001
SHMT2	0.359	1.432	1.275	1.609	<0.001
SLC16A3	0.202	1.224	1.138	1.318	<0.001
SLC25A1	0.298	1.347	1.203	1.508	<0.001
SLC38A10	0.294	1.342	1.200	1.502	<0.001
SLC7A2	(0.099)	0.905	0.875	0.937	<0.001
SPATA18	(0.248)	0.780	0.714	0.853	<0.001
SPATA4	(0.792)	0.453	0.335	0.611	<0.001
SQLE	0.190	1.209	1.126	1.298	<0.001
STARD3	0.204	1.226	1.133	1.328	<0.001
STAT5A	(0.321)	0.725	0.647	0.813	<0.001
STAT5B	(0.428)	0.652	0.565	0.752	<0.001
STIP1	0.421	1.524	1.353	1.716	<0.001

SUSD3	(0.142)	0.868	0.833	0.905	<0.001
SYTL1	(0.360)	0.698	0.610	0.798	<0.001
TIMELESS	0.336	1.399	1.243	1.575	<0.001
TK1	0.218	1.244	1.155	1.340	<0.001
TMEM101	(0.196)	0.822	0.766	0.882	<0.001
TMEM132A	0.201	1.223	1.133	1.319	<0.001
TMEM26	(0.192)	0.826	0.772	0.883	<0.001
TOP2A	0.177	1.194	1.123	1.268	<0.001
TPX2	0.282	1.326	1.223	1.438	<0.001
TRIB3	0.242	1.273	1.164	1.393	<0.001
TRIM4	(0.377)	0.686	0.607	0.775	<0.001
TRIP13	0.203	1.225	1.138	1.319	<0.001
TROAP	0.319	1.376	1.260	1.502	<0.001
TTK	0.200	1.221	1.130	1.320	<0.001
TUBA1B	0.483	1.620	1.387	1.893	<0.001
TUBA3D	(0.139)	0.870	0.828	0.915	<0.001
TUBA3E	(0.293)	0.746	0.671	0.829	<0.001
UBE2C	0.189	1.208	1.144	1.275	<0.001
UHRF1	0.248	1.282	1.187	1.384	<0.001
USP30	0.786	2.194	1.697	2.838	<0.001
UTP23	0.392	1.481	1.297	1.691	<0.001
VEGFA	0.281	1.325	1.212	1.447	<0.001
WDR19	(0.347)	0.707	0.625	0.799	<0.001
WNT7B	0.287	1.333	1.193	1.488	<0.001
ZBTB4	(0.368)	0.692	0.601	0.797	<0.001
ZCCHC24	(0.249)	0.780	0.710	0.856	<0.001
ZIC2	0.176	1.192	1.116	1.273	<0.001
ZNF214	(0.788)	0.455	0.334	0.618	<0.001
ZWINT	0.285	1.329	1.208	1.462	<0.001

Abbreviations: HR, hazard ratio; LCI, lower limit of confidence interval; UCI, upper limit of confidence interval.

Supplementary table 3 Differences between ABAT high expression group and ABAT low expression group

Variables	ABAT high expression group (N=1068)	ABAT low expression group (N=1068)	P value
ER			
negative	46 (4.3%)	394 (36.9%)	<0.001
positive	984 (92.1%)	524 (49.1%)	
Missing	38 (3.6%)	150 (14.0%)	
Size			
Mean (SD)	2.53 (1.36)	2.72 (1.72)	0.005
Median [Min, Max]	2.20 [0, 18.0]	2.30 [0, 18.2]	

Missing	30 (2.8%)	134 (12.5%)	
Node status			
Mean (SD)	0.444 (0.497)	0.513 (0.500)	0.002
Median [Min, Max]	0 [0, 1.00]	1.00 [0, 1.00]	
Missing	19 (1.8%)	125 (11.7%)	
Age (years)			
Mean (SD)	63.2 (12.5)	59.0 (13.2)	<0.001
Median [Min, Max]	64.3 [26.4, 92.1]	59.8 [21.9, 96.3]	
Missing	2 (0.2%)	11 (1.0%)	
Grade			
Mean (SD)	2.20 (0.657)	2.64 (0.556)	<0.001
Median [Min, Max]	2.00 [1.00, 3.00]	3.00 [1.00, 3.00]	
Missing	75 (7.0%)	159 (14.9%)	
OS time (Month)			
Mean (SD)	105 (59.1)	91.3 (59.9)	<0.001
Median [Min, Max]	99.1 [0.100, 307]	77.7 [0.267, 306]	
Missing	20 (1.9%)	127 (11.9%)	

Supplementary table 4. Comparison between CBioExlorer and other web-based tool for exploring the biomarkers of human cancers.

	CBioEx plorer	GEPIA/G EPIA2	PROGg eneV2	SurvEx press	KM_P lotter	UAL CAN	GSC ALite	CAS- viewer	Onco Lnc	CaP SSA	LO Gpc	PRE COG	cBioP ortal
Data													
Sample size	47,210	10,588	28,503	39,325	14912	7,233	10,55 8	10558	8,616	10,2 06	31,3 10	19,16 8	
Disease type	45	33	27	26	21	33	33	33	21	27	27	39	32
Data set	268	33	193	225	45	35	63	63	21	28	209	165	>100
Availability of Curated data	Yes	NO	No	No	No	No	No	No	No	No	No	Yes	Yes
Normalized gene expression data	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes
Customized data input and analysis	Yes	Yes	No	No	No	No	No	No	No	Yes	No	No	No
Features													
Dimensionality reduction	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No
Cross-validation or nested cross-validation	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Construction and evaluation of Predictive model	Yes	No	No	No	No	No	No	No	No	Yes	No	No	No
Nomogram	Yes	No	No	No	No	No	No	No	No	No	No	No	No
KM curve	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Optimal cutoff	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Univariate CoxPH	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No
Multivariate CoxPH	Yes	No	Yes	No	No	No	No	No	No	No	No	No	No

|													
Time-dependent ROC analysis	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Correlation with clinical features	Yes	No	No	No	No	No	No	No	No	Yes	No	No	No
Differential expression analysis	Yes	Yes	No	No	No	Yes	Yes	Yes	No	Yes	No	No	No
Correlation with other genes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No
Immune cell infiltration analysis	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No
Stemness score	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Multi-omics analysis	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No
Pan-cancer analysis	No	Yes	No	No	No	No	No	No	Yes	No	No	No	Yes
Enrichment analysis													
GSEA	Yes	No	No	No	No	No	No	No	No	No	No	No	No
ORA	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Gene ontology	Yes	No	No	No	No	No	No	No	No	No	No	No	No
KEGG	Yes	No	No	No	No	No	No	No	No	No	No	No	No
MsigDB	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Reactome pathway	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Software													
Log in required	No	No	No	No	No	No	No	No	No	No	No	Yes	No
Web interface	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standalone application	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Interactive results	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Availability of source code	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Result download	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
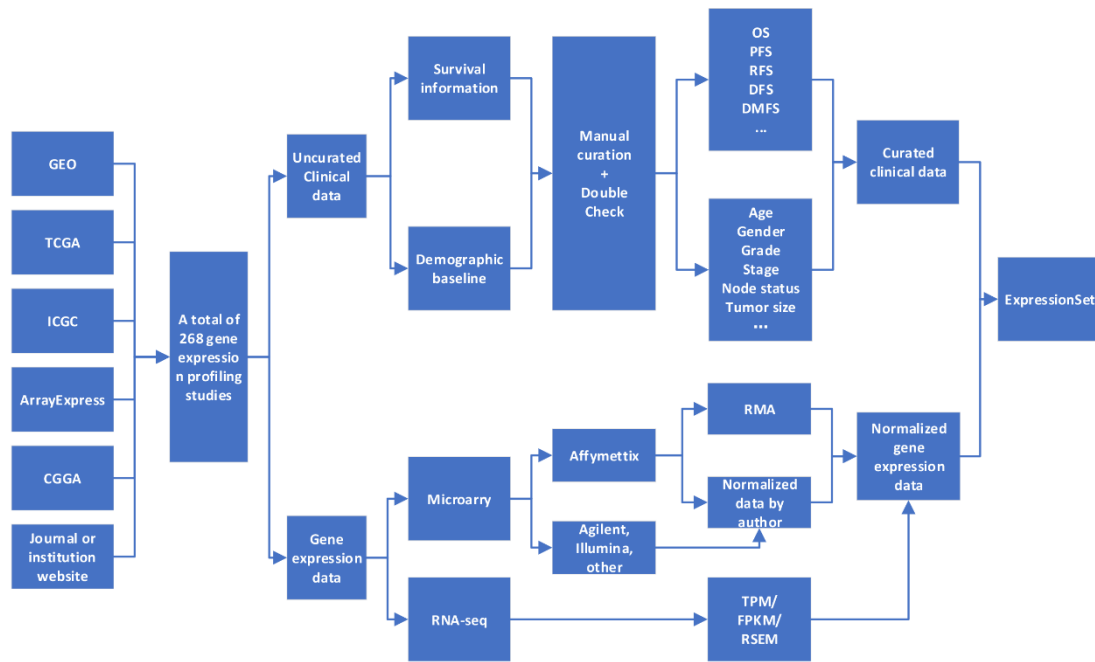
Supplementary figure 1. Flow chart of curation of public gene expression studies integrated in CuratedCancerPrognosisData
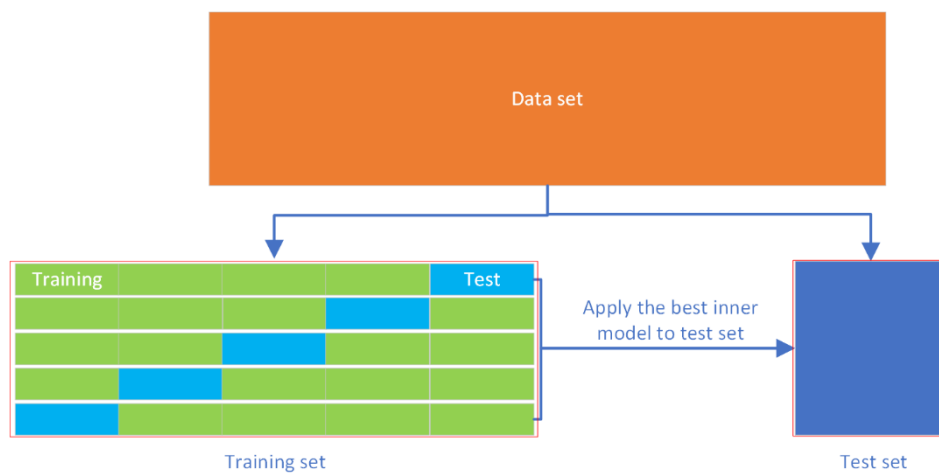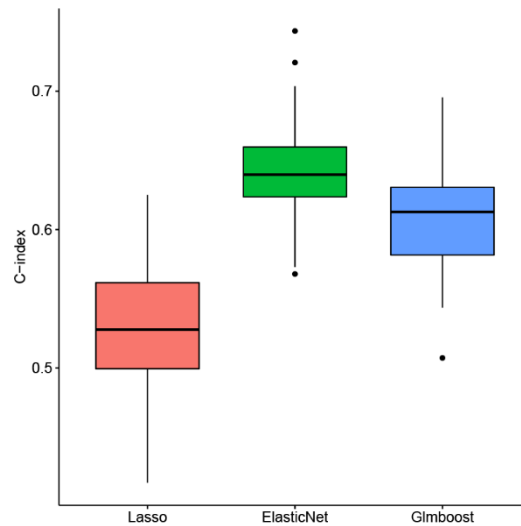

Supplementary figure 2. Workflow of cross validation
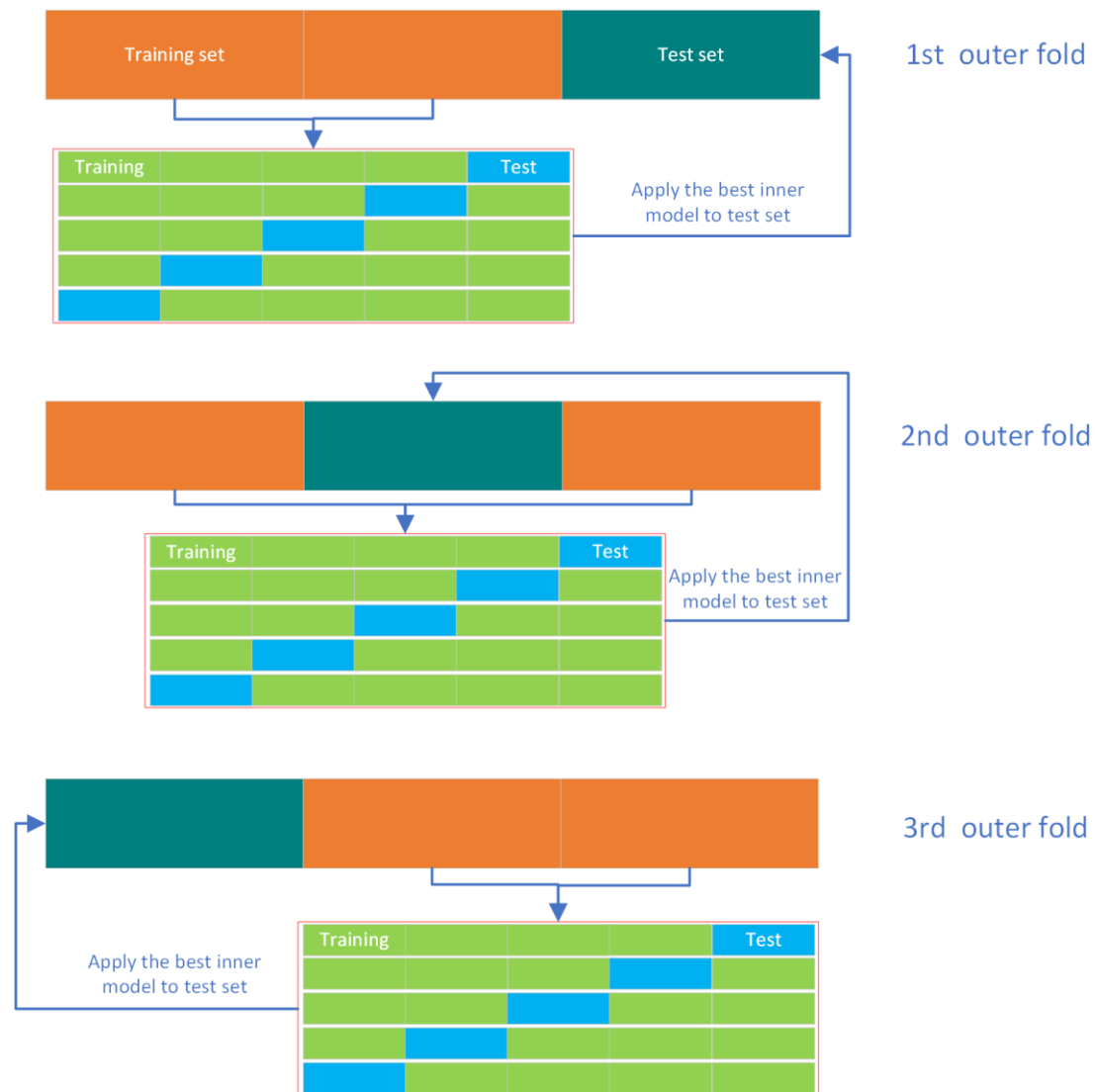
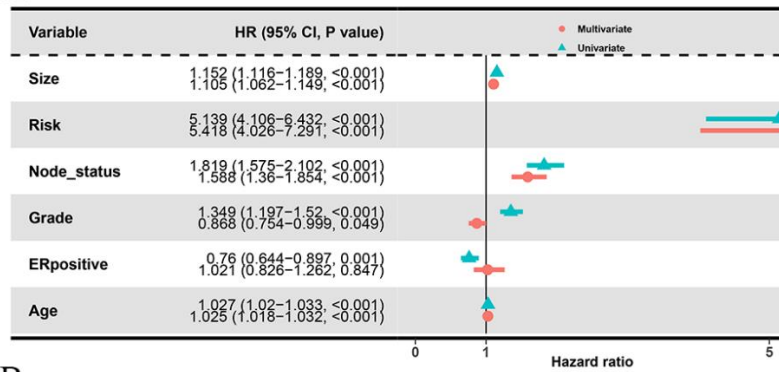**Supplementary figure 3.** Workflow of nested cross validation
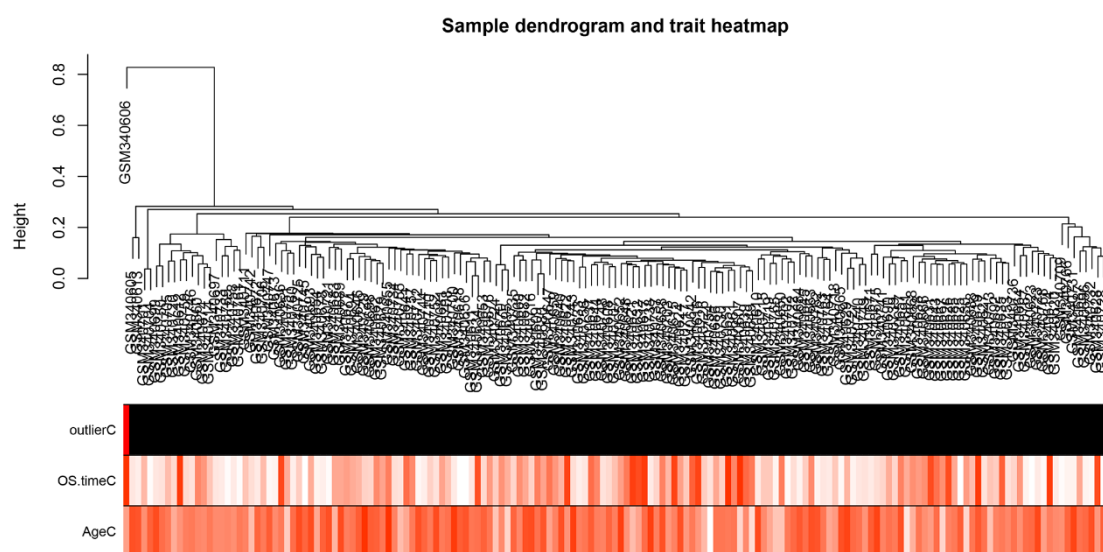
Supplementary figure 4. Detection of outliers based on sample dendrogram.
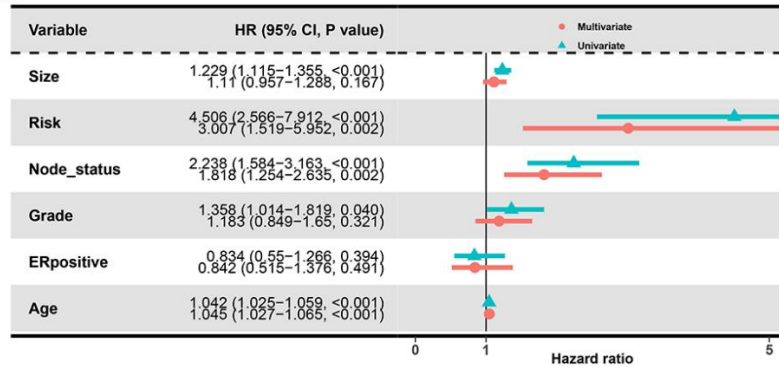
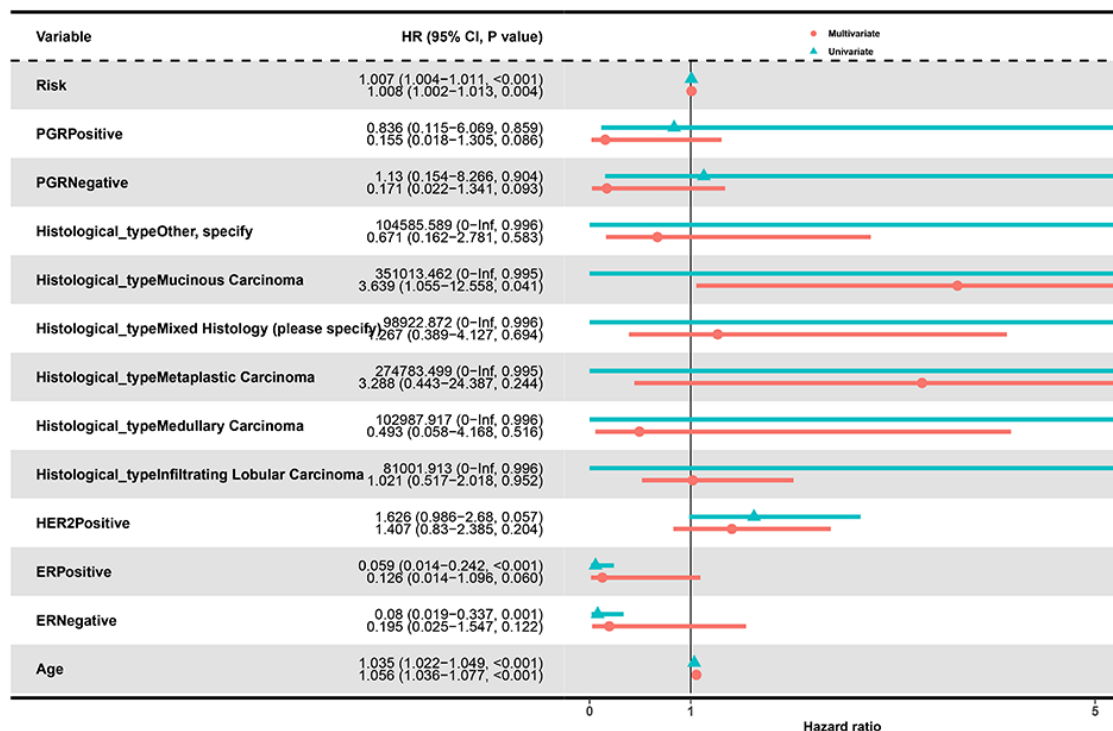**Supplementary figure 5.** Comparison of C-index of three survival learners based on cross validation and bootstraps.
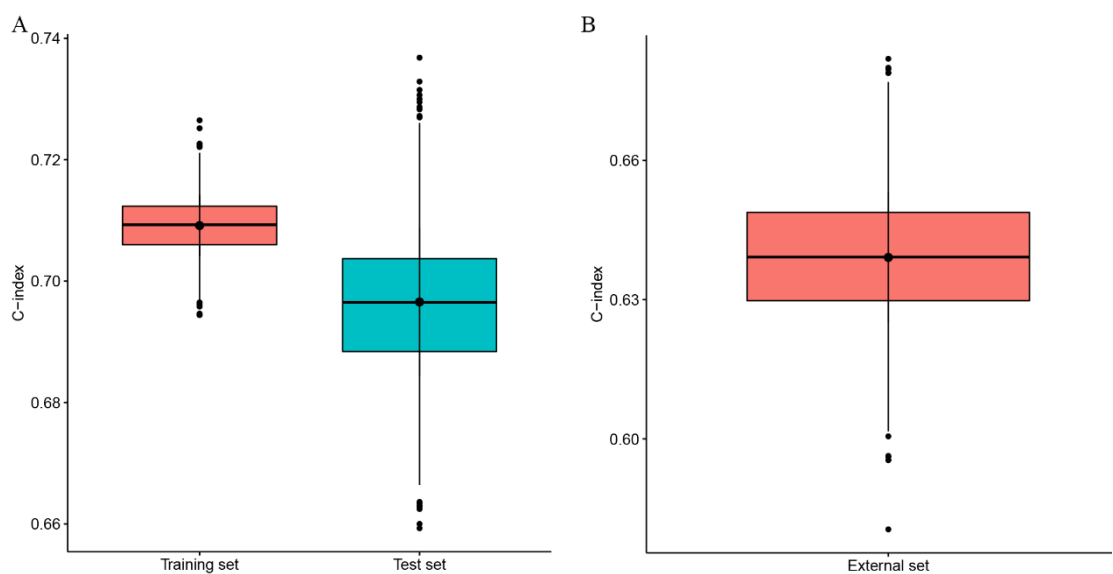

Supplementary figure 6. Cox proportion hazards regression model identifying independent prognostication role of the risk score in the training set (A) and test set (B).

Supplementary figure 7. Cox proportion hazards regression model identifying independent prognostication role of the risk score in the validation set.

Supplementary figure 8. Internally (A) and externally (B) validation of the nomogram.
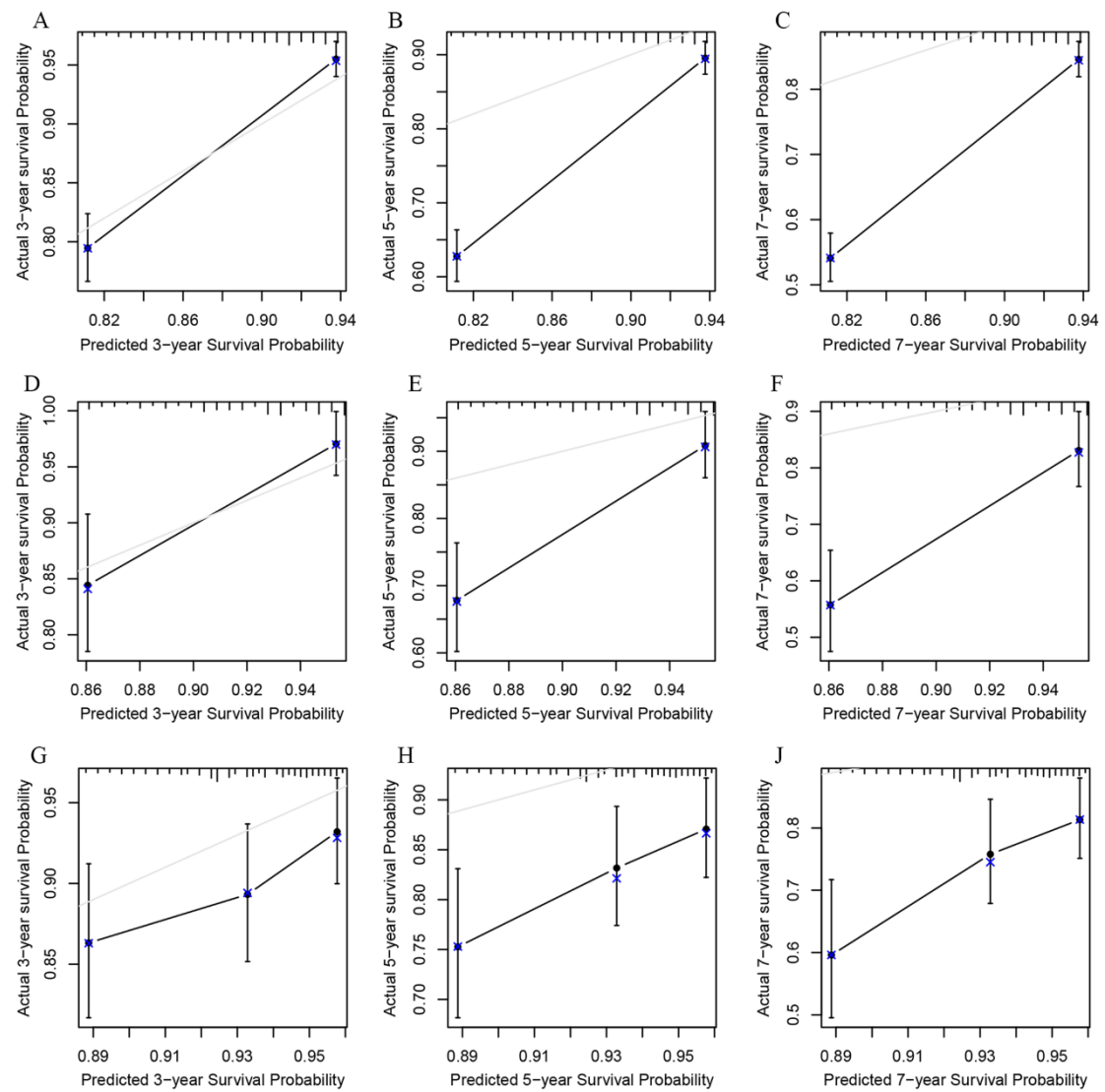
Supplementary figure 8. Internally and externally calibration of the nomogram in the training set (A-C), test set (D-F) and validation set (G-J).