

# Simulated examples for the survival ensemble methods

**Select data type**

ELSA\_Diabetes\_Study2 ▼

**Time point for event prediction:**

10

**Random seed for calibration and validation**

42

**K\_Outer loop CV (for validation)**

3

**K\_inner CV folds (model tuning)**

3

**Simulated data: random seed (generation):**

4242

**Sample size:**

150

**Observation time**

5

**Expected event prevalence by study end**

0.5

**Expected drop out rate**

0.3

**Custom data: path to data file**

~/Desktop/Study\_KCL/PhD P

**Predictors to use in the model**

"baseline\_age\_", "genderdun

**Time variable name**

time

**Event indicator variable name**

event

Sample statistics

CoxPH

SRF

Ens1: CoxPH->SRF

Ens2: CoxPH in clusters

Ens3: extended CoxPH

Summary

Conclusions

Internally cross-validated results:

Show10▼entries

Search:

	AUCROC	BS	BS_scaled	C_score	Calib_slope	Calib_alpha	T
test	0.7778	0.0775	0.0946	0.7473	1.0887	0.0462	10
train	0.7872	0.0764	0.1086	0.7549	1.1331	0.0453	10

Showing 1 to 2 of 2 entries

Previous

1

Next

Internally cross-validated Test results for each CV fold:

Show10▼entries

Search:

	AUCROC	BS	BS_scaled	C_score	Calib_slope	Calib_alpha	T
test.1	0.7711	0.0867	0.0952	0.7428	1.0928	0.0734	10
test.2	0.7611	0.076	0.0718	0.7206	0.9451	0.0374	10
test.3	0.8013	0.0698	0.1167	0.7785	1.2281	0.0278	10

Showing 1 to 3 of 3 entries

Previous

1

Next

CoxPH coefficients:

Show25▼entries

Search:

	coef	exp(coef)	se(coef)	z	Pr(> z )
sz20_	0.0322	1.0327	0.0455	0.7075	0.4792
pc1_	0.0253	1.0256	0.0452	0.5601	0.5754
pc2_	-0.0651	0.9369	0.0451	-1.4432	0.149
pc3_	0.0068	1.0068	0.0472	0.1435	0.8859
pc4_	-0.0111	0.9889	0.0443	-0.2512	0.8016
age_	0.176	1.1925	0.0565	3.1175	0.0018
sex	-0.3102	0.7333	0.1027	-3.019	0.0025
bmi_0_	0.4794	1.6151	0.0415	11.5451	0
hyp_0	0.5093	1.6642	0.094	5.4163	0

	coef <small>⬆</small>	exp(coef) <small>⬆</small>	se(coef) <small>⬆</small>	z <small>⬆</small>	Pr(> z ) <small>⬆</small>
cvd_0	-0.0261	0.9742	0.1326	-0.1972	0.8437
B_dep_0	0.3015	1.3518	0.1215	2.4816	0.0131
trig_0	0.108	1.1141	0.0284	3.7987	0.0001
baseline_hdl	-0.4483	0.6387	0.1565	-2.8651	0.0042
stroke_0	0.4717	1.6027	0.2252	2.0942	0.0362
B_smokstatus_0	0.3797	1.4619	0.1193	3.1815	0.0015
exercise_light	-0.0306	0.9699	0.195	-0.1568	0.8754
exercise_vig	-0.2421	0.785	0.1105	-2.1913	0.0284
EduLevel_low	0.3944	1.4835	0.1788	2.206	0.0274
EduLevel_med	0.2889	1.3349	0.1691	1.7084	0.0876
wealth_med	0.0135	1.0136	0.1226	0.1104	0.9121
wealth_low	0.1641	1.1783	0.1257	1.3059	0.1916
t2dm_	0.3107	1.3644	0.0466	6.6616	0

Other results:

```
|
|
| 0%
|
|=====| 20%
|
|=====| 40%
|
|=====| 60%
|
|=====| 80%
|
|=====| 100%

$test
  T    AUCROC      BS  BS_scaled  C_score
1 10 0.7710866 0.08671481 0.09523727 0.7427818
2 10 0.7611132 0.07604008 0.07184847 0.7205960
3 10 0.8012918 0.06981161 0.11672662 0.7784708
  Calib_slope Calib_alpha test cv_n
1    1.0927587 0.07336480    1    1
2    0.9451367 0.03742542    1    2
3    1.2281164 0.02778624    1    3

$train
  T    AUCROC      BS  BS_scaled  C_score
1 10 0.7913817 0.07189087 0.11073732 0.7551647
2 10 0.7939409 0.07726153 0.11842344 0.7660196
3 10 0.7763437 0.07991728 0.09655729 0.7434453
  Calib_slope Calib_alpha test cv_n
1    1.145628 0.04191689    0    1
2    1.117113 0.04608451    0    2
3    1.136618 0.04798016    0    3

$testaverage
      T    AUCROC      BS  BS_scaled
10.00000000 0.77783056 0.07752216 0.09460412
  C_score Calib_slope Calib_alpha      test
0.74728287 1.08867061 0.04619215 1.00000000

$trainaverage
      T    AUCROC      BS  BS_scaled
10.00000000 0.78722211 0.07635656 0.10857268
  C_score Calib_slope Calib_alpha      test
0.75487650 1.13311944 0.04532719 0.00000000

$model_list
$model_list[[1]]
Call:
coxph(formula = as.formula(paste("Surv(df_train$time, df_train$event) ~",
  paste(predict.factors, collapse = "+"))), data = df_train,
  x = TRUE)

      coef exp(coef) se(coef)      z
sz20_ -0.028345 0.972053 0.055940 -0.507
pc1_   0.019860 1.020059 0.057520 0.345
```

pc2_	-0.091728	0.912353	0.057566	-1.593
pc3_	-0.030055	0.970392	0.060139	-0.500
pc4_	-0.039656	0.961120	0.058256	-0.681
age_	0.146215	1.157445	0.071871	2.034
sex	-0.369663	0.690967	0.131234	-2.817
bmi_0_	0.469355	1.598963	0.055908	8.395
hyp_0	0.561660	1.753581	0.120607	4.657
cvd_0	-0.028051	0.972339	0.168995	-0.166
B_dep_0	0.379309	1.461275	0.150855	2.514
trig_0	0.102940	1.108425	0.035682	2.885
baseline_hdl	-0.299819	0.740952	0.196535	-1.526
stroke_0	0.330850	1.392151	0.303594	1.090
B_smokstatus_0	0.378792	1.460519	0.151321	2.503
exercise_light	0.007339	1.007366	0.245089	0.030
exercise_vig	-0.343196	0.709499	0.145865	-2.353
EduLevel_low	0.366422	1.442564	0.220926	1.659
EduLevel_med	0.178474	1.195392	0.210247	0.849
wealth_med	0.085987	1.089792	0.159795	0.538
wealth_low	0.230550	1.259292	0.162348	1.420
t2dm_	0.319483	1.376417	0.059816	5.341

p

sz20_	0.61237
pc1_	0.72988
pc2_	0.11106
pc3_	0.61724
pc4_	0.49605
age_	0.04191
sex	0.00485
bmi_0_	< 2e-16
hyp_0	3.21e-06
cvd_0	0.86817
B_dep_0	0.01192
trig_0	0.00392
baseline_hdl	0.12713
stroke_0	0.27581
B_smokstatus_0	0.01231
exercise_light	0.97611
exercise_vig	0.01863
EduLevel_low	0.09720
EduLevel_med	0.39595
wealth_med	0.59050
wealth_low	0.15558
t2dm_	9.24e-08

Likelihood ratio test=257.7 on 22 df, p=&lt; 2.2e-16

n= 3978, number of events= 305

\$model\_list[[2]]

Call:

```
coxph(formula = as.formula(paste("Surv(df_train$time, df_train$event) ~",
  paste(predict.factors, collapse = "+"))), data = df_train,
  x = TRUE)
```

	coef	exp(coef)	se(coef)	z
sz20_	0.05845	1.06019	0.05648	1.035
pc1_	0.02100	1.02122	0.05473	0.384

pc2_	-0.01324	0.98684	0.05520	-0.240
pc3_	0.03302	1.03358	0.05921	0.558
pc4_	0.03321	1.03376	0.05220	0.636
age_	0.24278	1.27478	0.06807	3.566
sex	-0.27238	0.76156	0.12402	-2.196
bmi_0_	0.47992	1.61595	0.04985	9.628
hyp_0	0.56260	1.75523	0.11439	4.918
cvd_0	-0.04503	0.95597	0.15915	-0.283
B_dep_0	0.32119	1.37877	0.14796	2.171
trig_0	0.10805	1.11410	0.03174	3.404
baseline_hdl	-0.60433	0.54644	0.19215	-3.145
stroke_0	0.56474	1.75900	0.25133	2.247
B_smokstatus_0	0.36836	1.44537	0.14856	2.480
exercise_light	0.02430	1.02459	0.23387	0.104
exercise_vig	-0.21397	0.80738	0.13363	-1.601
EduLevel_low	0.42431	1.52853	0.22276	1.905
EduLevel_med	0.34439	1.41113	0.21110	1.631
wealth_med	-0.00133	0.99867	0.14948	-0.009
wealth_low	0.17185	1.18750	0.15294	1.124
t2dm_	0.31351	1.36823	0.05610	5.589

p

sz20_	0.300705
pc1_	0.701234
pc2_	0.810360
pc3_	0.577049
pc4_	0.524662
age_	0.000362
sex	0.028072
bmi_0_	< 2e-16
hyp_0	8.73e-07
cvd_0	0.777231
B_dep_0	0.029943
trig_0	0.000663
baseline_hdl	0.001661
stroke_0	0.024641
B_smokstatus_0	0.013154
exercise_light	0.917258
exercise_vig	0.109347
EduLevel_low	0.056809
EduLevel_med	0.102802
wealth_med	0.992903
wealth_low	0.261150
t2dm_	2.28e-08

Likelihood ratio test=310.2 on 22 df, p=&lt; 2.2e-16

n= 3979, number of events= 336

\$model\_list[[3]]

Call:

```
coxph(formula = as.formula(paste("Surv(df_train$time, df_train$event) ~",
  paste(predict.factors, collapse = "+"))), data = df_train,
  x = TRUE)
```

	coef	exp(coef)	se(coef)	z
sz20_	0.065039	1.067201	0.055387	1.174
pc1_	0.032623	1.033161	0.054653	0.597

pc2_	-0.089742	0.914167	0.053964	-1.663
pc3_	0.016462	1.016598	0.054782	0.300
pc4_	-0.042009	0.958861	0.053633	-0.783
age_	0.142562	1.153225	0.068243	2.089
sex	-0.315088	0.729725	0.124046	-2.540
bmi_0_	0.486410	1.626467	0.048283	10.074
hyp_0	0.424529	1.528870	0.112253	3.782
cvd_0	-0.006792	0.993231	0.161008	-0.042
B_dep_0	0.234661	1.264480	0.148876	1.576
trig_0	0.117352	1.124515	0.040517	2.896
baseline_hdl	-0.433751	0.648073	0.189558	-2.288
stroke_0	0.472919	1.604671	0.282917	1.672
B_smokstatus_0	0.387311	1.473014	0.141095	2.745
exercise_light	-0.129075	0.878908	0.240804	-0.536
exercise_vig	-0.186913	0.829516	0.129140	-1.447
EduLevel_low	0.391485	1.479176	0.214204	1.828
EduLevel_med	0.333466	1.395797	0.201573	1.654
wealth_med	-0.038004	0.962709	0.143598	-0.265
wealth_low	0.097583	1.102503	0.148756	0.656
t2dm_	0.294049	1.341849	0.056125	5.239

p

sz20_	0.240290
pc1_	0.550563
pc2_	0.096314
pc3_	0.763797
pc4_	0.433472
age_	0.036703
sex	0.011082
bmi_0_	< 2e-16
hyp_0	0.000156
cvd_0	0.966352
B_dep_0	0.114976
trig_0	0.003775
baseline_hdl	0.022124
stroke_0	0.094607
B_smokstatus_0	0.006051
exercise_light	0.591947
exercise_vig	0.147794
EduLevel_low	0.067605
EduLevel_med	0.098063
wealth_med	0.791276
wealth_low	0.511828
t2dm_	1.61e-07

Likelihood ratio test=257.4 on 22 df, p=< 2.2e-16  
 n= 3979, number of events= 345

\$time

Time difference of 5.365462 secs