Supplementary Material: A Hierarchical Bayesian Model for Personalized Survival Predictions

I. Appendix

This appendix includes further predictive performance evaluations of HBS and all baseline algorithms on two additional medical datasets stemming from the UK Biobank and the German Breast Cancer Study Group (GBSG2).

A. UK Biobank

We conducted a second experiment on a subset of the UK Biobank dataset[1] based on which we are interested in predicting the onset of myocardial infraction and ischemic heart disease, leading causes of morbidity and mortality, accounting for a global figure of 30% of deaths in the world [1]. UK Biobank is a large national and international health resource comprising patients from primary care with a wide range of medical conditions. We extracted a cohort of $1,500$ patients with no previous history of cardiovascular diseases (CVD) followed for 10 years since registration. Patients are associated with eight core covariates: gender, age, smoking status, systolic blood pressure, blood pressure treatment, total cholesterol, HDL cholesterol and diabetes. We evaluated HBS with a subgroup partition determined by 4 groups according to the cholesterol and hypertension treatment patients received. Parameter values resulting in the performance figures for HBS shown on Tables I and II were estimated with 5000 iterations of our posterior sampling algorithm in addition to 1000 iterations as burn-in. Performance is computed at the 0.25, 0.5 and 0.75 time quantiles of the observed event times.

TABLE I
C-index figures (High C-index better) at specified time quantiles on the UK Biobank dataset.

Models	0.25	0.5	0.75	⏱
Cox	0.653±0.078	0.694±0.048	0.692±0.060	< 1s
Weibull	0.651±0.072	0.693±0.040	0.693±0.061	< 1s
Aalen	0.650±0.073	0.689±0.043	0.688±0.063	5s
CoxBoost	0.658±0.068	0.697±0.044	0.697±0.052	14.9s
SRF	0.595±0.088	0.631±0.058	0.639±0.064	6.15s
CForest	0.634±0.078	0.660±0.050	0.660±0.054	4.41s
BART	0.633±0.079	0.650±0.050	0.652±0.052	1116s
Wei-Tree	0.649±0.082	0.690±0.051	0.691±0.065	1564s
HBS	0.665±0.078	0.695±0.052	0.701±0.055	1780s

TABLE II
Brier Score (Low BS better) at specified time quantiles on the UK Biobank dataset.

Models	0.25	0.5	0.75	⏱
Cox	0.216±0.020	0.227±0.051	0.296±0.021	< 1s
Weibull	0.217±0.020	0.227±0.050	0.298±0.022	< 1s
Aalen	0.218±0.021	0.229±0.052	0.298±0.021	5s
CoxBoost	0.215±0.018	0.225±0.050	0.295±0.021	14.9s
SRF	0.230±0.020	0.242±0.051	0.317±0.023	6.15s
CForest	0.217±0.017	0.228±0.052	0.301±0.022	4.41s
BART	0.217±0.017	0.227±0.053	0.299±0.022	1116s
Wei-Tree	0.215±0.016	0.229±0.053	0.301±0.022	1564s
HBS	0.214±0.017	0.226±0.053	0.297±0.023	1780s

TABLE III
C-index figures (High C-index better) at specified time quantiles on the GBSG2 dataset.

Models	0.25	0.5	0.75	⏱
Cox	0.710±0.034	0.680±0.028	0.665±0.025	< 1s
Weibull	0.705±0.033	0.677±0.025	0.664±0.023	< 1s
Aalen	0.705±0.034	0.675±0.026	0.662±0.025	4.5s
CoxBoost	0.714±0.034	0.687±0.028	0.667±0.025	42.5s
SRF	0.710±0.034	0.677±0.028	0.670±0.026	8.8s
CForest	0.728±0.030	0.691±0.028	0.676±0.027	1.1s
BART	0.726±0.031	0.697±0.025	0.673±0.026	381s
Wei-Tree	0.728±0.031	0.704±0.026	0.679±0.028	590s
HBS	0.737±0.032	0.725±0.026	0.686±0.027	646s

B. German Breast Cancer Study Group (GBSG2)

In a third experiment we consider a prospective controlled clinical trial conducted to assess the effect on survival of a breast cancer treatment, hormonal therapy, by the German Breast Cancer Study Group (GBSG2) [2], [3]. This data set is publicly available and contains 686 women of whom 171 died, followed-up for a median time of nearly 5 years. In addition to the hormonal treatment considered, the women are described by 8 features relating to tumor size and clinical measurements relevant to breast cancer. Further pre-processing details can be found in [2]. We evaluated HBS and baseline algorithms with the previous specifications, performance results are given in Tables III and IV, computed at the 0.25, 0.5 and 0.75 time quantiles of the observed event times. HBS is competitive on every time horizon and outperforms in terms of C-index.

Supplementary material submitted January 15, 2018

[1] Availabble at http://www.ukbiobank.ac.uk

References

[1] S. Mendis, *Global status report on noncommunicable diseases 2014*. World health organization, 2014.

[2] M. Schumacher, C. Schmoor, W. Sauerbrei, A. Schauer, L. Ummenhofer, W. Gatzemeier, and H. Rauschecker, "The prognostic effect of histological tumor grade in node-negative breast cancer patients," *Breast cancer research and treatment*, vol. 25, no. 3, pp. 235–245, 1993.

TABLE IV
BRIER SCORE (LOW BS BETTER) FIGURES AT SPECIFIED TIME QUANTILES ON THE GBSG2 DATASET.

Models	**0.25**	**0.5**	**0.75**	⏱
CPH	0.146±0.011	0.199±0.010	0.214±0.010	$<$ 1s
Weibull	0.147±0.011	0.199±0.010	0.215±0.011	$<$ 1s
Aalen	0.148±0.011	0.199±0.010	0.215±0.010	4.5s
CoxBoost	0.146±0.011	0.200±0.009	0.218±0.009	42.5s
SRF	0.145±0.012	0.201±0.013	0.219±0.016	8.8s
CForest	0.140±0.011	0.194±0.012	0.216±0.013	1.1s
BART	0.138±0.009	0.194±0.010	0.217±0.010	381s
Wei-Tree	0.147±0.010	0.196±0.011	0.218±0.013	590s
HBS	0.139±0.010	0.194±0.010	0.217±0.012	646s

[3] W. Sauerbrei, P. Royston, H. Bojar, C. Schmoor, and M. Schumacher, "Modelling the effects of standard prognostic factors in node-positive breast cancer. german breast cancer study group (gbsg)." *British Journal of Cancer*, vol. 79, no. 11-12, pp. 1752–1760, 1999.