**Appendix A: Machine Learning Models**

**Cox Lasso**

The Cox Lasso applies Lasso (L1) regularization to the Cox proportional hazards model for regression on right-censored time-to-event outcomes. The method performs variable selection by applying a penalty during model fitting that sets less important predictor coefficients to zero. The remaining (non-zero) coefficients comprise the selected predictors. A tuning parameter controls the extent of this shrinkage: larger values of the tuning parameter correspond to more shrinkage and thus the selection of fewer predictors. We fit the Cox Lasso using the *glmnet* package in R, with the tuning parameter selected via cross-validation to balance model simplicity and fit.1

**Survival Random Forest**

The survival random forest, as implemented in the *randomForestSRC* R package, uses an ensemble tree method designed for right-censored time-to-event data. A log-rank split rule is used, and the estimates associated with each terminal node are computed using the Kaplan-Meier estimator (survival estimate) and the Nelson-Aalen estimator (cumulative hazard estimate). Estimates for an individual are averaged over all bootstrap samples for which the individual is out of bag (OOB). Prediction error for the forest is measured by 1-C, where C is Harrell’s concordance index, a measure of accuracy in ranking pairs in terms of their predicted and actual survival.2

**Generalized additive model**

A generalized additive model (GAM) is a regression model that allows for non-linear relationships between predictors and the outcome. In the R package *mgcv*, which we used for our model, smooth terms are fit using penalized regression splines. The generalized additive model accommodates right-censored time-to-event data by fitting a Cox proportional hazards model with the smooth terms incorporated in the partial likelihood.3

**Gradient boosted regression**

Gradient boosting uses an iterative method to fit a regression function to the data. At each iteration, the gradient, or the derivative of the loss function with respect to the current regression function, is calculated. The regression function is then updated in the direction of this gradient, improving the fit. Gradient boosted regression as implemented in the R package *gbm*, which we used for our model, uses regression trees as the functions. To accommodate right-censored time-to-event data, the model uses the negative log partial likelihood under the Cox proportional hazards model as the loss function.4,5

**REFERENCES**

1. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent. *J Stat Softw*. 2011;39(5). doi:10.18637/jss.v039.i05

2. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. 2008;2(3):841-860. doi:10.1214/08-AOAS169

3. Wood SN. *Generalized Additive Models: An Introduction with R*. 2nd ed. Chapman and Hall/CRC; 2017. doi:10.1201/9781315370279

4. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat*. 2001;29(5). doi:10.1214/aos/1013203451

5. Friedman JH. Stochastic gradient boosting. *Comput Stat Data Anal*. 2002;38(4):367-378. doi:10.1016/S0167-9473(01)00065-2