

Modeling Mortality of Pediatric Patients Undergoing Hematopoietic Stem Cell Transplantation Using Supervised Machine Learning

Albert Cao¹, Allison Dunn¹, Jogarao Gobburu¹, Janel Long-Boyle^{2,3}, Rahul Goyal¹

¹Center for Translational Medicine, School of Pharmacy, University of Maryland Baltimore, Maryland

²Department of Pediatrics, Division of Allergy, Immunology, and Bone Marrow Transplantation, University of California San Francisco, San Francisco, California

³Department of Clinical Pharmacy, University of California San Francisco, San Francisco, California



SCAN ME

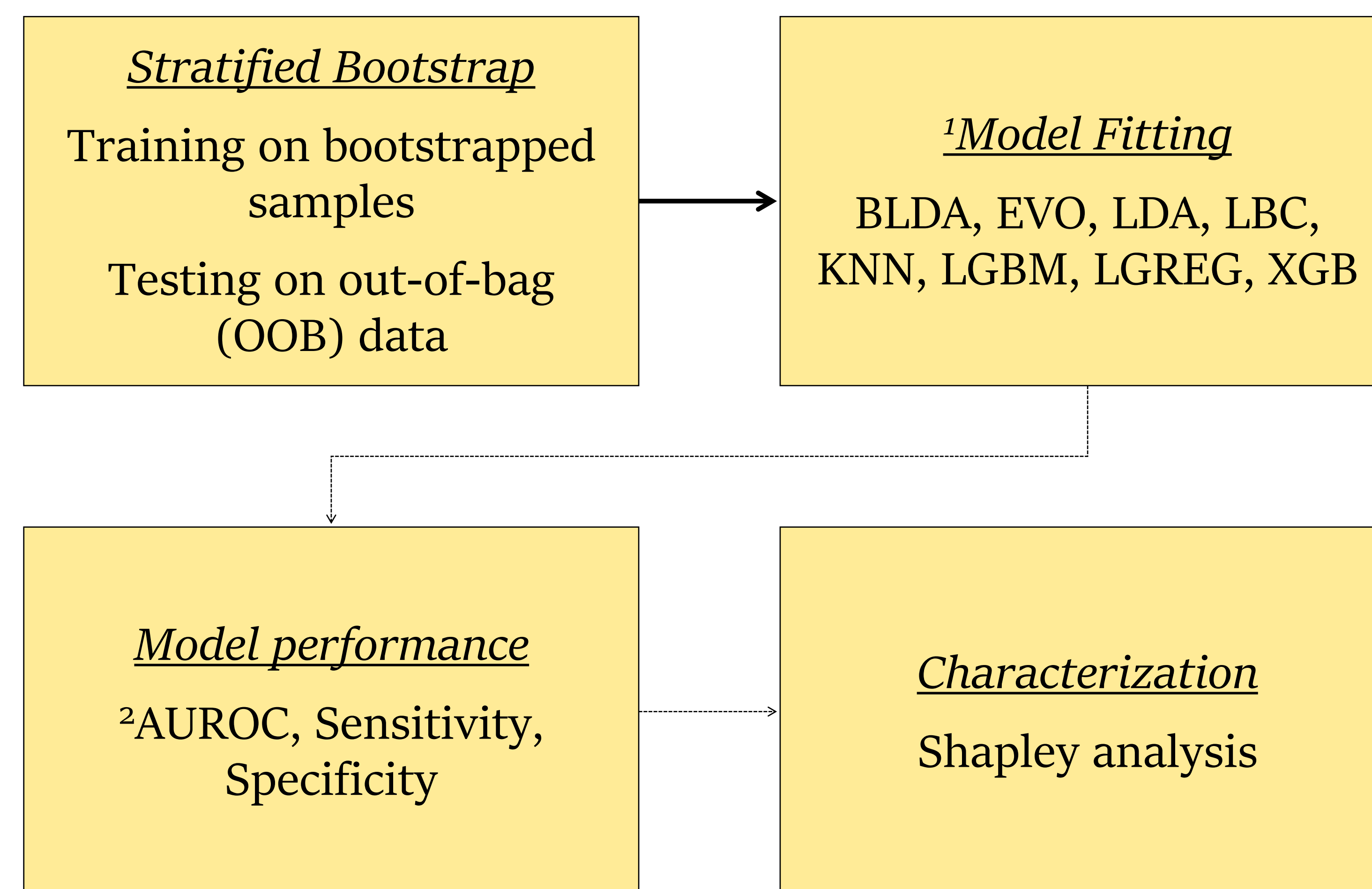
Background

- Busulfan is an alkylating agent used as pretreatment for hematopoietic stem cell transplantation (HSCT) in both hematologic malignant and non-malignant disorders in pediatric patients.
- The relationship between transplantation parameters and mortality in pediatric patients with malignant disorders has not yet been well characterized.

Objective

- Identify risk factors of 1-year mortality in pediatric patients that undergo HSCT using machine learning.

Methods



¹BLDA – Bayesian linear discriminant analysis (LDA), ET – Evo trees, LBC – linear binary classifier, KNN – k-nearest neighbors, LGBM – Light GBM, LGREG – logistic classifier, XGB – XGBoost
²AUROC – area under the receiver operator characteristic curve

Software: Pumas 2.3.1 (pumas.ai)

Low Busulfan exposure and increased degree of HLA mismatch with donor cells are the most important risk factors of mortality at 1-year in pediatric patients receiving hematopoietic progenitor cell transplant

Results

Figure 1 – Model performance as shown by sensitivity. Each boxplot is generated based on the distribution of 200 sensitivities calculated from out-of-bag data. Dotted horizontal line represents a dummy classifier

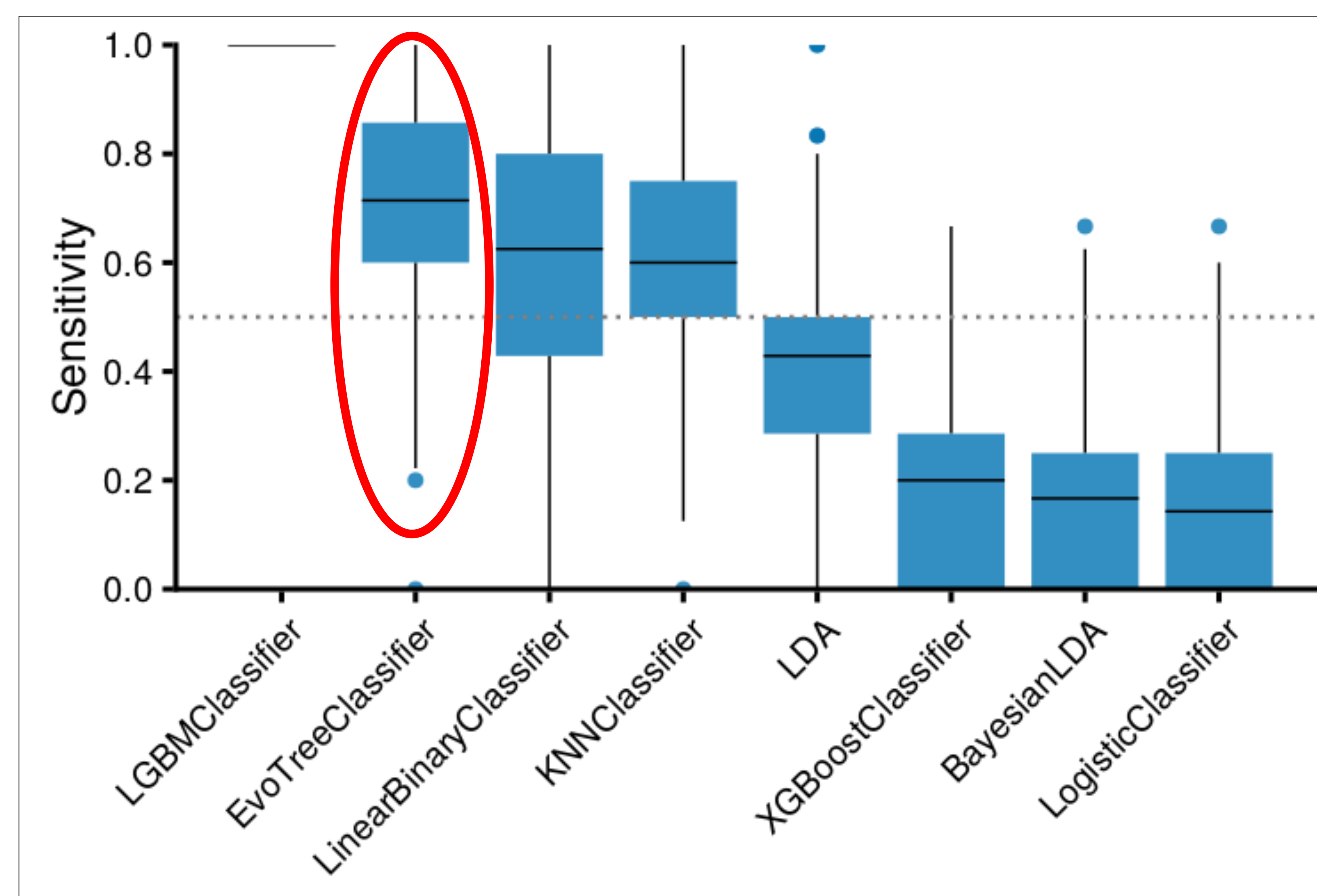
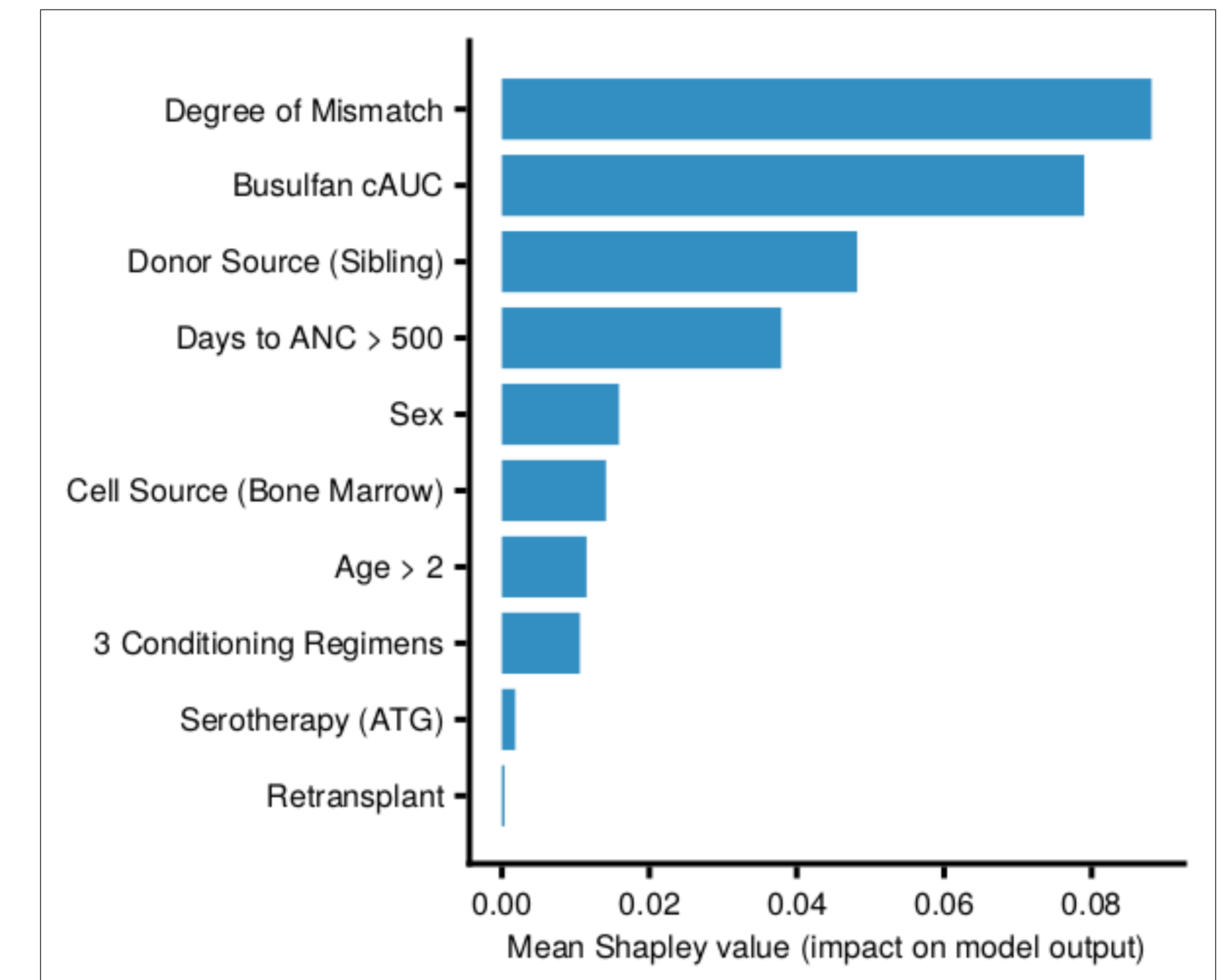


Figure 2 – Feature importance obtained from the final Evo Trees Classifier model. Feature importance was calculated as the mean Shapley value across all bootstrapped resamples for each feature value



Median performance of final Evo Trees algorithm on out-of-bag data: AUROC – 0.53; Sensitivity – 0.71; Specificity – 0.34

Conclusion

- Busulfan cumulative AUC < 85 mg.hr/L and high degree of HLA mismatch are associated with increased odds of mortality at 1-year. Patients who meet this criteria may benefit from more rigorous monitoring.
- With availability of new data, performance of the final model can be improved and applied in real-time.