Interpretable Treatment Prioritization Rule Identifies Diabetic Patients Who Can Benefit from Prompt Coronary Revascularization

Chirag Nagpal and Artur Dubrawski
Auton Lab, Carnegie Mellon University

Background:
The BARI2D [1] trial (n=2368) was originally inconclusive (HR: 0.875-1.217, 95% CI) in determining optimal treatment (Medical Therapy vs. Revascularization) for Type-2 Diabetics with ischemic heart disease to reduce major adverse cardiovascular events.

Methods:
We applied a data driven machine learning approach (Sparse Cox Mixtures) [2, 3] to the BARI2D clinical study data to identify a selection rule to prioritize patients based on response to medical therapy vs. early revascularization, and found evidence suggesting heterogeneous benefits.

Results:
We trained our interpretable rule model on 1184 participants (~50%) of the study stratified by treatment assignment and event incidence, and validated on the remainder of the study population. Our model discovered 5 covariates (Serum Potassium, Total and LDL Cholesterol, Age and use of Sulfonylurea) as effective predictors of treatment benefit.

Conclusion:
We identified a High Risk group (n=395, ~30%) that were harmed by medical therapy (HR: 1.67±0.66, 95% CI) while the rest benefitted (HR: 0.72±0.25, 95% CI). Validation of the rule on held out participants supports the hypothesis of heterogeneous benefit. Among the held out set, medical therapy harmed the High Risk group (n=395, ~30%) (HR: 1.34±0.51, 95% CI), while the remaining validation set participants benefited (HR: 0.86±0.27, 95% CI).

References:
[1] BARI 2D Study Group. "A randomized trial of therapies for type 2 diabetes and coronary artery disease." *New England Journal of Medicine* 360.24 (2009): 2503-2515.
[2] Nagpal, Chirag, et al. "Recovering Sparse and Interpretable Subgroups with Heterogeneous Treatment Effects with Censored Time-to-Event Outcomes." *arXiv preprint arXiv:2302.13457* (2023).
[3] Nagpal, Chirag, et al. "auton-survival: an Open-Source Package for Regression, Counterfactual Estimation, Evaluation and Phenotyping with Censored Time-to-Event Data." *Machine Learning for Healthcare Conference*. PMLR, 2022. (2022).

Figure 1: **A)** The discovered sparse prioritization rule. **B)** The population level survival estimates for the training and validation cohorts in the BARI2D trial stratified by treatment assignment. As evidenced from the figures, significant differences in treatment benefit were not observed at a population level. **B)** The discovered **High Risk** patients phenogroups on the training and validation cohorts. Notice that Medical Therapy **harms** the patients as evidenced by higher hazard rates. **C)** The rest of the study population **benefit** from **Medical Therapy** with lower hazard rates in both the training and held-out validation cohort.

Learnt Interpretable Rule by Sparse Cox Mixtures					
Feature	**Std. Age**	**Std. Serum Potassium**	**Std.Total Cholesterol**	**Std. LDL Cholesterol**	**Sulfonylurea Use**
Coefficient	**(+)0.308**	**(+)0.152**	**(+)0.006**	**(+)0.035**	**(+)0.052**