Risk Prediction for Atherosclerotic Cardiovascular Disease

Jonas Kemp (jbkemp7), Aditya Kanukurthy (sang33t1)
CS 229, Fall 2016

Motivation

Risk prediction for atherosclerotic cardiovascular disease (ASCVD) is an active area of research. Current risk calculators used in clinical practice are based on the ASCVD pooled cohort equations (PCEs), a set of Cox proportional hazards models developed by the American Heart Association and American College of Cardiology in 2013. However, independent validation studies have suggested that these equations overestimate risk for minority groups, which may lead to additional health and financial burdens from overtreatment.[1] Here, we aim to develop a new model for ASCVD risk prediction that mitigates the risk overestimation issue.

Data and Features

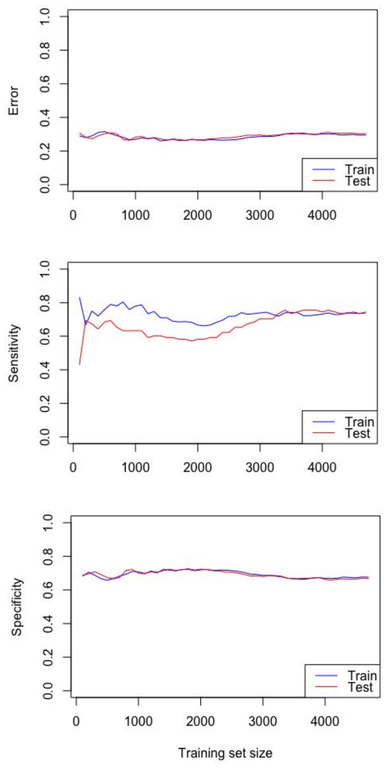
We fit our models using patient data from eight cohort studies, five of which were used to fit the original pooled cohort equations. Outcomes were defined as an ASCVD event (heart attack, stroke, or death from coronary heart disease) within the first 10 years of any study. Features recorded for each patient included age, gender, race, BMI, cholesterol levels, blood pressure, hypertension treatment status, smoking status, and diabetes status. We also considered a quadratic feature for age, an interaction feature for age with each of the medical variables, and an interaction between blood pressure and hypertension treatment. The final pooled dataset included over 29,000 patients.

Selected References

Thanks to Dr. Sanjay Basu for providing access to the data and for guidance throughout the project.

1. DeFilippis et al., 2016. *European Heart Journal.*
2. De Bin, 2016. *Computational Statistics 31* (2): 513-31.
3. Ishwaran et al. 2008. *Annals of Applied Statistics 2* (3): 841-60.

Error Analysis of Baseline Model



The original PCEs include one equation for each combination of race (black or white) and gender (male or female). Hazard ratios are converted into risk scores using the following formula:

$$p_i = 1 - S_0(t)^{\exp(\beta^T(X_i - \overline{X}))}$$

where S_0 is the baseline hazard at time t, β are the model parameters, X_i are the predictors for patient i, and $\overline{X}$ are the mean predictor values. The typical "high-risk" threshold is a 10-year risk of 7.5%, the recommended clinical threshold for prescribing statin therapy.

Other researchers have hypothesized that the PCEs may suffer from overfitting to small minority samples or from cohort effects. We initially attempted to fit regularized and mixed-effects Cox models to improve on these problems, with little success. However, error analysis of the baseline model fit to the full eight-cohort dataset revealed a much more substantial bias problem (left).

Models

To attempt to reduce bias, we fit the following ensemble survival models for our analysis:

- *Boosted Cox models:* Cox regression models fit using either gradient- or likelihood-based boosting approaches. Implemented in the R packages mboost and CoxBoost.[2]
- *Random survival forests:* an extension of Breiman's random forests to handle right-censored survival data and estimate cumulative hazard. Implemented in the R package randomForestSRC.[3]

Results

Men	*Train*			*Test*		
	Sens.	Spec.	AUC	Sens.	Spec.	AUC
Cox	0.89	0.43	0.74	0.88	0.42	0.72
Gradient Boosting	0.86	0.39	0.70	0.86	0.38	0.68
Likelihood Boosting	0.86	0.44	0.73	0.87	0.43	0.72
Random Forest	1	0.68	1	0.85	0.41	0.70

Women	*Train*			*Test*		
	Sens.	Spec.	AUC	Sens.	Spec.	AUC
Cox	0.78	0.65	0.79	0.71	0.66	0.72
Gradient Boosting	0.49	0.85	0.76	0.42	0.84	0.69
Likelihood Boosting	0.70	0.71	0.78	0.65	0.70	0.73
Random Forest	1	0.84	1	0.65	0.69	0.72

* Sensitivity and specificity are reported at the 7.5% risk threshold.

Discussion and Future Directions

Our analysis reveals significant challenges in ASCVD risk prediction. In the classification setting, we failed to find evidence of risk overestimation among minority groups, though we did observe consistently high false positive rates among men relative to women. Future research should further investigate this discrepancy in order to reconcile our work with previous literature, perhaps through unsupervised techniques to better understand underlying structure in the data.

None of our ensemble methods succeeded in outperforming the baseline Cox model, although those that matched performance frequently favored higher specificity at the recommended risk threshold. Thus, these models might decrease overtreatment under current guidelines. Whether the lost sensitivity is acceptable is a matter for clinical expertise. We expect that an expanded feature set would likely offer a boost in performance and reduce the need for tradeoffs.