**Minutes of the Population-Level Estimation Workgroup**

November 16, 2016

Present: Alejandro Schuler, Ajou students, Martijn Schuemie, Yuriy Khoma, Lan Kelly, Nicole Pratt, Yaromir Shpilevskiy

Today Alejando Schuler from Nigam Shah’s group (Stanford) presented his ongoing research on propensity scores.

Alejandro started by focusing on a particular clinical questions: “will AB or CD drug be more effective in lowering the blood pressure of newly diagnosed hypertension patients?”

About 10,000 features were generated from the data (not using FeatureExtraction). Fitting PS model using elastic net.

PS model evaluation framework

* Metrics of fit
* Balance metrics

When minimizing the cross-validation error to find the optimal hyperparameter, Alejandro observed overfitting and overcorrection in terms of covariate balance.

Alejandro introduced the 1-SE rule: find model within 1 SD of mininum CV error, so deliberately underfitting the model from the CV optimum.

Alejandro also proposed the notion of only computing PS out of the training sample.

Alejandro then focused on another clinical question: “Causal effect of type of organ failure during sepsis care on long-term survival for patients who survive sepsis”

In this study Alejandro will use plasmode simulation, so he can compare estimated to true propensity (amongst others).

Martijn wonders why we see much better after-matching balance in the CohortMethod package. Nicole points out that Alejandro uses non-binary covariates, whereas CohortMethod mainly creates binary covariates. Martijn and Alejandro agree to investigate the different results.