

# Should a propensity score model be super? The utility of machine learning procedures for causal adjustment

Shomoita Alam \*

[shomoita.alam@mail.mcgill.ca](mailto:shomoita.alam@mail.mcgill.ca)

---

In investigations of the effect of treatment on outcome, the propensity score is a tool to eliminate imbalance in the distribution of confounding variables between treatment groups. Recent work has suggested that Super Learner, an ensemble method, outperforms logistic regression in non-linear settings however experience with real data analyses tend to show overfitting of the propensity score model using this approach. We investigated wider range of settings of varying complexities including simulations based on real data to compare the performances of logistic regression, generalized boosted models, and Super Learner in providing balance and for estimating the average treatment effect via propensity score regression, propensity score matching, and inverse probability of treatment weighting. We found that Super Learner and logistic regression are comparable in terms of covariate balance and mean squared error, however Super Learner is computationally very expensive and may induce positivity violations in real data settings. Approaches based on generalized boosted models were inferior to both logistic regression and Super Learner in terms of both balance and mean squared error. We also found that propensity score regression adjustment was superior to either matching or inverse weighting when the form of the dependence on the treatment on the outcome is correctly specified. Finally, we note that to fully understand a complex estimation procedure, simulations based on both real and entirely synthetic data may be needed.

*This is joint with Erica E.M. Moodie (Department of Epidemiology, Biostatistics and Occupational Health, McGill University) and David A. Stephens (Department of Mathematics & Statistics, McGill University).*

---

\*Department of Mathematics & Statistics, McGill University, 805 Sherbrooke Street West, Montréal, QC H3A 0B9, CANADA