Examining the Effect of Exercise on the Progression and Severity of Huntington's Disease Using Different Covariate Balancing Methods and Simulated Data Derived from the PACE-HD Study

Andreas Markoulidakis, Khadijeh Taiyari, Peter Holmans, Philip Pallmann, Monica Busse-Morris, Cardiff University, UK;
Beth Ann Griffin, RAND, Arlington, VA

EHDN
September 09 – 11, 2021

Background

- Regular activity and structured exercise are hypothesized to lessen the impact of Huntington's disease (HD)
- Randomized controlled trials (RCTs) are the gold standard for measuring the causal effects of interventions on clinical outcomes
- Causal effects can also be inferred from observational studies but care must be taken to account for imbalance in confounders in groups being compared
- Propensity score (PS) and balance weighting are useful for reducing imbalances; they weight groups to look alike on observed confounders
- No guidance available on which balancing methods are best for handling trade-off between balance and power and small sample sizes, common in HD

Objective

To provide step-by-step guidelines for estimating causal effects when using observational data and comparing the performance of multiple balancing methods using data derived from Physical Activity and Exercise Outcomes in HD (PACE-HD).

Pace-HD

- PACE-HD is a RCT done within an observational cohort study comparing structured exercise (treatment) to "exercise as usual" (control)
- There are 111 participants; 28 in treatment and 83 in control
- Our primary outcome is the cUHDRS (a composite measure that is used to measure HD progression in terms of functioning, motor, and cognitive decline)
- We are considering 3 pretreatment confounders measured at baseline:
  1. 6' Walk: the distance (in meters) one is able to walk in 6 minutes
  2. Vo2max: maximum oxygen consumption during exercise
  3. cUHDRS: Here, captures baseline severity of the disease for each individual prior to the launch of the study

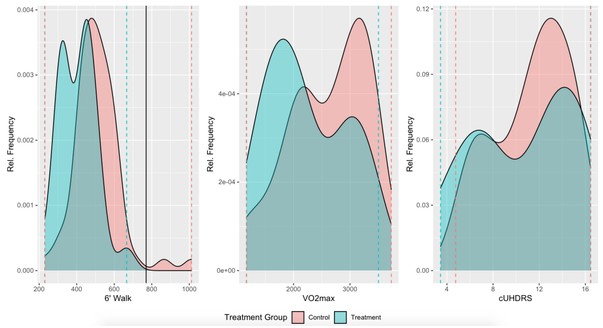
The 6 Key Steps

Step 1: Choose which estimand one is interested in (ATE, ATT, ATC)

- ATE (average treatment effect on the population) measures the effect of exercise for all people with HD like those enrolled in the study
- ATT (average treatment effect on the treated) measures the effect of exercise for individuals with HD who are exercising
- ATC (average treatment effect on the controls) is the opposite of ATT

Since individuals who exercise will tend to be different from those who do not (e.g., healthier), we opted to estimate ATT to understand the impact of exercise on HD for individuals like those who tend to exercise.

Step 2: Assess sample for any obvious overlap concerns and adjust sample size as needed



- By definition, the overlap assumption requires that all individuals in the study have a non-zero probability of receiving treatment (i.e., $\Pr(T_i = 1|\boldsymbol{X}_i)>0$)
- The figures above depict the smooth density plot (kernel approximation) for the treatment and control group for each confounding pretreatment covariate
- The dotted lines represent the minimum and maximum values for each group while the solid black like in the 6' Walk plot represents the point beyond which there is no mass for treatment group (thus it is not represented)
- There seems to be an overlap concern with 6' Walk Distance thus we removed the 2 individuals from the control group who are problematic

Step 3: Estimation of propensity scores or balancing weights, ideally using multiple methods

Key Terms	Definition
Logistic Regression (LR)	The basic LR model for estimating PS assumes that the *logit* of the probability of receiving treatment is equated with a linear combination of covariates $X_i^T\beta$
Covariate Balance Propensity Score (CBPS)	Fits a penalized version of LR that estimates the model subject to a constraint that prioritizes balancing confounders as well as model fit
Generalized Boosted Model (GBM)	Flexible, nonparametric machine learner that fits a piecewise-constant model, constructed as a combination of simple regression trees
Entropy Balance (EB)	Calculates weights through a reweighting scheme, until adequate balance in the pre-specified moments is achieved

Step 4: Assess balance and effective sample size for all methods and choose the best one for the final outcome analysis

			ES			
	Unweigh	**LR**	**GBM_ES**	**GBM_KS**	**CBPS**	**EB #1**
SIX_MIN	0.91	0.07	0.06	0.04	0.04	0.00
VO2max	0.52	0.02	0.13	0.17	0.03	0.00
cUHDRS_Y1	0.12	0.11	0.05	0.04	0.01	0.00
MEAN	0.52	0.07	0.08	0.08	0.03	0.00
MAX	0.91	0.11	0.13	0.17	0.04	0.00
			KS			
	Unweigh	**LR**	**GBM_ES**	**GBM_KS**	**CBPS**	**EB #1**
SIX_MIN	0.43	0.15	0.18	0.17	0.19	0.18
VO2max	0.37	0.10	0.14	0.16	0.11	0.10
cUHDRS_Y1	0.21	0.17	0.12	0.12	0.14	0.13
MEAN	0.34	0.14	0.15	0.15	0.15	0.14
MAX	0.43	0.17	0.18	0.17	0.19	0.18
			ESS			
	Unweigh	**LR**	**GBM_ES**	**GBM_KS**	**CBPS**	**EB #1**
	56.0	20.2	15.0	16.5	23.2	21.4

Key definitions:
ES = Effect size difference; mean difference between 2 groups divided by standard deviation
KS = Kolmogorov-Smirnov statistic; maximum difference in the empirical cumulative distribution functions between the 2 groups
ESS = Effective sample size; reduced sample size expected due balancing weights

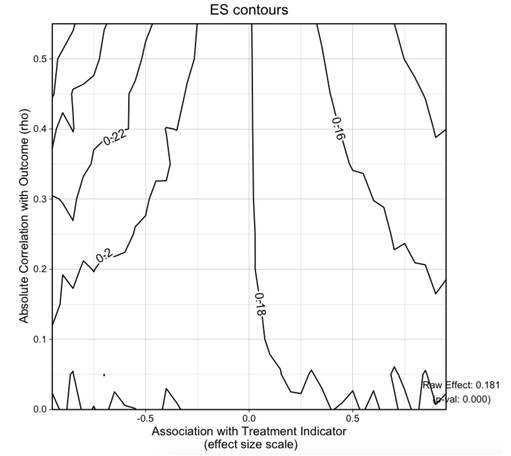
- EB and CBPS achieve adequate balance on ES (all < 0.1); however, no method achieves good balance on the KS-statistic (all > 0.1)
- This could occur due to low sample size (only 26 in treatment group)
- As expected, all methods meaningful reduce the ESS due to the design effect of using weights (a natural consequence of minimizing confounding)
- EB (controlling for 1 moment) performs best overall in terms of ES, KS and ESS, thus we will move to our outcome analysis with this weighting algorithm

Step 5: Model outcome and estimate the causal treatment effect

	Estimate	Standard Error	P-value	CI Lower	CI Upper
(Intercept)	-3.30	0.07	<2e-16	-3.44	-3.16
treat	0.18	0.03	1.75e-07	0.12	0.24
SIX_MIN	0.00	0.00	0.008	0.00	0.00
VO2max	0.00	0.00	0.591	-0.00	0.00
cUHDRS_Y1	0.26	0.00	<2e-16	0.26	0.27

- Outcome model controls for EB weights as survey weights and baseline confounders as classic control covariates
- Significant evidence (in our simulated data) that exercise group doing better; individuals in the exercise group have higher cUHDRS values than those who do not (by an average of 0.181), a year after the initial measurement

Step 6: Assess sensitivity of the results to unobserved confounding



- The contour lines shown represent the sensitivity of our estimated treatment effect to a potential unobserved confounder. The x-axis displays the ES between the treated and control groups on the unobserved confounder and the y-axis displays the absolute correlation of the unobserved confounder with the outcome covariate.
- The sign of the estimated treatment effect is expected to remain consistent across a range of possible values for the unobserved confounder. In the most extreme observed case, the estimated effect size would be reduced by 7 percent.
- Statistical significance is very robust. In the most extreme observed case, the p-value would be expected to increase from 0.000 to 0.002.

Conclusions

- We present a step-by-step guide for estimating causal intervention effects using observational data
- Guidelines show how to choose the best method for estimating the balancing weights, interpret the results, and assess key assumptions like overlap and unobserved confounding
- Lessons can apply more broadly to other HD studies using observational data

Acknowledgements

- DOMINO-HD is funded though the EU joint program for Neurodegenerative Disease Research as part of the JPND funding call in to Health and Social Care (2019) with funding from Alzheimer's Society, Secretary of State for Health and Social Care, Health and Care Research Wales, Public Health Agency Northern Ireland, Jacques and Gloria Gossweiler Foundation, Bundesministerium für Bildung und Forschung, Narodowe Centrum Badań i Rozwoju, Swiss National Science Foundation (SNF), 32ND30_185548 and Health Research Board (JPND-HSC-2018-003).
- Funding also provide by grant R01DA045049 (PI Griffin) through the National Institute of Drug Abuse.