

# Imperfect imputation: Adjusting for the error incurred when we impute.

**Kyle Grosser** 

## What is Huntington's Disease?

... a **neurodegenerative disease** that leads to **motor**, **cognitive**, and **psychiatric** impairments



Figure: George Huntington, MD. 1872

## How do we model Huntington's Disease?

Model of interest:

$$E[Y_{ij}|\cdot] = \beta X_{ij} + \alpha T_i + b_i$$

- *Y<sub>ij</sub>*: Cognitive impairment
- X<sub>ij</sub>: Striatum volume
- T<sub>i</sub>: Age at disease onset, possibly censored by C<sub>i</sub>
- $b_i$ : Random intercept

The challenge:

#### Estimate this model despite covariate censoring.

#### How do we overcome censored covariates?

Using conditional mean imputation<sup>1</sup>, we can replace censored  $T_i$  with  $T_i^*$ :

 $T_i^* = \mathcal{E}(T_i | T_i > C_i, X_i).$ 

However, this method can produce biased parameter estimates.

We reduce this bias using a measurement error approach:

 $T = T^* + U.$ 

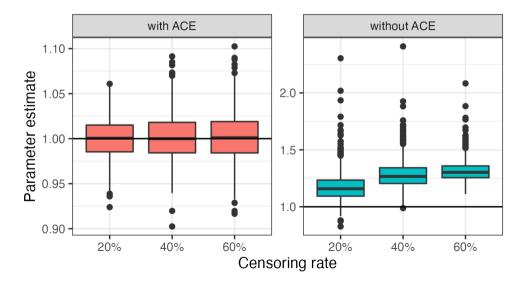
We allow U to follow any distribution  $f_U(\cdot)$ .

<sup>&</sup>lt;sup>1</sup>Atem, F. D., Sampene, E., & Greene, T. J. (2019). Improved conditional imputation for linear regression with a randomly censored predictor. *Statistical methods in medical research*, 28(2), 432-444.

We dub this method:

Actively correcting for error in imputation, or, ACE Imputation.

#### How does ACE perform in simulation?



#### What do we do next?

- See how well ACE Imputation performs when the linear predictor is mis-specified in the imputation model
- Apply this method to data from an observational study of Huntington's Disease.
- Extend our work to **nonlinear models**.

#### Acknowledgements

- My wonderful PhD advisor, Dr. Tanya Garcia
- My good friend and co-author, Dr. Sarah Lotspeich
- Funding:
  - National Institutes of Environmental Health Sciences (T32ES007018)
  - National Institute of Neurological Disorders and Stroke (K01NS09934)

#### Thank you! Any questions?

kylefg95@live.unc.educonserKylewww.github.com/GrosserKyle