Right Treatment, Right Patient, Right Time, Wrong Data? Measurement Error and Dynamic Treatment Regimes

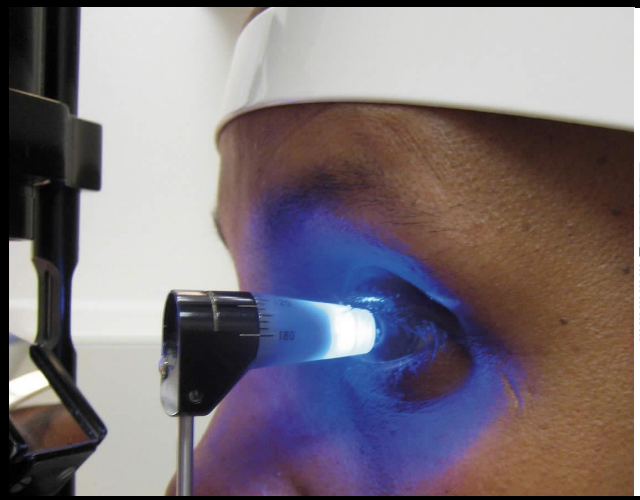
Michael Wallace, University of Waterloo

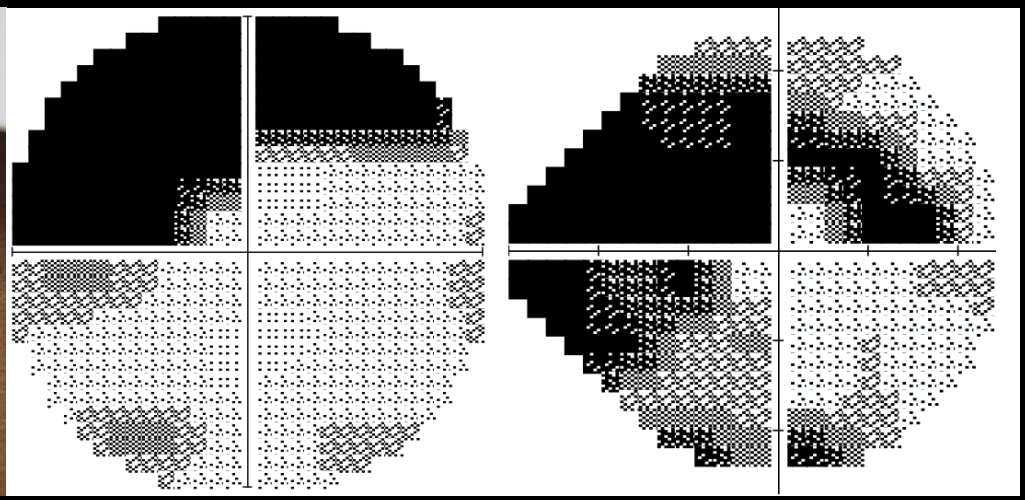
Slide deck and Shiny app links available at:
mpwallace.github.io

Glaucoma: One Disease, Many Treatments

Glaucoma: group of eye diseases associated with elevated intraocular pressure (IOP).

Elevated IOP can lead to vision loss.





Treatment via eye drops: when to switch to/add a new medication?

Example: Patient is currently taking one eye drop (Azarga). A personalized treatment rule could be:

> "*If current IOP exceeds 15, add second eye drop (Alphagan), otherwise continue with only current eye drop.*"

- Question: How do we choose the best decision rule?

  Should our IOP cut-off be 13, 15, 20?

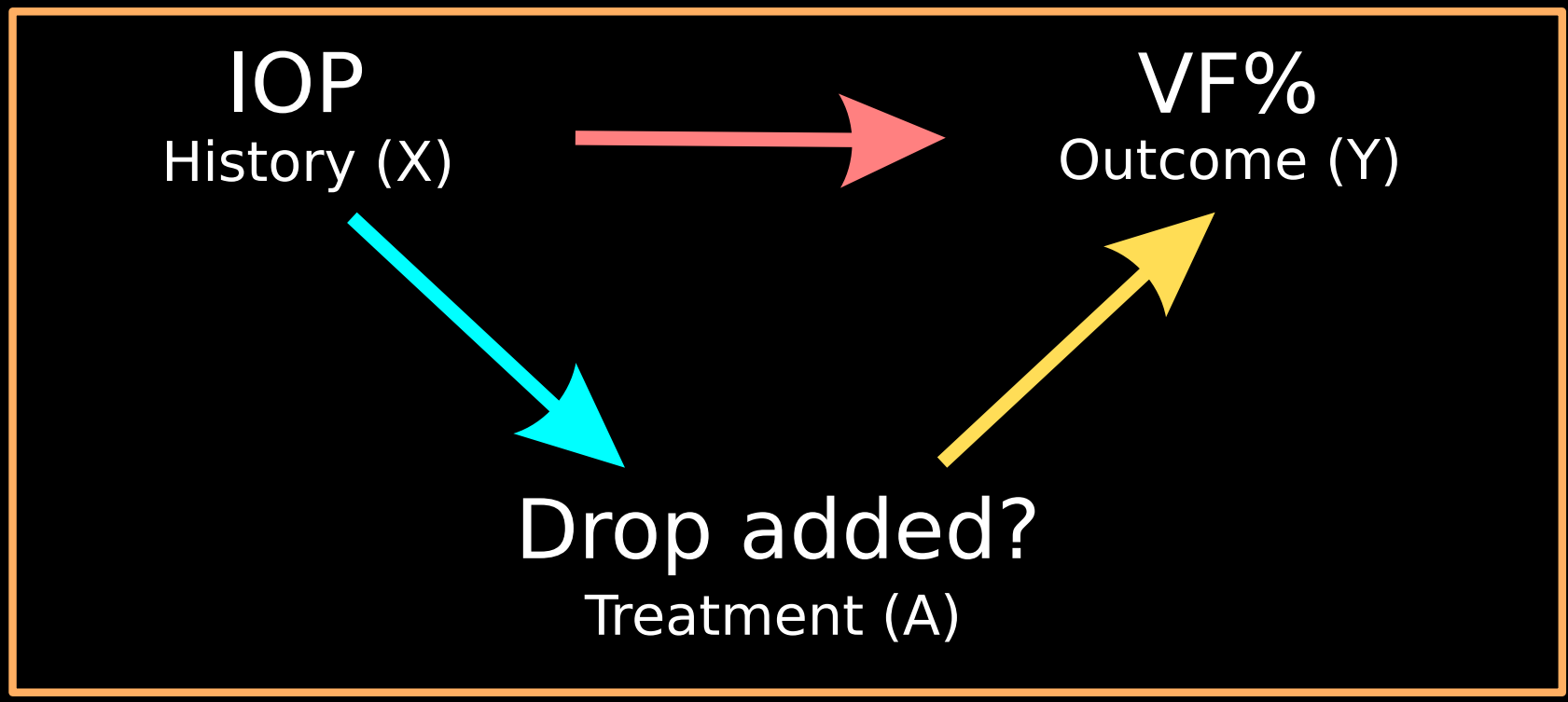
The Data

Some hypothetical data:

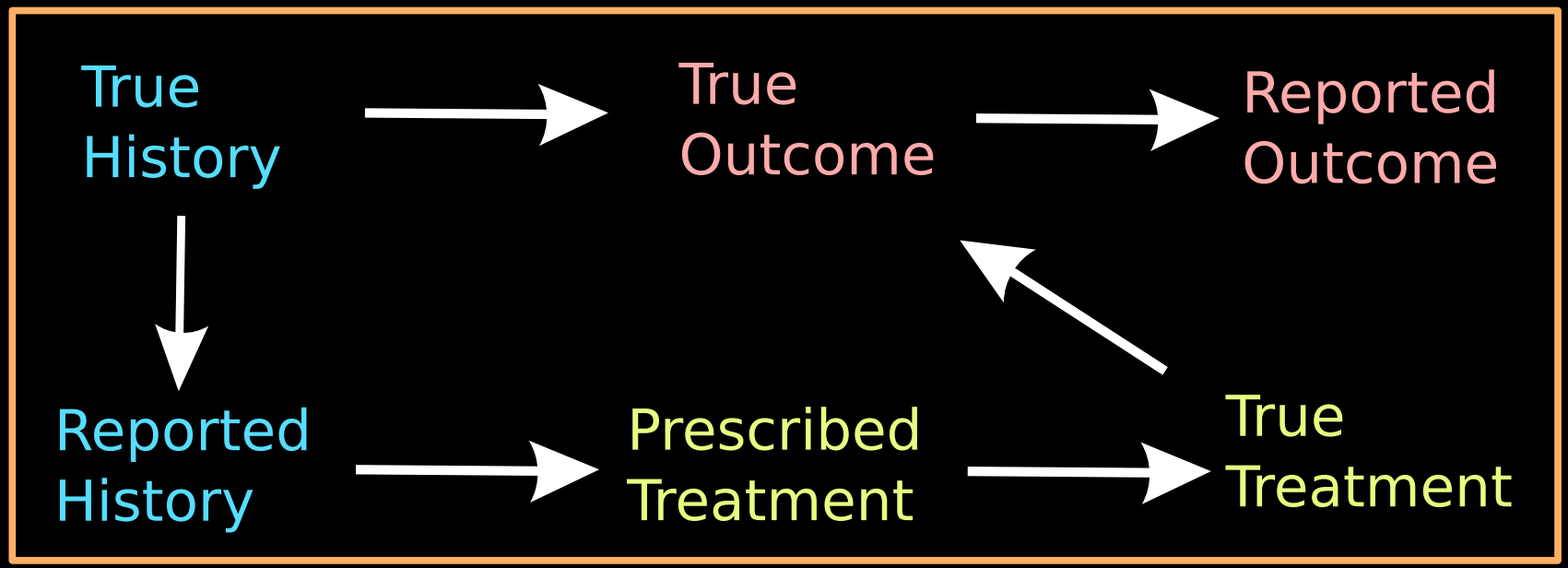
Patient	Observed IOP	Drop added?	VF% at 3 months
1	16	No	73
2	20	Yes	55
3	21	Yes	50
4	16	Yes	61
5	15	No	42
...	...	...	...

VF% = Visual Field Percentage

Question: How do these variates relate?



Goal: Identify treatment A that optimizes $E[Y|X, A]$



Problem: Measurement error

Identifying the best treatment regime

$$\underbrace{E[Y|X,A]}_{\text{Expected outcome (to be maximized)}} \qquad A \in \{0,1\}$$

- We might propose the following model

$$E[Y|X,A;\beta,\psi] = \beta_0 + \beta_1 \text{IOP} + A(\psi_0 + \psi_1 \text{IOP})$$

"Add drop ($A = 1$) if $\psi_0 + \psi_1 \text{IOP} > 0$"

- More generally:

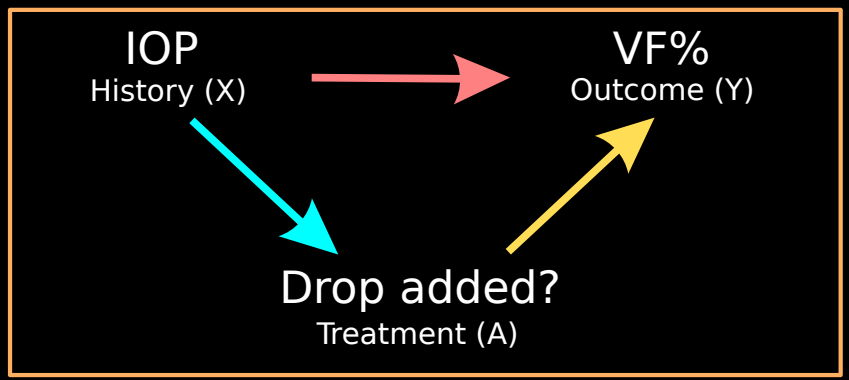
$$\overbrace{E[Y|X,A;\beta,\psi]}^{\text{Expected outcome (to be maximized)}} = \overbrace{G(X;\beta)}^{\text{Treatment-free}} + \overbrace{\gamma(X,A;\psi)}^{\text{Blip}}$$

- Simplifies focus: choose A that maximizes $\gamma(X,A;\psi)$.
- But: what if treatment-free model mis-specified?

Dynamic WOLS (dWOLS)

$$E[Y|X, A; \beta, \psi] = G(X; \beta) + \gamma(X, A; \psi)$$

- We specify a third model, the treatment model:
  1. Treatment-free model: $G(X; \beta)$.
  2. Blip model: $\gamma(X, A; \psi)$.
  3. Treatment model: $P(A = 1|X; \alpha)$.
- Estimate ψ via WOLS of Y on covariates in blip and treatment-free models, with weights $w = |A - P(A = 1|X; \hat{\alpha})|$.



Identifying the best treatment regime

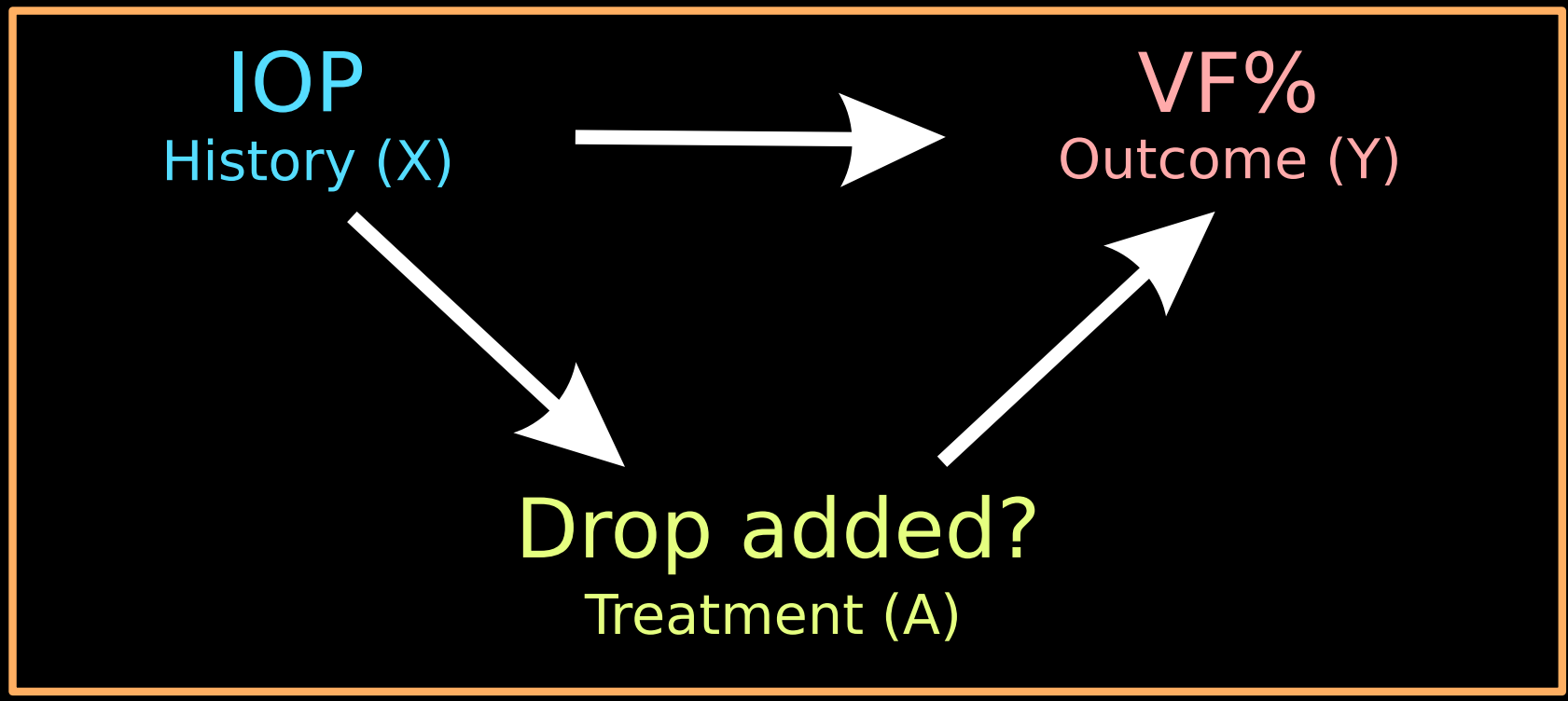
- Suppose the true outcome model is:

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 \text{IOP} + \beta_2 \text{IOP}^2 + A(\psi_0 + \psi_1 \text{IOP})$$

- Option 1: OLS with correctly specified outcome model.
- Option 2: Mis-specify treatment-free model, e.g.:

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 \text{IOP} + A(\psi_0 + \psi_1 \text{IOP})$$

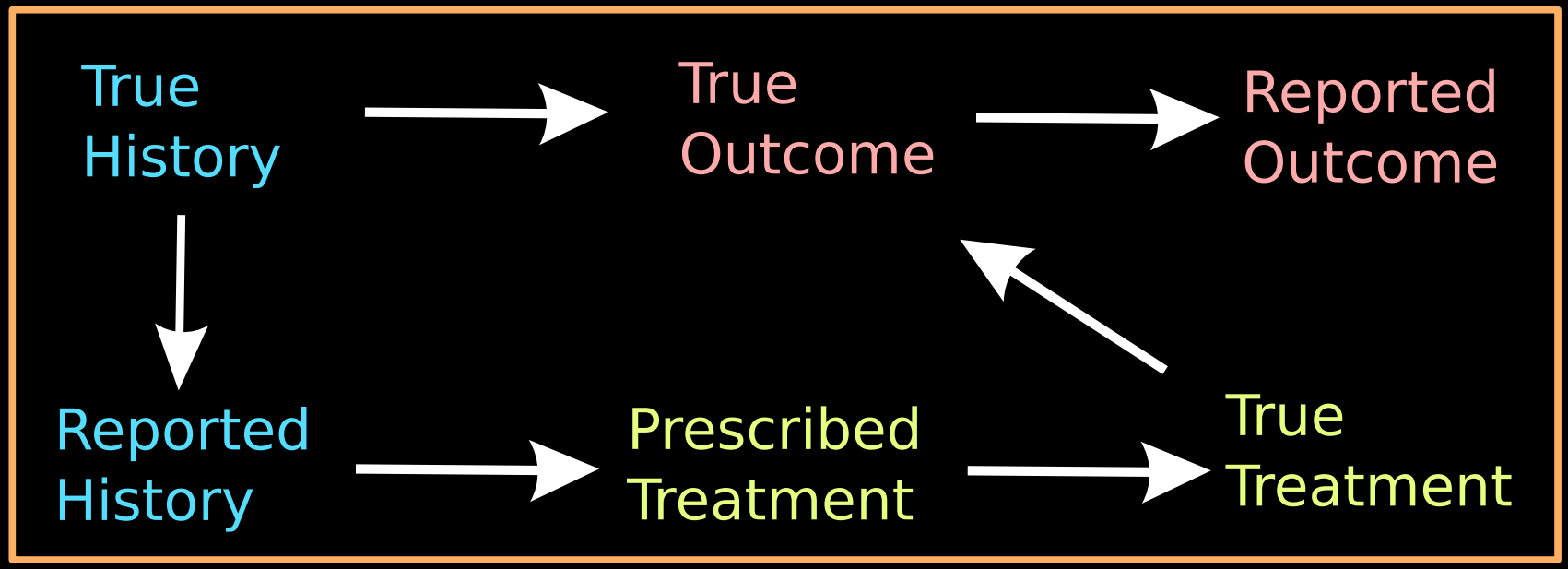
  and perform WOLS with weights $w = |A - P(A = 1|X; \hat{\alpha})|$ will still yield consistent estimators of ψ_0, ψ_1.
- Estimators are "doubly robust": consistent if at least one of treatment-free or treatment components correctly specified.
- Key point: Our treatment decisions depend only on ψ in a correctly specified blip.



Suppose the true outcome model is (e.g.):

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$

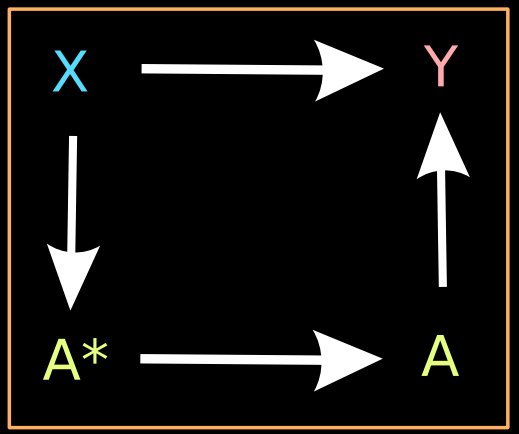
dWOLS gives doubly-robust estimators of ψ_0, ψ_1.



Suppose the true outcome model is (e.g.):

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$

What if instead of X, A, Y we observe error-prone X^*, A^*, Y^*?



Suppose the true outcome model is:

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$

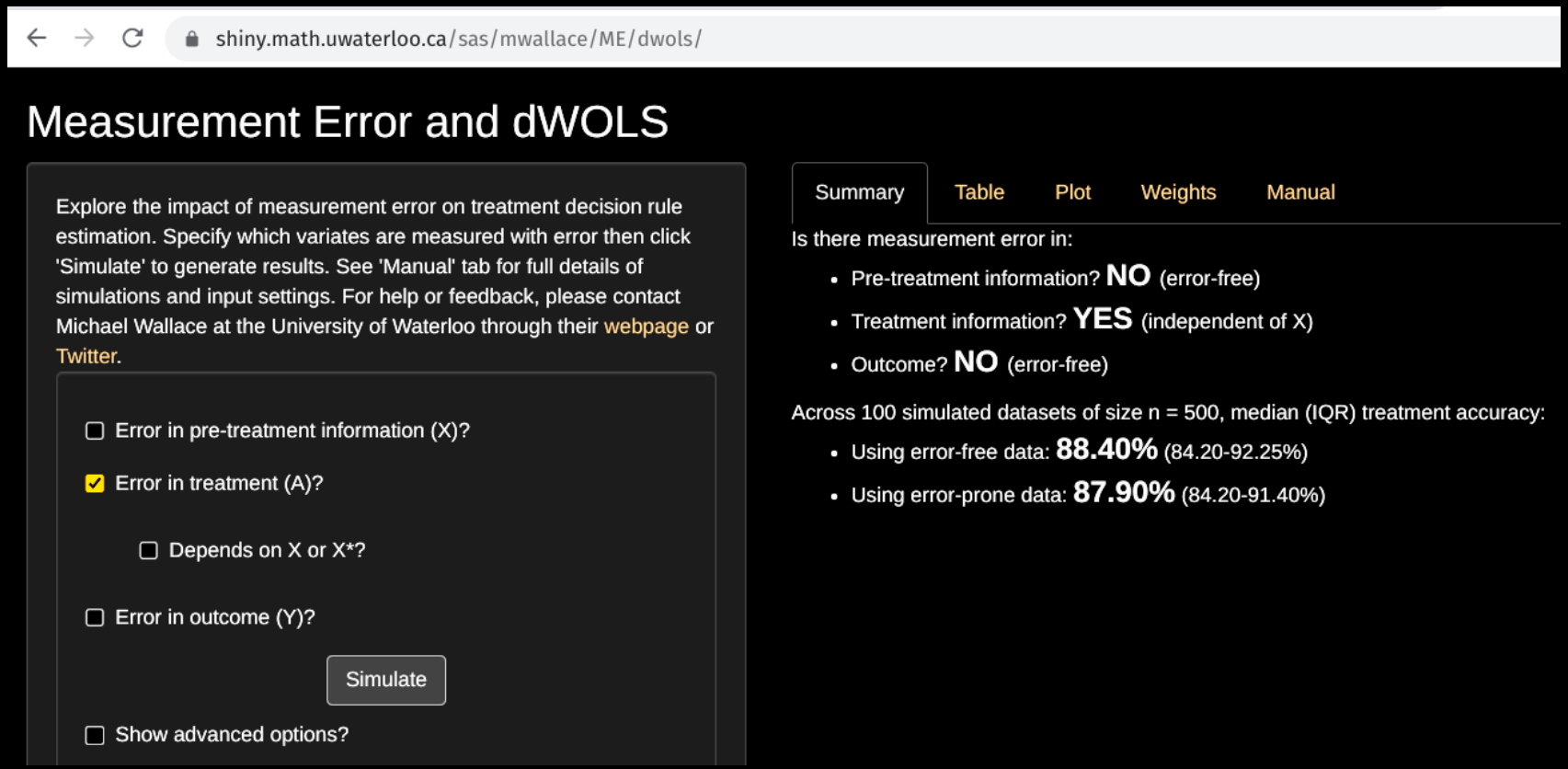
but we observe an error-prone A^*.

For binary A, misclassification can be characterized by the positive and negative predictive values:

$$PPV = P(A = 1|A^* = 1) \qquad NPV = P(A = 0|A^* = 0)$$

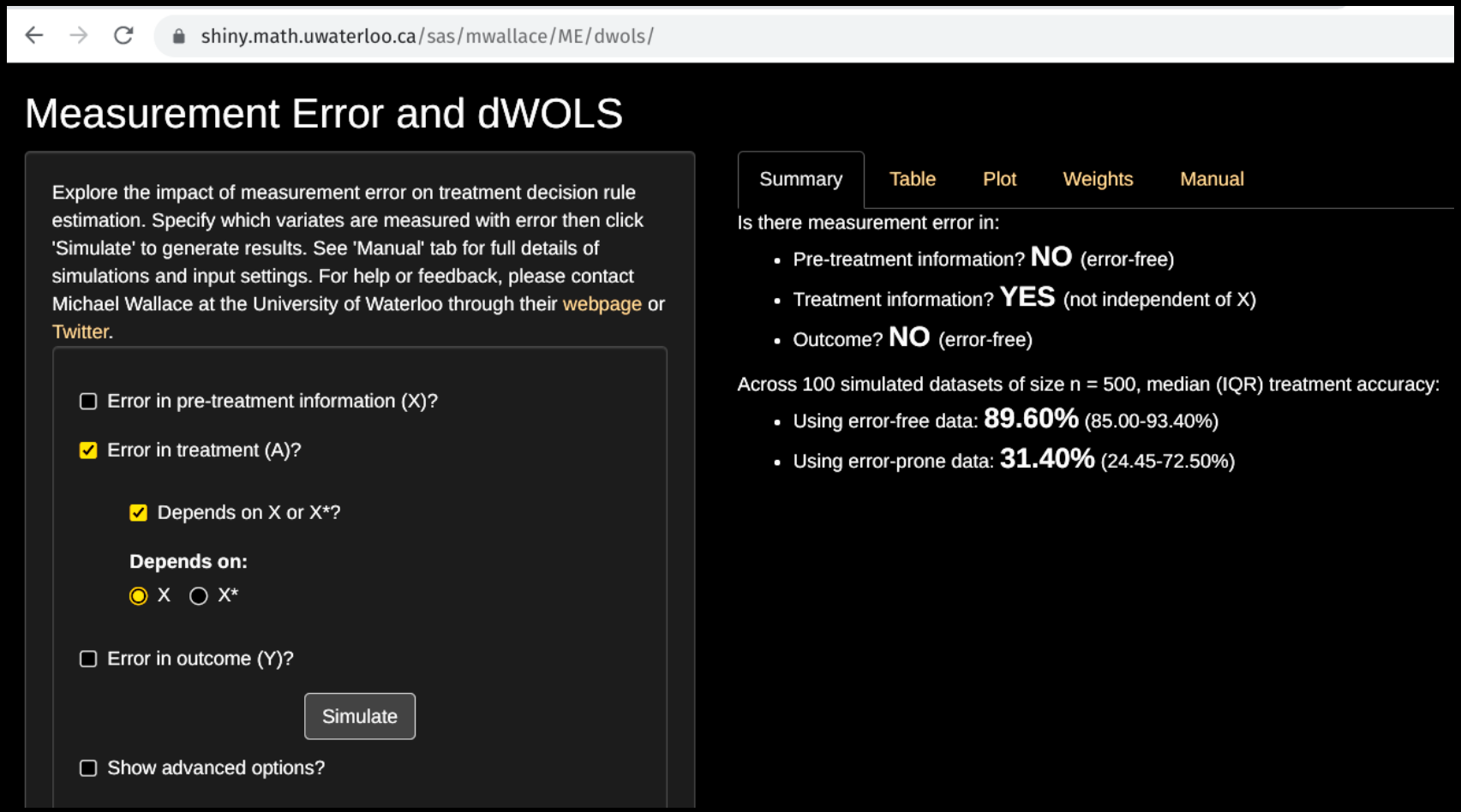
Key Question: Do these depend on X or X^*?

Shiny App: Error in A



All links available at `https://mpwallace.github.io/`

Shiny App: Error in A



All links available at `https://mpwallace.github.io/`

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$

If misclassification does not depend on X, then our estimates of ψ_0, ψ_1 will be biased:

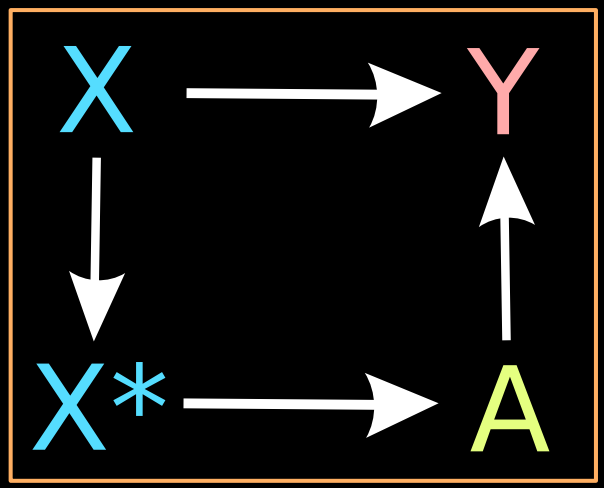
$$\psi_0^* = (PPV + NPV - 1)\psi_0 \qquad \psi_1^* = (PPV + NPV - 1)\psi_1$$

However: our treatment rule is of the form

$$A = 1 \text{ if } \psi_0 + \psi_1 X > 0$$

which is unaffected if ψ_0, ψ_1 are biased by the same factor.

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X) \qquad X^* = X + U$$



Summary | Table | Plot | Weights | Manual

Is there measurement error in:

- Pre-treatment information? **YES** (error-prone)
- Treatment information? **NO** (error-free)
- Outcome? **NO** (error-free)

Across 100 simulated datasets of size n = 500, median (IQR) treatment accuracy:

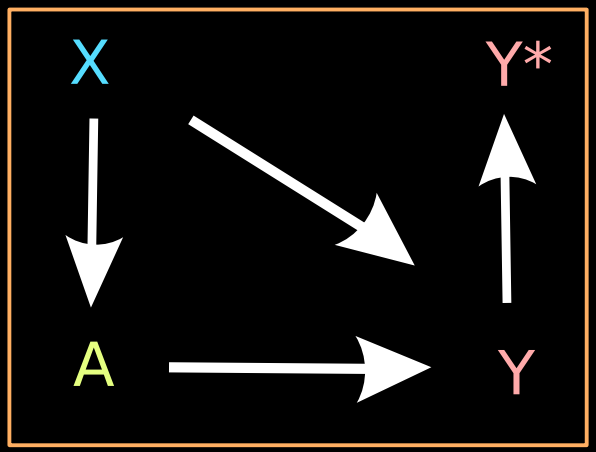
- Using error-free data: **88.10%** (84.95-91.85%)
- Using error-prone data: **82.20%** (80.80-83.40%)

Error in X leads to mis-estimated treatment rules.

Result: Regression calibration can be used with dWOLS and preserves double robustness.

Error in Y

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X) \qquad Y^* = Y + U$$



Error in Y^* independent of A:

Across 100 simulated datasets of size n = 500, median (IQR) treatment accuracy:

- Using error-free data: **87.80%** (84.15-91.65%)
- Using error-prone data: **87.80%** (84.00-91.85%)

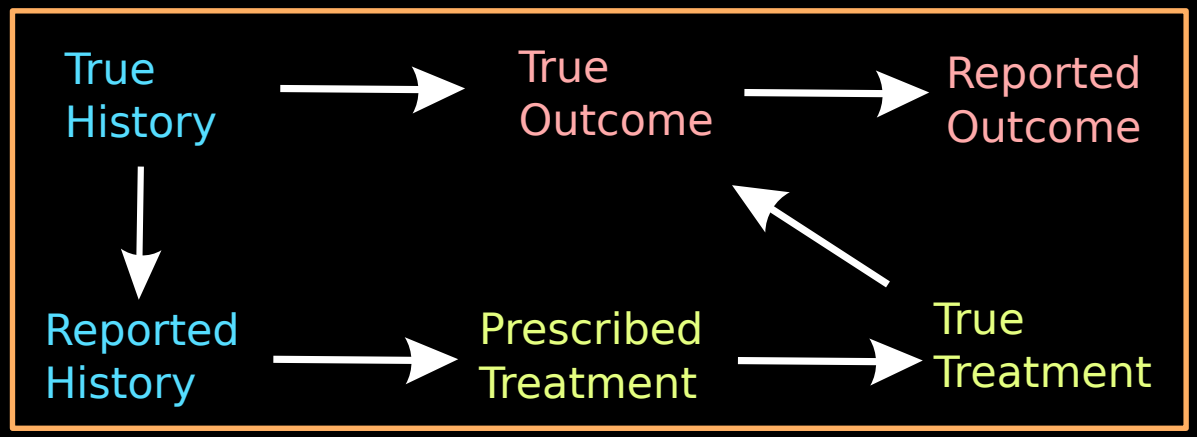
Error in Y^* not independent of A:

Across 100 simulated datasets of size n = 500, median (IQR) treatment accuracy:

- Using error-free data: **87.80%** (84.15-91.65%)
- Using error-prone data: **56.20%** (45.75-67.80%)

Where are we now?

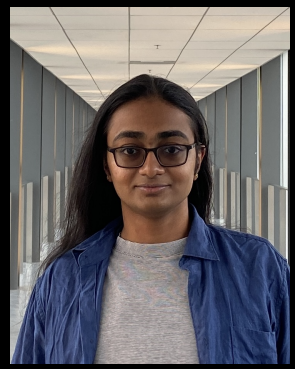
- Measurement error an important consideration in all elements of precision medicine problems.
- There are some special cases where errors have limited impact, or may be corrected for with standard theory.
- But: many more cases to explore.



Looking Ahead

- Size of error can depend on numerous factors, including sociodemographic status, symptom severity/disability.
- Larger measurement error increases probability of mis-treatment.
- When and how to account for this?
- Who should we prioritize re-measuring?

Ken Mawer



Jay Siva

- **dWOLS**: M. P. Wallace and E. E. M. Moodie (2015). Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics* **71(3)** 636-644.

- **Precision Medicine and Measurement Error in Tailoring Variates**: D. Spicker and M. P. Wallace (2020). Measurement error and precision medicine: error-prone tailoring covariates in dynamic treatment regimes. *Statistics in Medicine* **39(26)**

- **R Package DTRreg**: Available on CRAN.

- **Precision Medicine and Measurement Error More Broadly**: M. P. Wallace. Measurement error and precision medicine. In Cai T., Chakraborty B., Laber E., Moodie E. and van der Laan M. (Eds), Handbook of Statistical Methods for Precision Medicine. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. 2025.

michael.wallace@uwaterloo.ca
mpwallace.github.io

NSERC CRSNG