All else being equal: Implications of measurement error for precision medicine and health equity

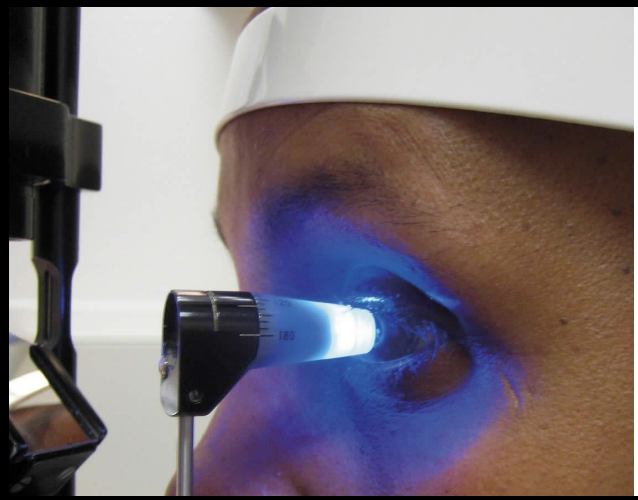
Michael Wallace, University of Waterloo

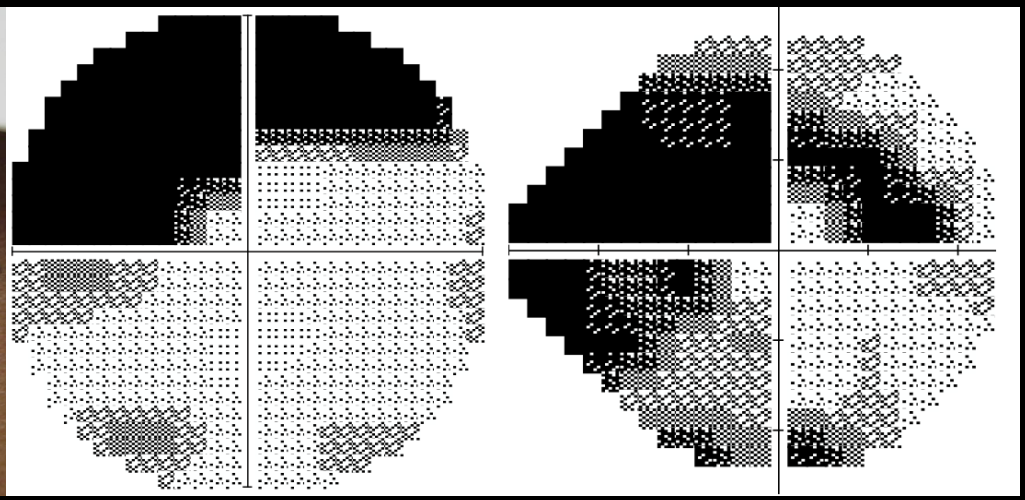
Slide deck and Shiny app links available at:
mpwallace.github.io

Glaucoma: One Disease, Many Treatments

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

Elevated IOP can lead to vision loss.





Glaucoma: One Disease, Many Treatments

Treatment options attempt to lower IOP (and by extension preserve visual field), they include:

- Lifestyle changes.
- Eye drops (numerous options).
- Surgery.

Treatment decisions based on numerous factors.

Example: Patient is currently taking Azarga eye drops. A personalized treatment rule could be:

"*If current IOP exceeds 15, add Alphagan eye drops, otherwise continue with only Azarga.*"

Example: Patient is currently taking Azarga eye drops. A personalized treatment rule could be:

"*If current IOP exceeds 15, add Alphagan eye drops, otherwise continue with only Azarga.*"

- 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

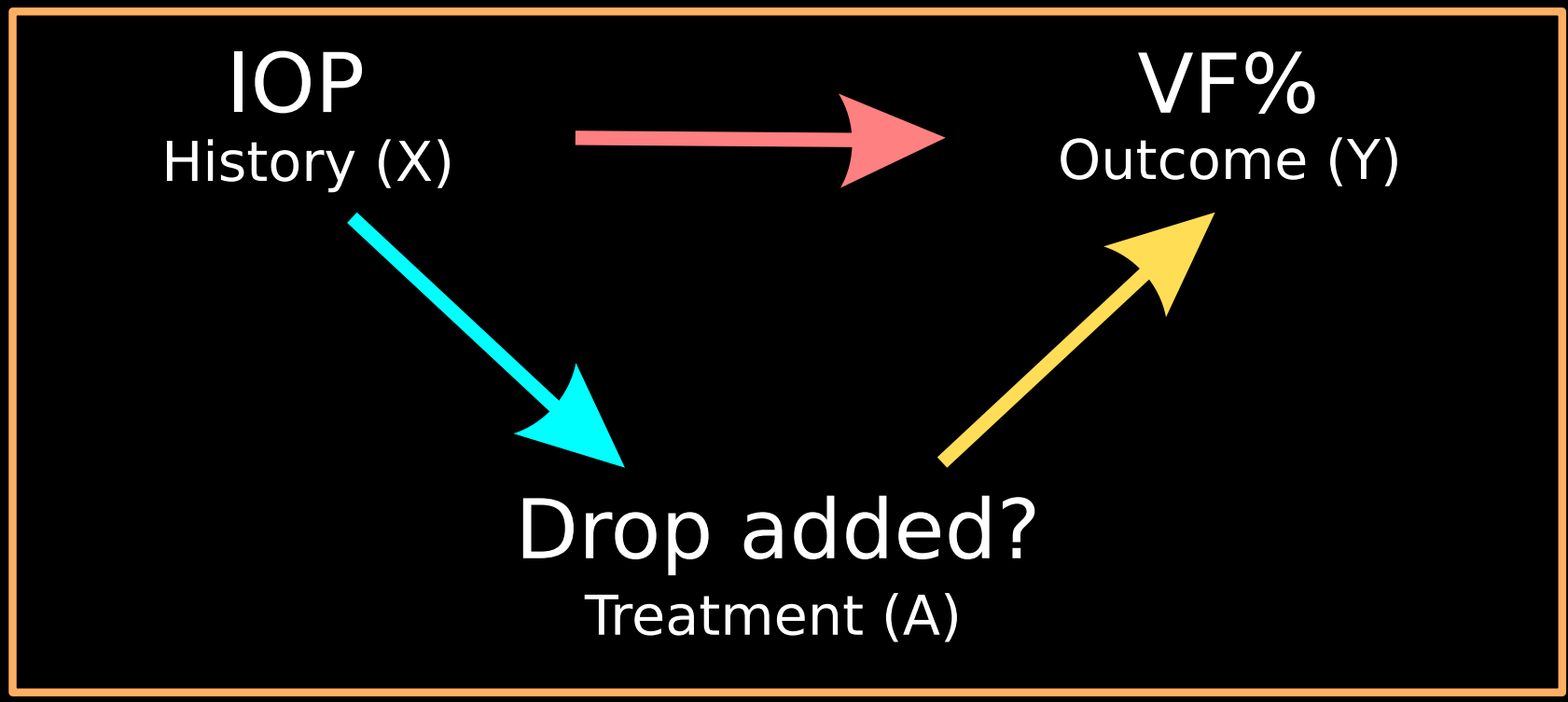
The Data

Some hypothetical data:

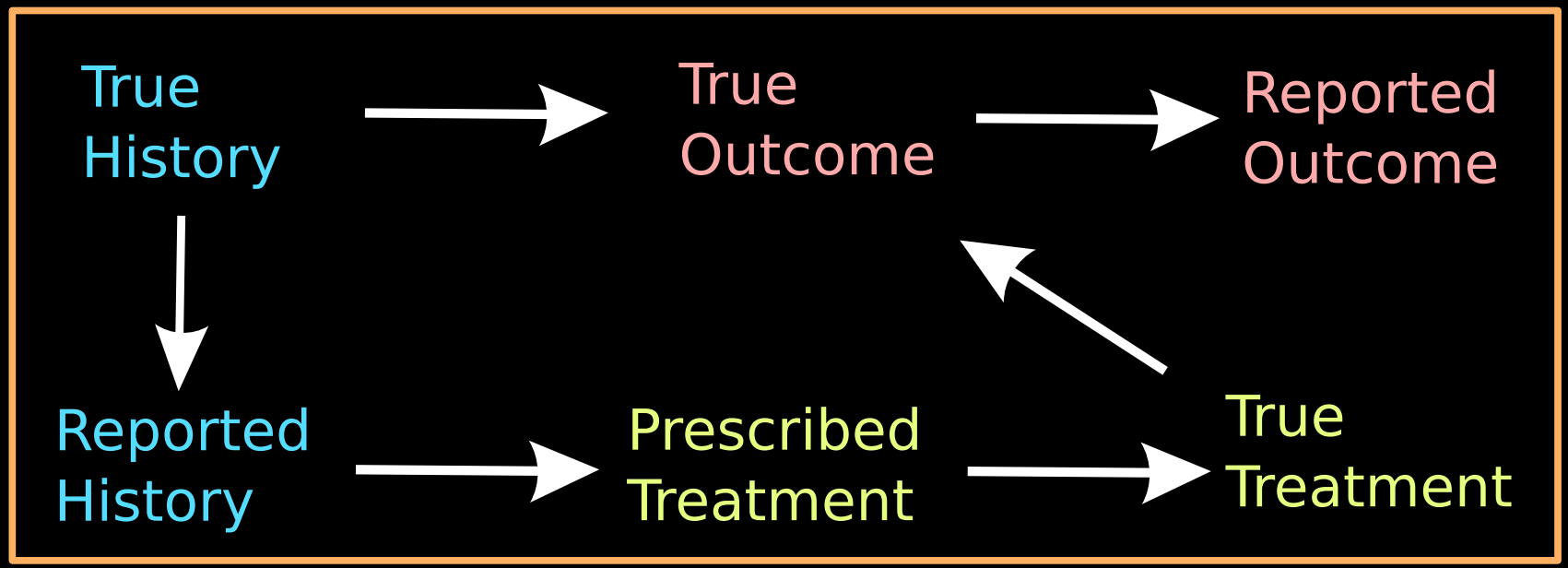
Patient	Observed IOP	Drop added?	VF% at 3 months
1	16	No	73
2	20	Yes	55
3	21	Yes	50
4	16	Yes	61
5	15	No	42
...	...	...	...

VF% = Visual Field Percentage

Question: How do these variates relate?

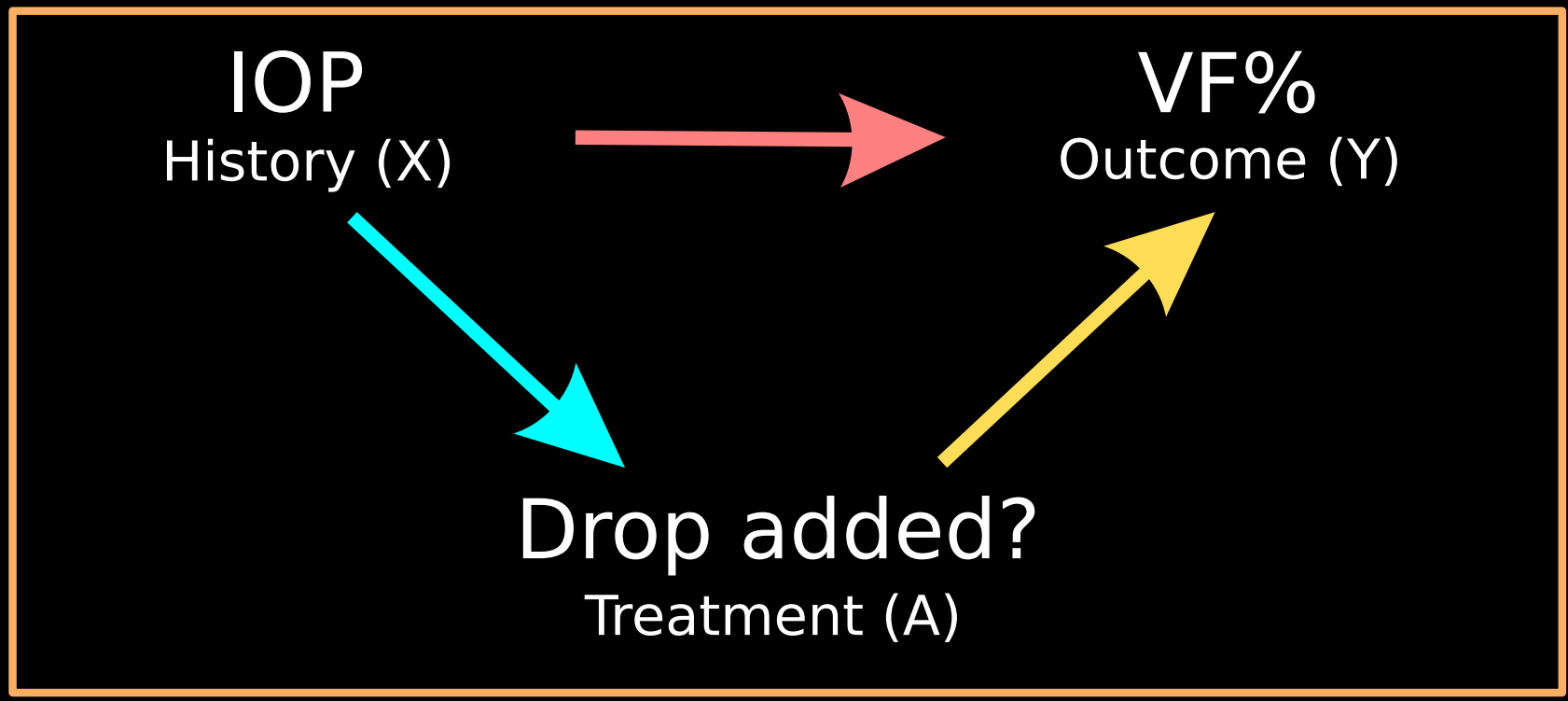


Goal: Identify treatment A that optimizes $E[Y|X, A]$



Problem: Measurement error

Data Structure: Error-free



First: error-free setting using dWOLS.

Identifying the best treatment regime

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

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

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

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 $\psi_0 + \psi_1 \text{IOP} > 0$"

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:

$$\overbrace{E[Y|X,A;\beta,\psi]}^{\text{Expected outcome (to be maximized)}} = \overbrace{G(X;\beta)}^{\text{Impact of patient history in the absence of treatment}} + \overbrace{\gamma(X,A;\psi)}^{\text{Impact of treatment on outcome}}$$

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:

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

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:

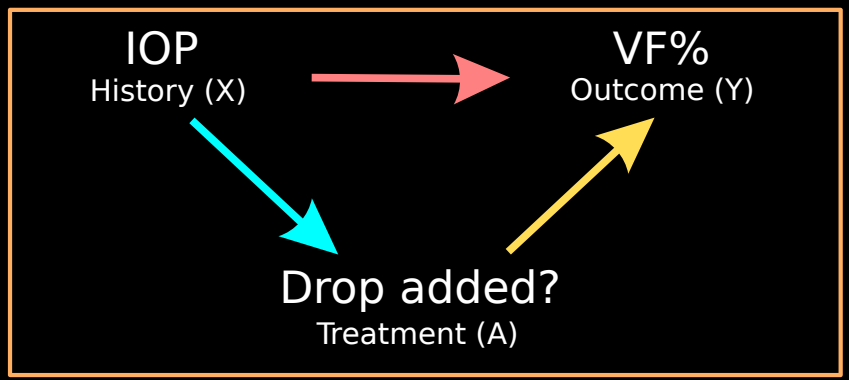
$$\overbrace{E[Y|X,A;\beta,\psi]}^{\text{Expected outcome (to be maximized)}} = \overbrace{G(X;\beta)}^{\text{Treatment-free}} + \overbrace{\gamma(X,A;\psi)}^{\text{Blip}}$$

- Simplifies focus: choose A that maximizes $\gamma(X,A;\psi)$.

Dynamic WOLS (dWOLS)

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

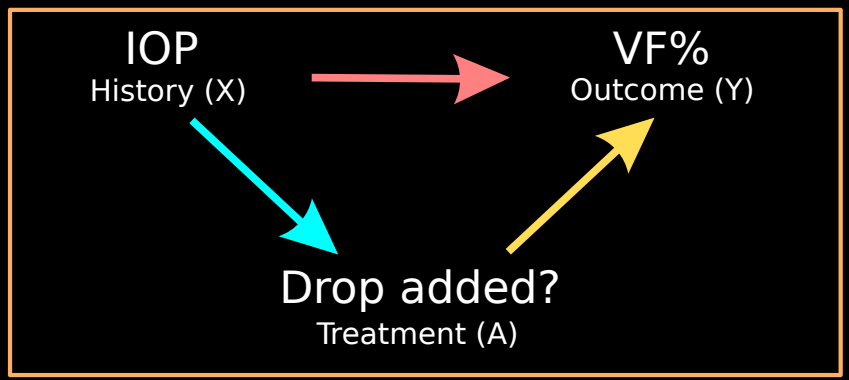
- We specify a third model, the treatment model:
  1. Treatment-free model: $G(X; \beta)$.
  2. Blip model: $\gamma(X, A; \psi)$.
  3. Treatment model: $P(A = 1|X; \alpha)$.



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:

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

Identifying the best treatment regime

- Suppose the true outcome model is:

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

- But we propose:

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

Identifying the best treatment regime

- Suppose the true outcome model is:

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

- But we propose:

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

- WOLS with weights $w = |A - P(A = 1|X; \hat{\alpha})|$ will still yield consistent estimators of ψ_0, ψ_1.

Identifying the best treatment regime

- Suppose the true outcome model is:

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

- But we propose:

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

- 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.
- The blip must always be correct.

Identifying the best treatment regime

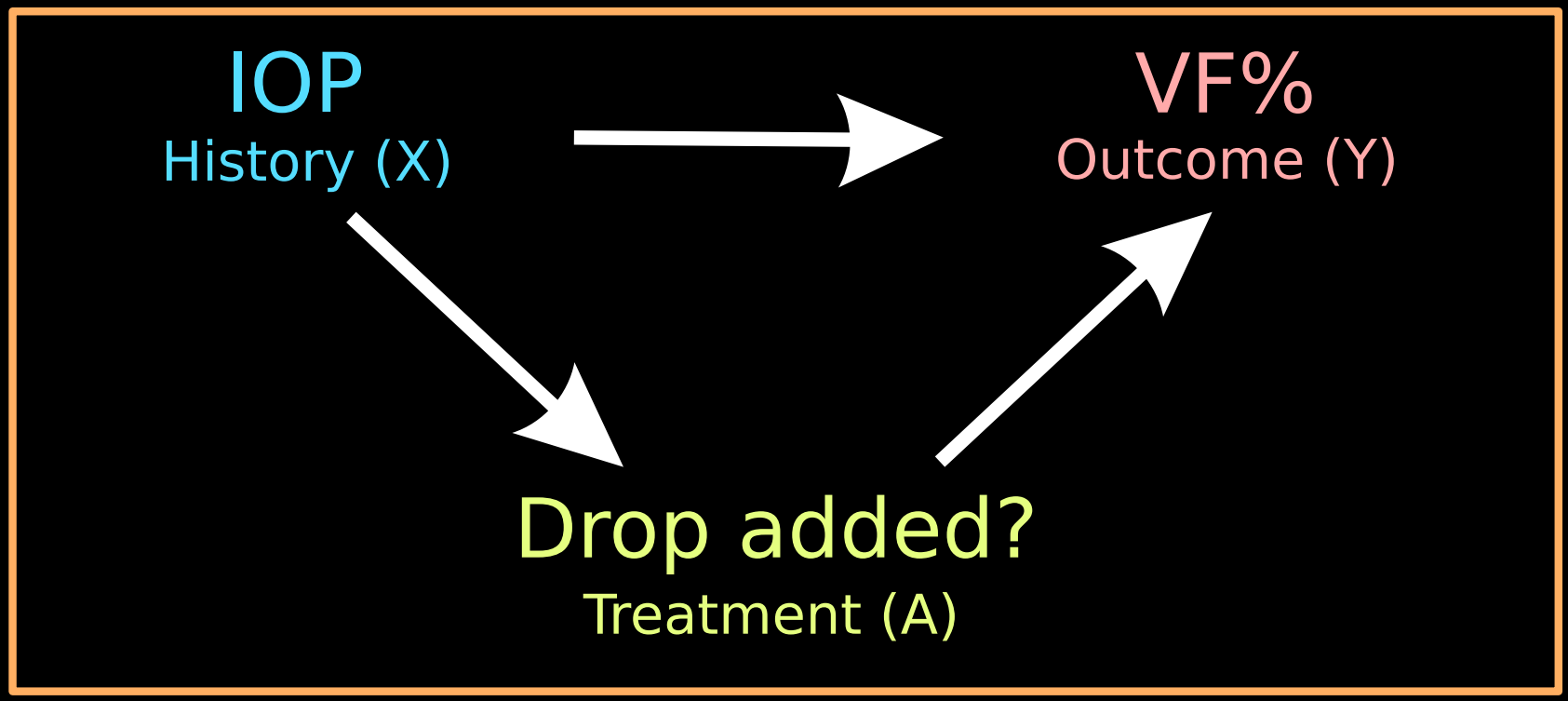
- Suppose the true outcome model is:

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

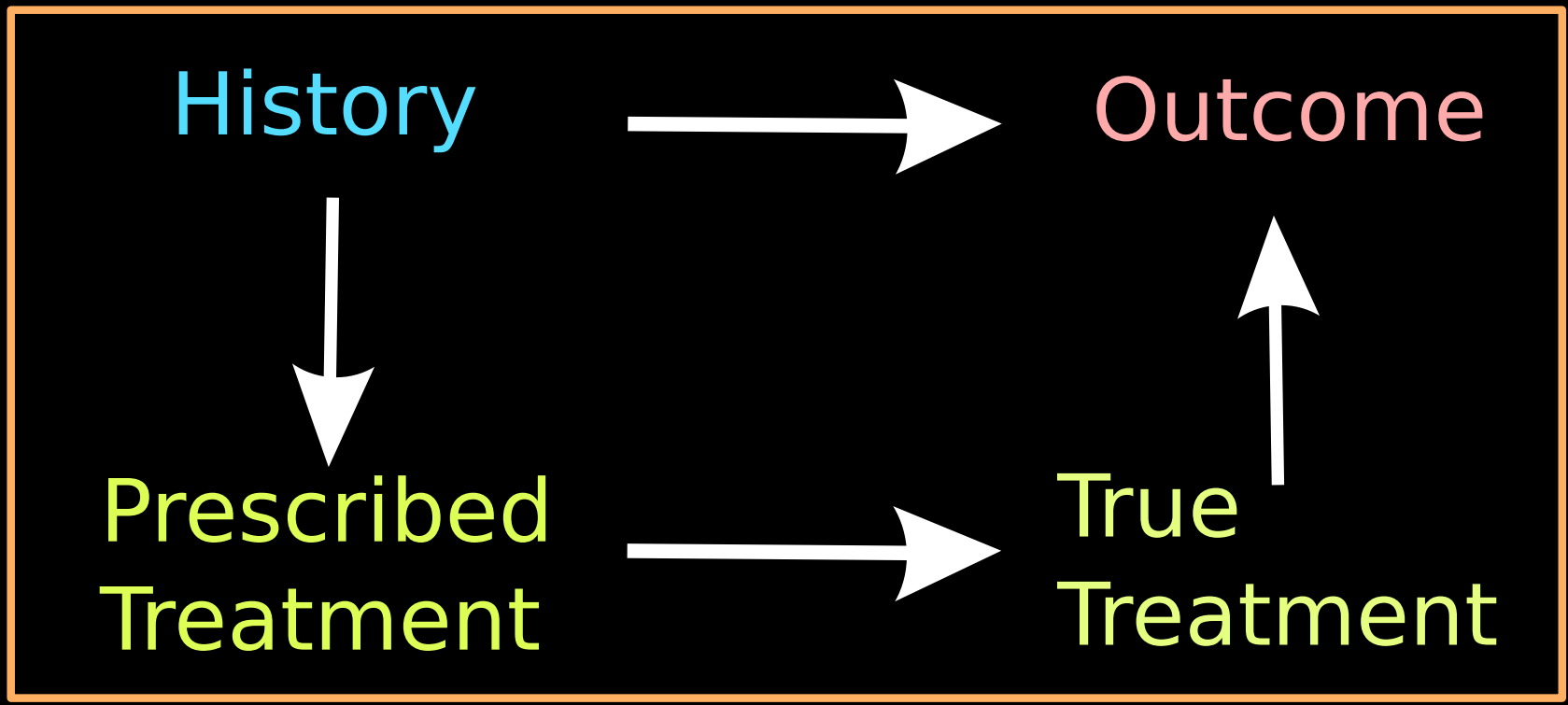
- But we propose:

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

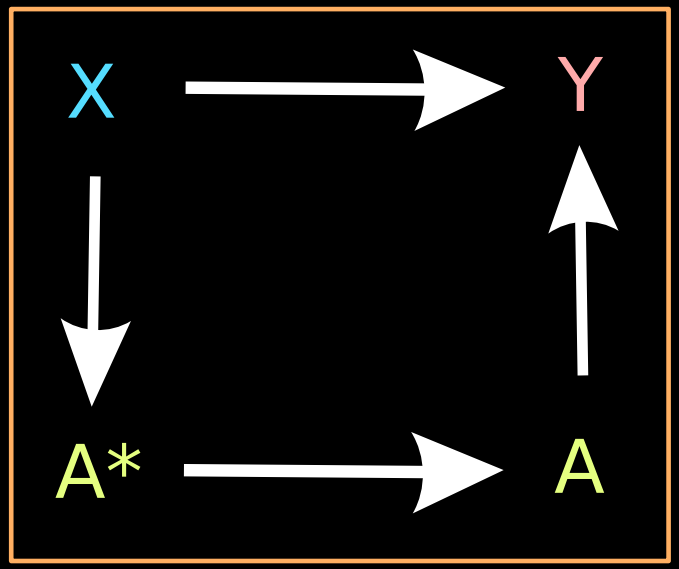
- WOLS with weights $w = |A - P(A = 1|X; \hat{\alpha})|$ will still yield consistent estimators of ψ_0, ψ_1.
- Estimators are "doubly robust": consistent if at least one of treatment-free or treatment components correctly specified.
- The blip must always be correct.
- Critical point: Our treatment decisions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


IOP
History (X)
VF%
Outcome (Y)
Drop added?
Treatment (A)

Error in Treatment A



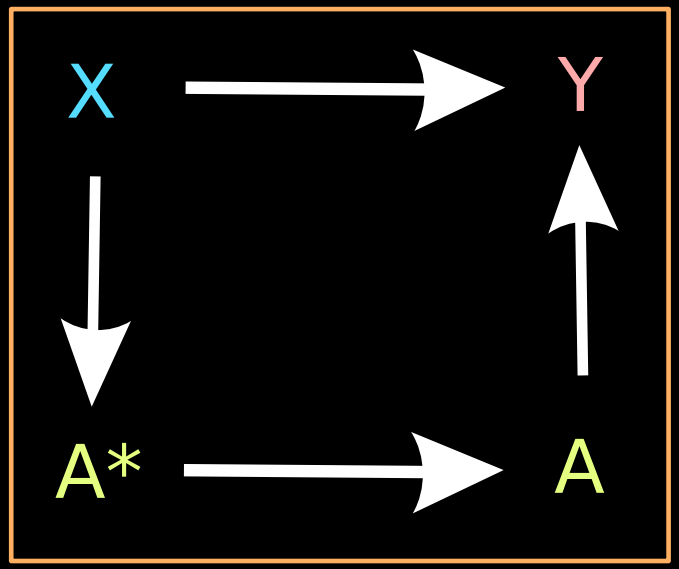
Error in Treatment A



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^*



Suppose the true outcome model is:

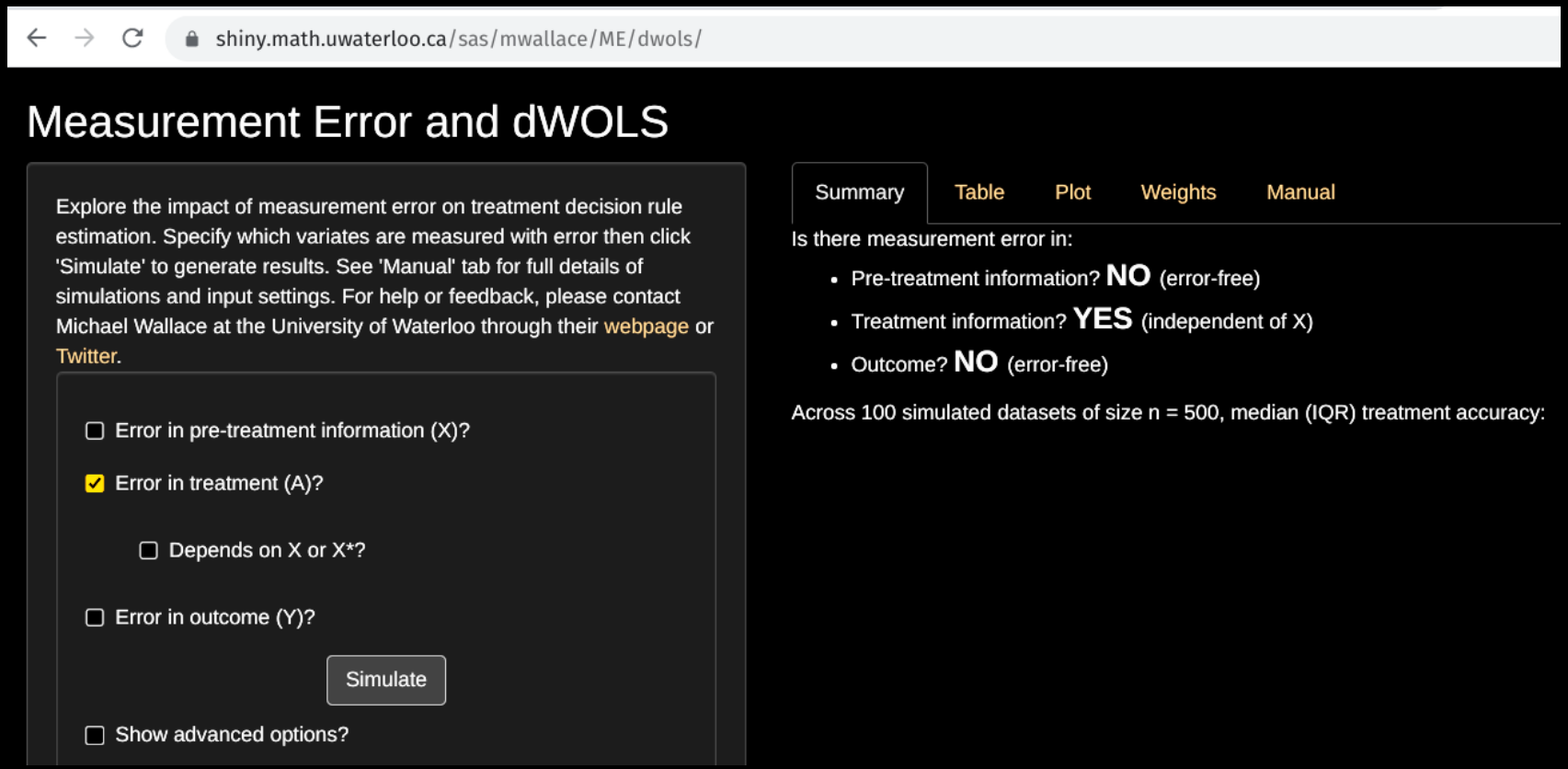
$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$

but we observe an error-prone A^*

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

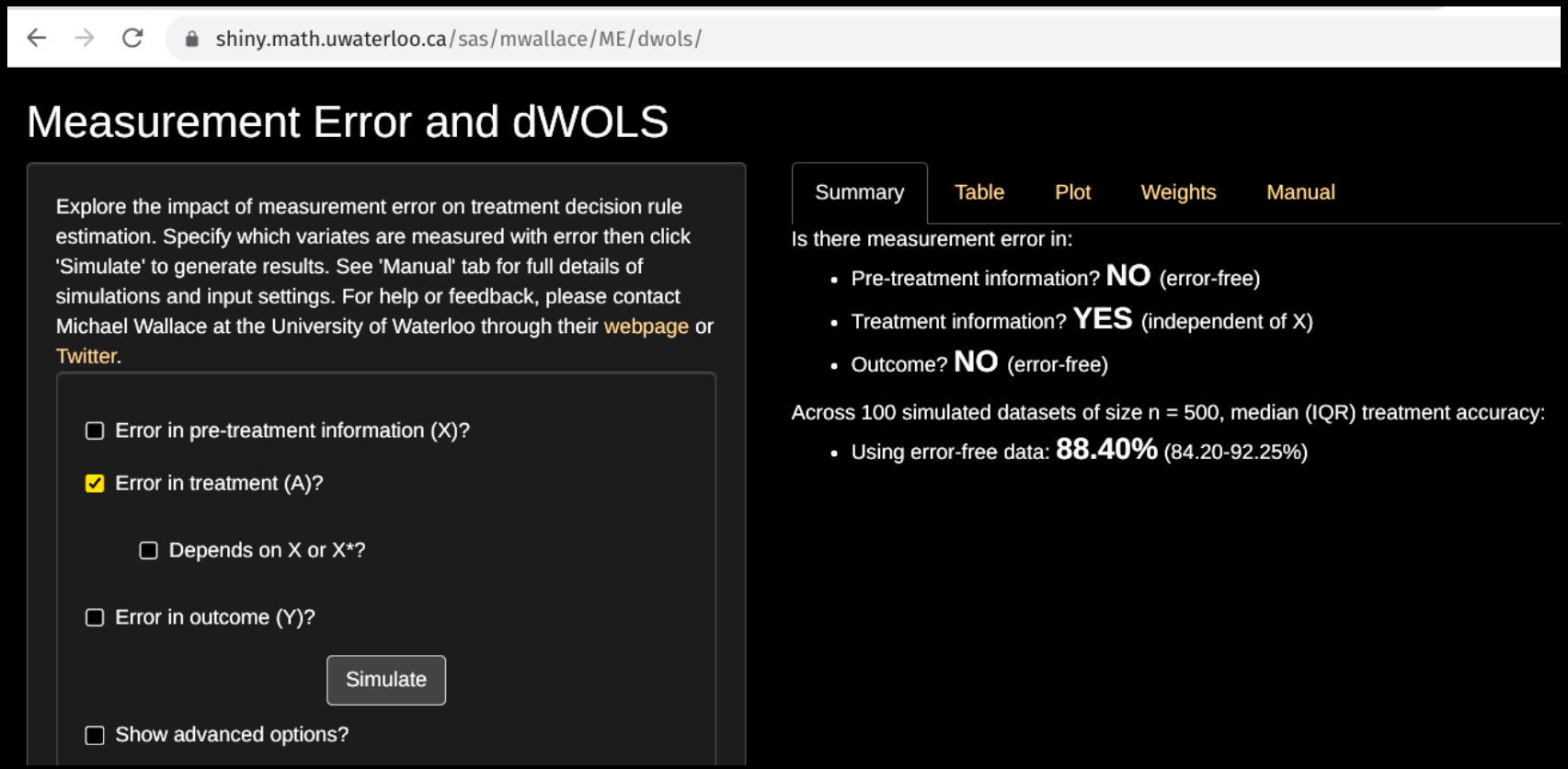
$$PPV = P(A = 1|A^* = 1) \qquad NPV = P(A = 0|A^* = 0)$$

Shiny App: Error in A



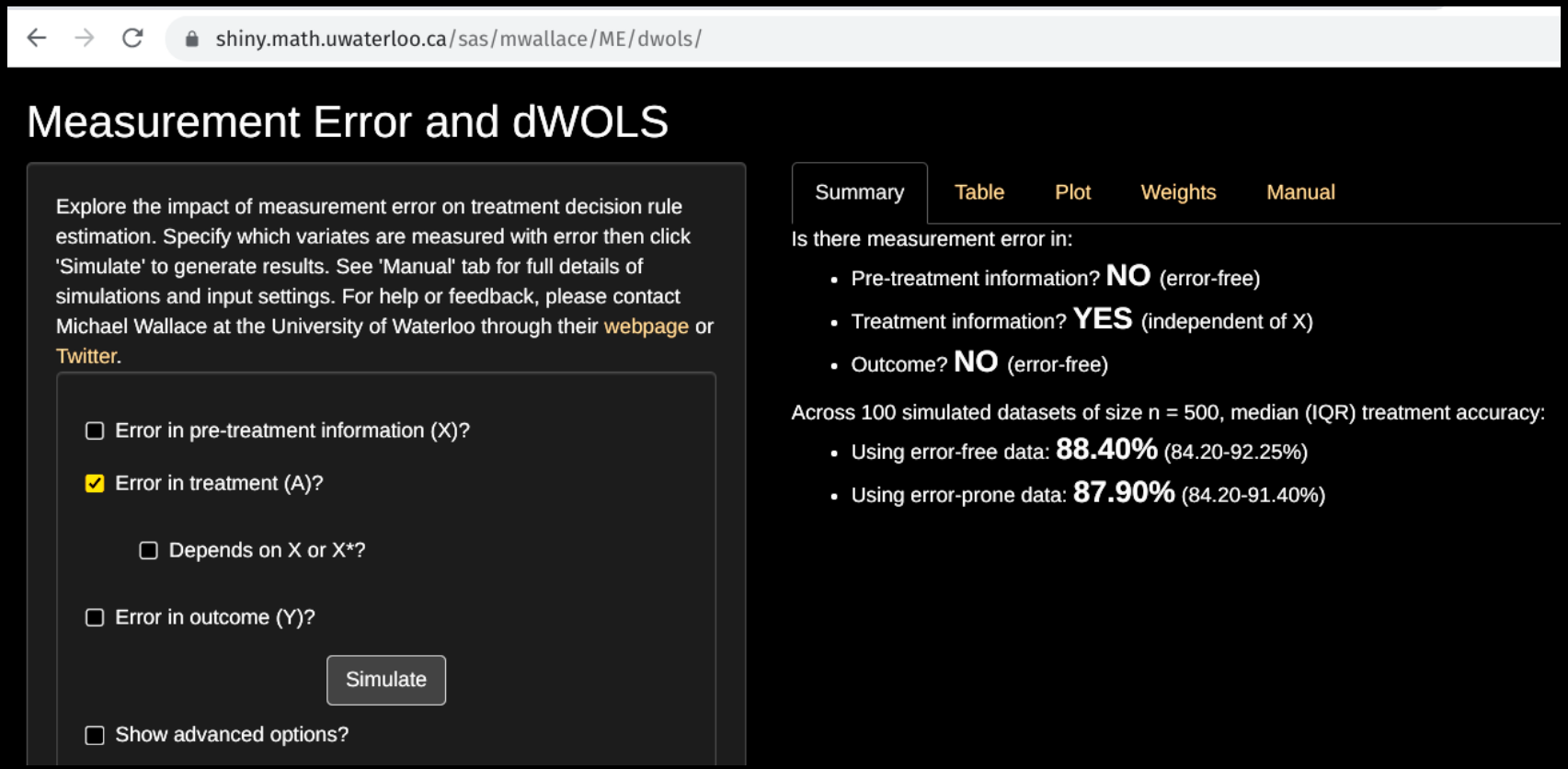
All links available at `https://mpwallace.github.io/`

Shiny App: Error in A



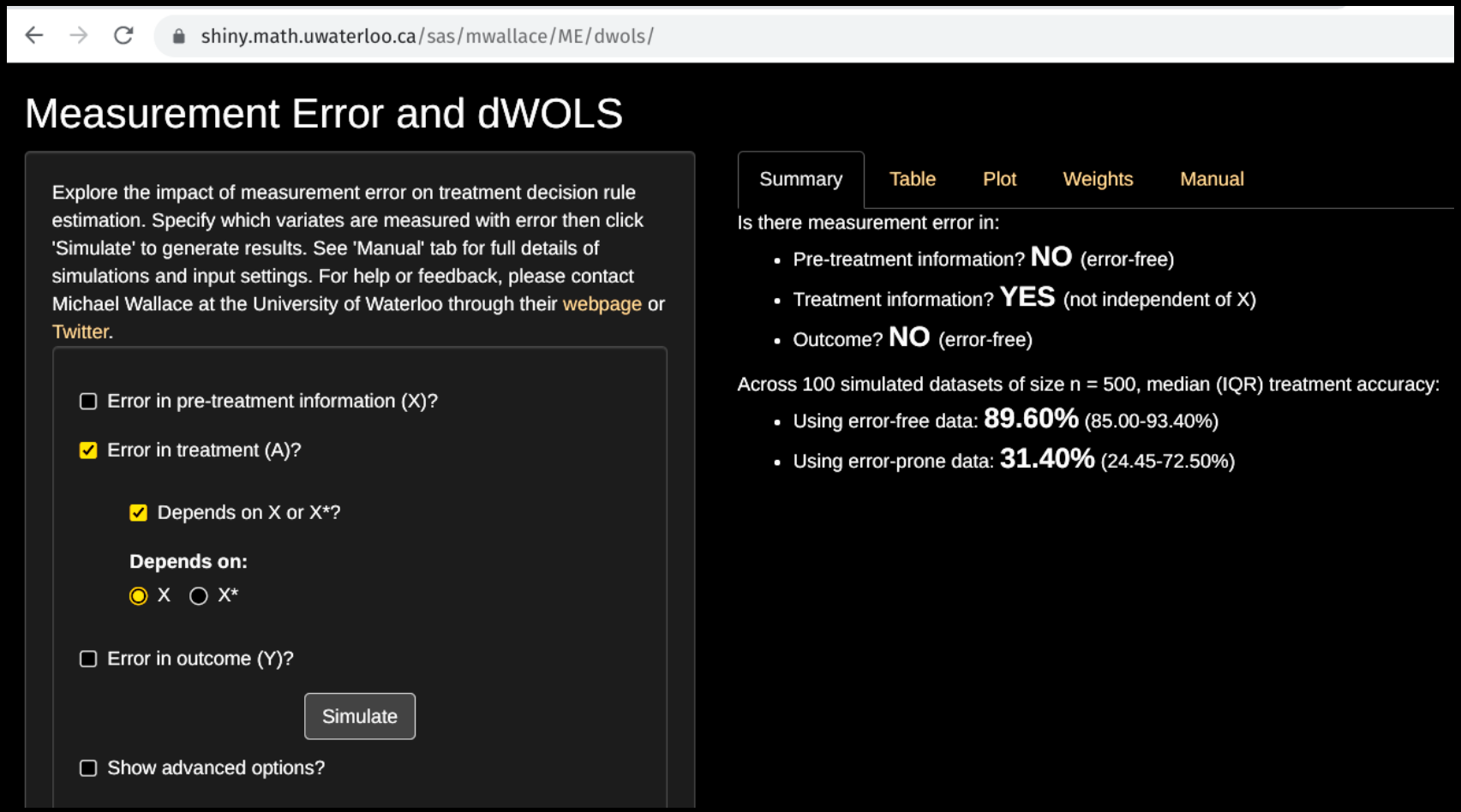
All links available at `https://mpwallace.github.io/`

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:

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

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

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.

We have $A = 1$ if $\psi_0 + \psi_1 X > 0$ or, if $\psi_1 > 0$

$$A = 1 \text{ if } X > -\frac{\psi_0}{\psi_1}$$

We call $\tau = -\frac{\psi_0}{\psi_1}$ a treatment threhsold.

(WLOG if $\psi_1 < 0$)

Looking Ahead: Future Treatment

Suppose we conclude that our treatment rule should be:

"If 3-month average IOP $>$ 15 add secondary drop, otherwise, maintain current treatment regime."

Looking Ahead: Future Treatment

Suppose we conclude that our treatment rule should be:

"If 3-month average IOP $>$ 15 add secondary drop, otherwise, maintain current treatment regime."

I go to the clinic and my IOP measurement is 16. Then what?

Looking Ahead: Future Treatment

Suppose we conclude that our treatment rule should be:

"If 3-month average IOP > 15 add secondary drop, otherwise, maintain current treatment regime."

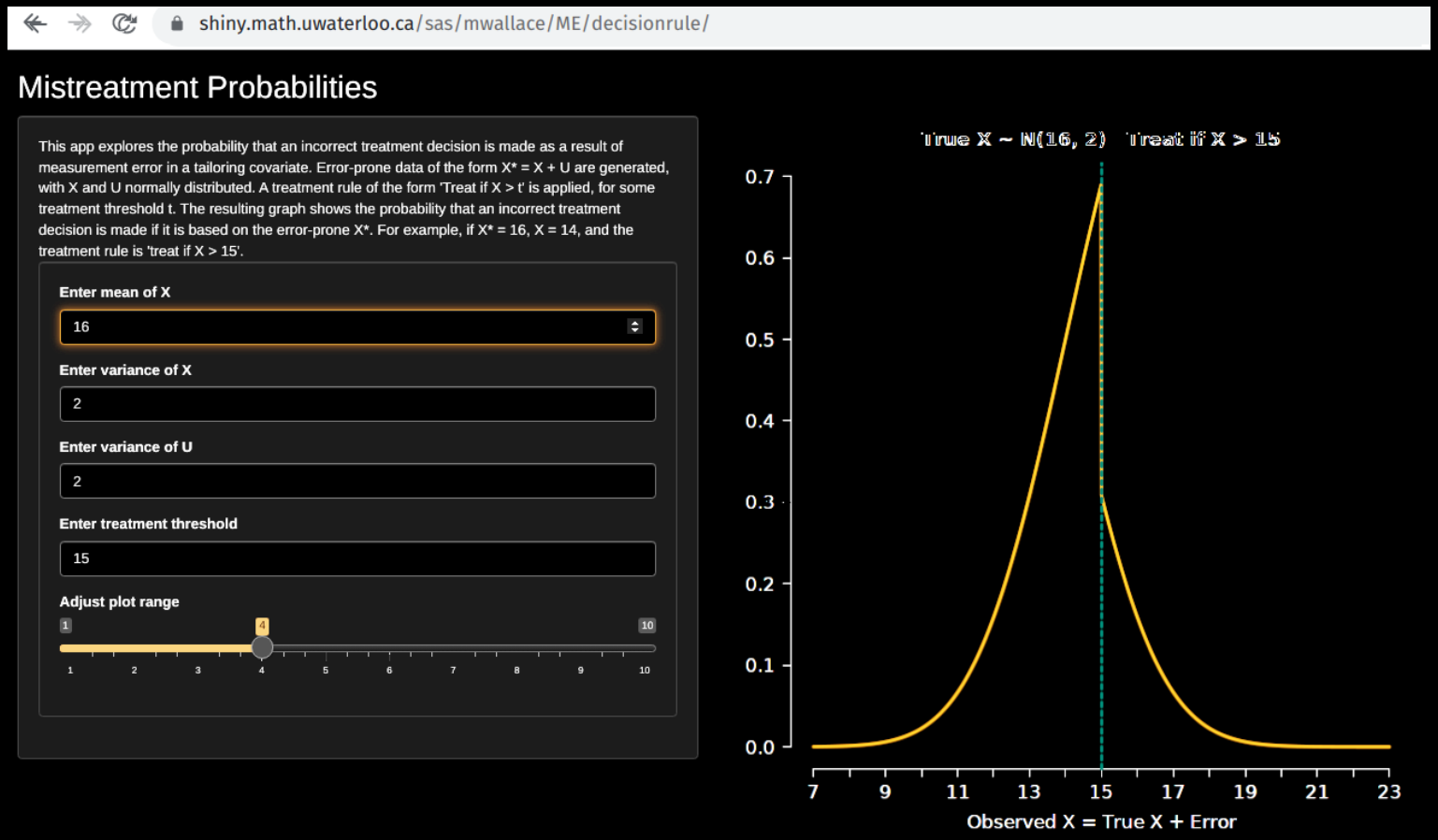
I go to the clinic and my IOP measurement is 16. Then what?

What is the probability I receive the wrong treatment?

$$P(X \leq 15 | X^* = 16)$$

Looking Ahead: Future Treatment

Exploring these probabilities through a Shiny app:



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

Looking Ahead: Health Equity

Essential consideration: What drives measurement error?

- Larger measurement error increases probability of mis-treatment.
- Size of error can depend on numerous factors, including sociodemographic status, symptom severity/disability, and tailoring variates themselves.
- Easy to show in simulation: Challenging to account for in practice.

Looking Ahead: Future Treatment

Further consideration: What if we estimate our treatment threshold based on error-prone data?

$$\begin{aligned} \text{True rule:} \quad & A = 1 \text{ if } X > \tau = 15 \\ \text{Estimated Rule:} \quad & A = 1 \text{ if } X > \tau^* = 12 \end{aligned}$$

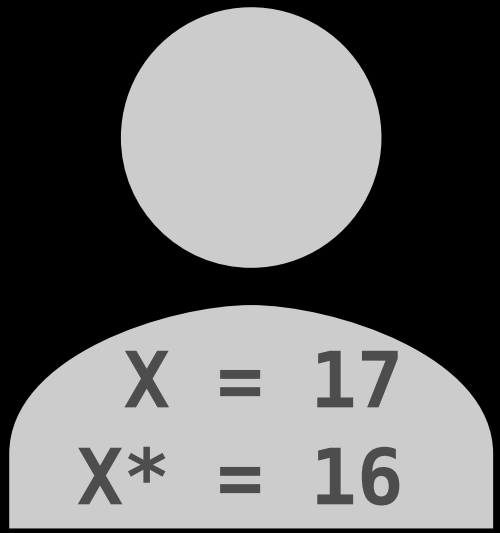
Looking Ahead: Future Treatment

Further consideration: What if we estimate our treatment threshold based on error-prone data?

$$\text{True rule:} \quad A = 1 \text{ if } X > \tau = 15$$

$$\text{Estimated Rule:} \quad A = 1 \text{ if } X > \tau^* = 12$$

Consider a patient with true $X = 17$, error-prone $X^* = 16$:



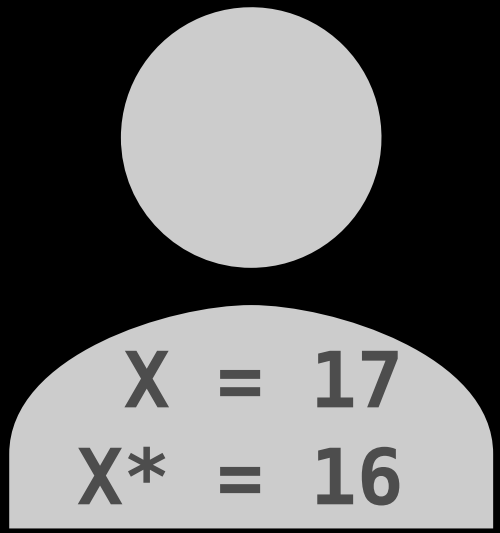
Looking Ahead: Future Treatment

Further consideration: What if we estimate our treatment threshold based on error-prone data?

$$\text{True rule:} \quad A = 1 \text{ if } X > \tau = 15$$

$$\text{Estimated Rule:} \quad A = 1 \text{ if } X > \tau^* = 12$$

Consider a patient with true $X = 17$, error-prone $X^* = 16$:



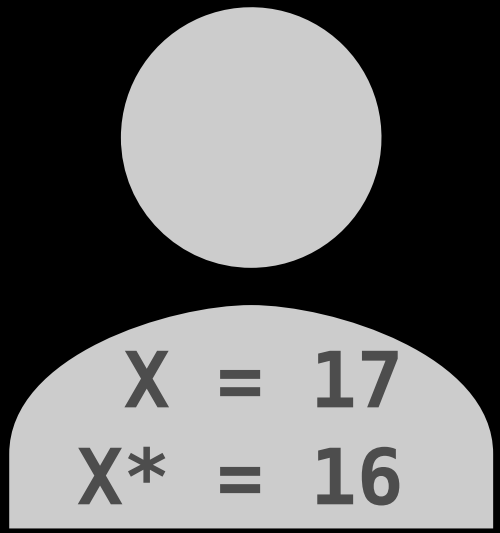
- $\boxed{X = 17,\ X > \tau = 15,\ A^{opt} = 1}$

Looking Ahead: Future Treatment

Further consideration: What if we estimate our treatment threshold based on error-prone data?

$$\text{True rule:} \quad A = 1 \text{ if } X > \tau = 15$$
$$\text{Estimated Rule:} \quad A = 1 \text{ if } X > \tau^* = 12$$

Consider a patient with true $X = 17$, error-prone $X^* = 16$:



- $\boxed{X = 17,\ X > \tau = 15,\ A^{opt} = 1}$
- $X = 17,\ X > \tau^* = 12 \implies A = 1$ ✓

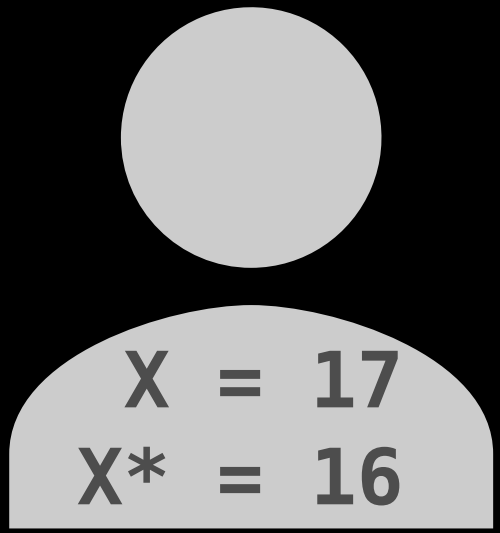
Looking Ahead: Future Treatment

Further consideration: What if we estimate our treatment threshold based on error-prone data?

True rule: $A = 1$ if $X > \tau = 15$

Estimated Rule: $A = 1$ if $X > \tau^* = 12$

Consider a patient with true $X = 17$, error-prone $X^* = 16$:



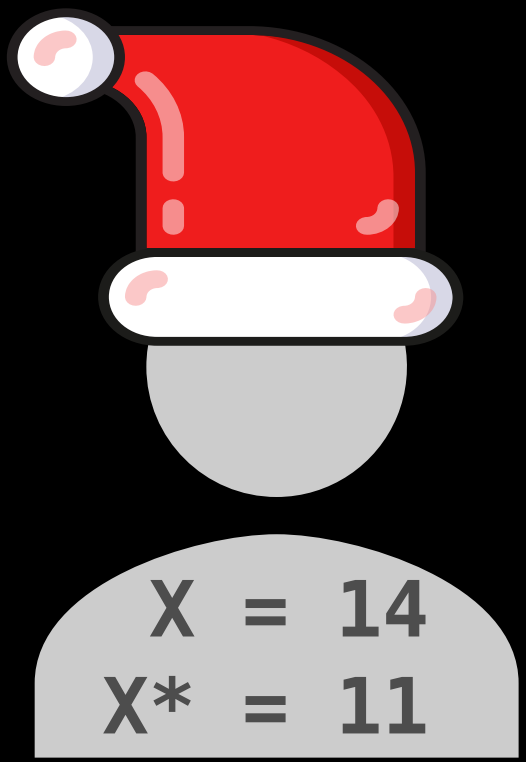
- $\boxed{X = 17,\ X > \tau = 15,\ A^{opt} = 1}$
- $X = 17,\ X > \tau^* = 12 \implies A = 1$ ✓
- $X^* = 16,\ X^* > \tau = 15 \implies A = 1$ ✓
- $X^* = 16,\ X^* > \tau^* = 12 \implies A = 1$ ✓
- The correct treatment is recommended in all scenarios!

Looking Ahead: Future Treatment

True rule: $A = 1$ if $X > \tau = 15$

Estimated Rule: $A = 1$ if $X > \tau^* = 12$

And sometimes, measurement error can even help!



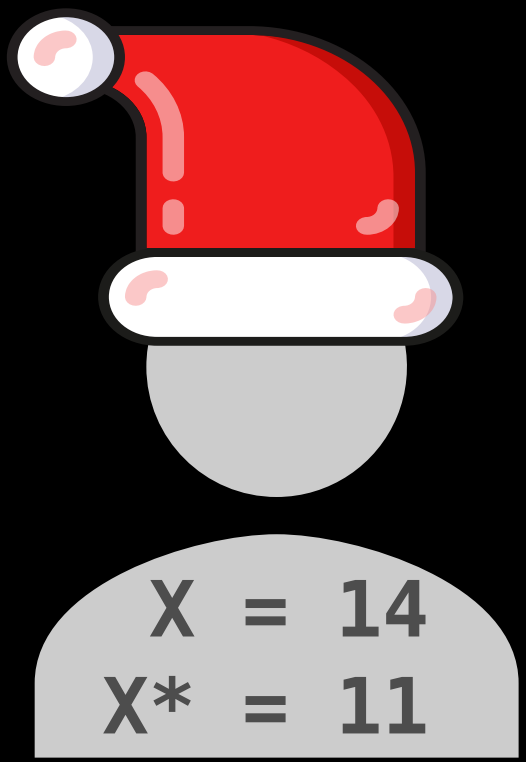
- $\boxed{X = 14,\ X < \tau = 15,\ A^{opt} = 0}$

Looking Ahead: Future Treatment

True rule: $A = 1$ if $X > \tau = 15$

Estimated Rule: $A = 1$ if $X > \tau^* = 12$

And sometimes, measurement error can even help!



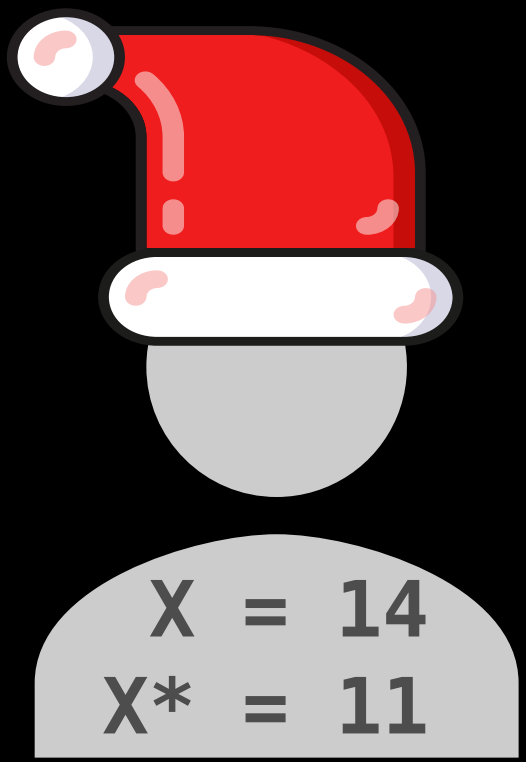
- $\boxed{X = 14,\ X < \tau = 15,\ A^{opt} = 0}$
- $X = 14,\ X > \tau^* = 12 \implies A = 1$ $\times$

Looking Ahead: Future Treatment

True rule: $A = 1$ if $X > \tau = 15$

Estimated Rule: $A = 1$ if $X > \tau^* = 12$

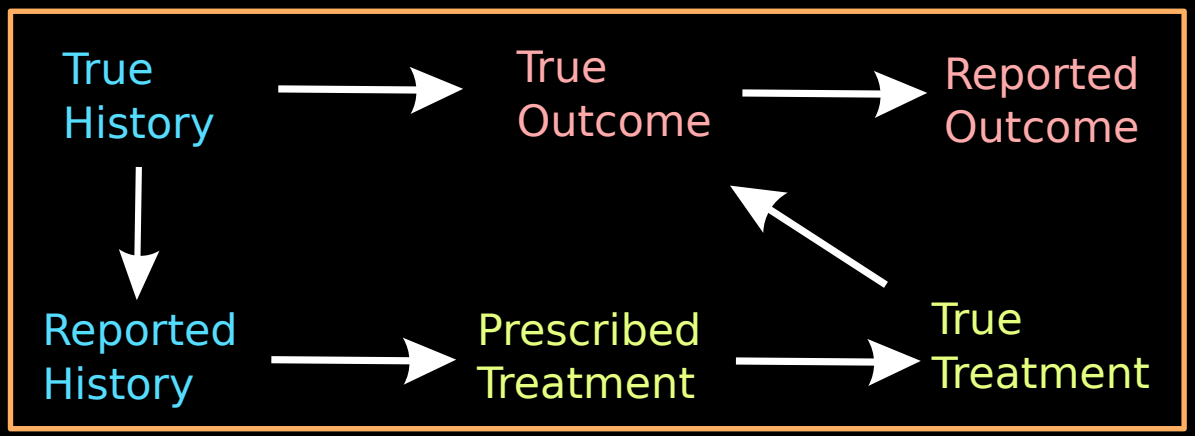
And sometimes, measurement error can even help!



- $\boxed{X = 14,\ X < \tau = 15,\ A^{opt} = 0}$
- $X = 14,\ X > \tau^* = 12 \implies A = 1$ ×
- $X^* = 11,\ X^* < \tau = 15 \implies A = 0$ ✓
- $X^* = 11,\ X^* < \tau^* = 12 \implies A = 0$ ✓

So 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.
- Critical to understand impact of using error-prone measurements for treatment decision making, especially where error may depend on individual characteristics.
- But: many more cases to explore.



- **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. 2024.

michael.wallace@uwaterloo.ca

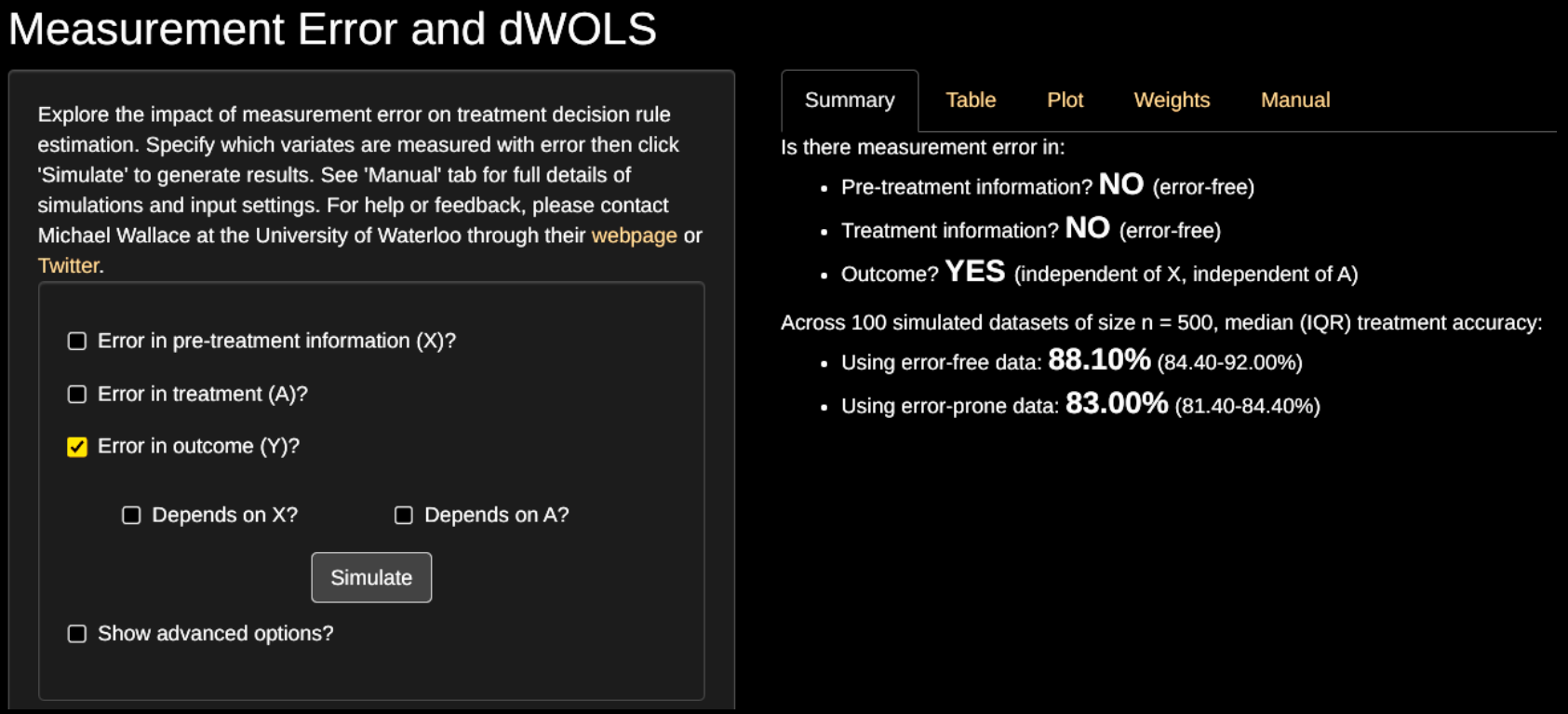
mpwallace.github.io

NSERC CRSNG CIHR IRSC Canadian Institutes of Health Research Instituts de recherche en santé du Canada

Shiny App: Error in Y

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

$$Y^* = Y + U \qquad U \sim N(\mu_{uy}, \sigma^2_{uy})$$

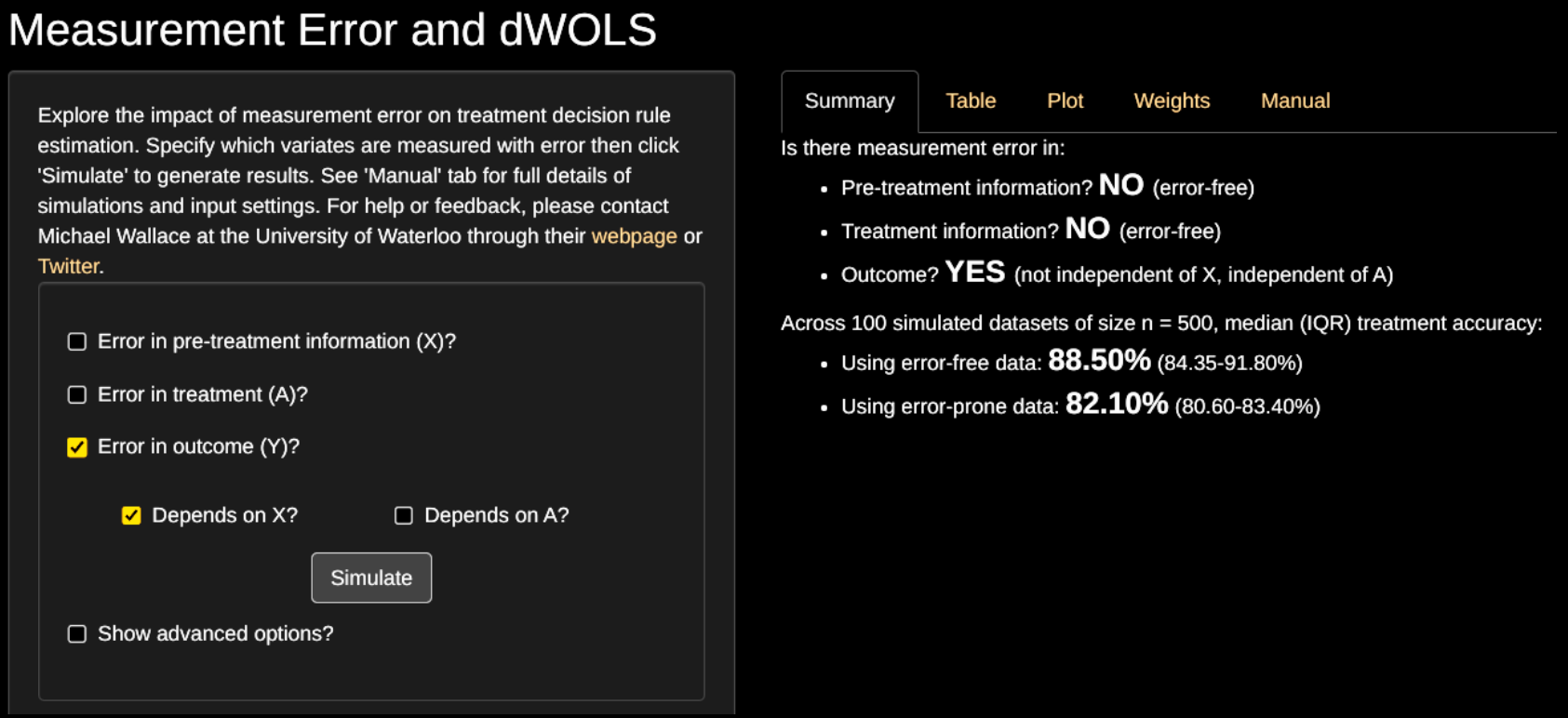


Error independent of X, A

Shiny App: Error in Y

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

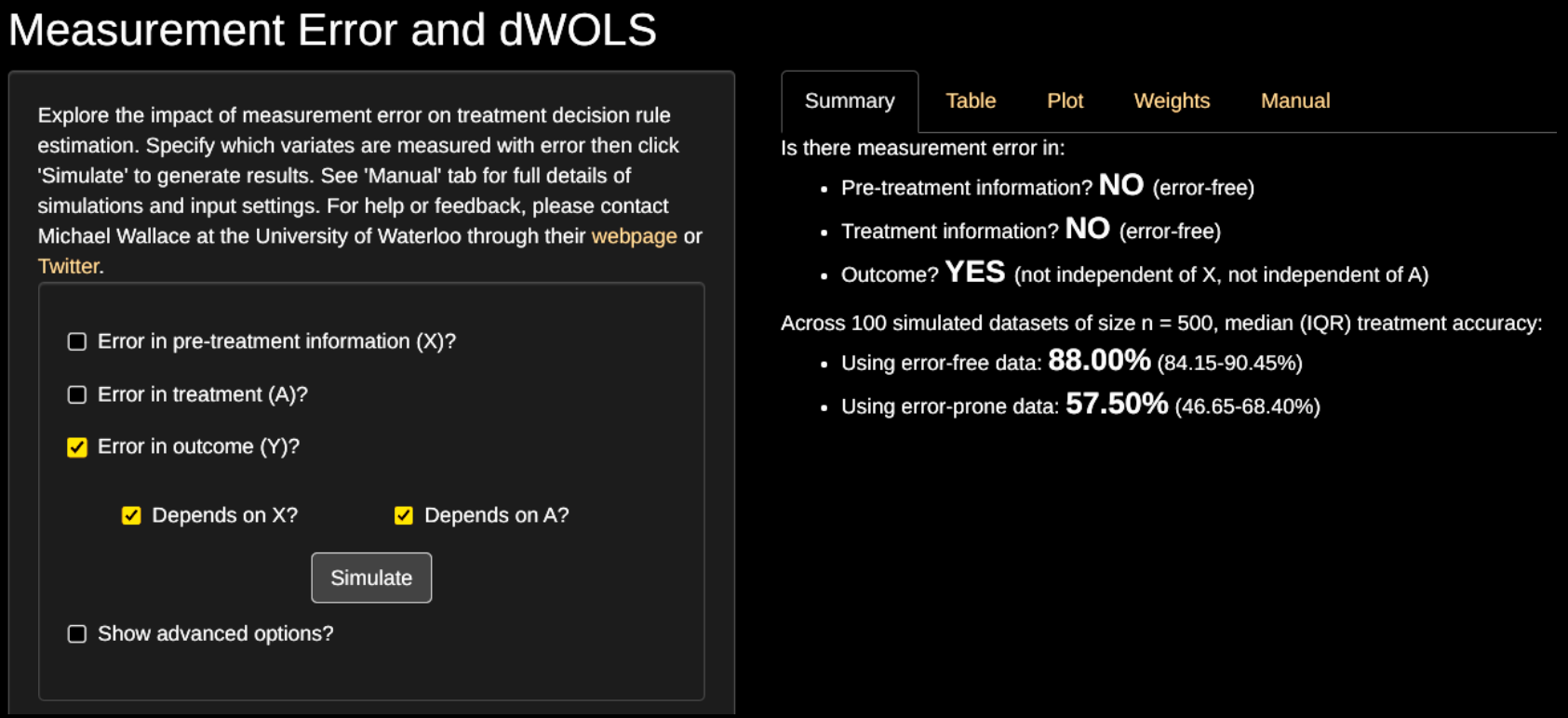
$$Y^* = Y + U \qquad U \sim N(\mu_{uy}, \sigma^2_{uy})$$



Error independent of A

Shiny App: Error in Y

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$
$$Y^* = Y + U \qquad U \sim N(\mu_{uy}, \sigma^2_{uy})$$



Error not independent of A

"Isn't this just a prediction problem?"

Data availability will vary by study and by variable:

Scenario	Analysis	Application
1: "We never observe the truth."	X^*	X^*

"Isn't this just a prediction problem?"

What if error-free data are possible, but expensive?

Scenario	Analysis	Application
1: "We never observe the truth."	X^*	X^*
2: "Past data are error-prone, but future data may not be."	X^*	X
3: "Past data are not error-prone, but future data may be."	X	X^*

"Isn't this just a prediction problem?"

Only Scenario 4 is well-studied.

Scenario	Analysis	Application
1: "We never observe the truth."	X^*	X^*
2: "Past data are error-prone, but future data may not be."	X^*	X
3: "Past data are not error-prone, but future data may be."	X	X^*
4: "We always observe the truth."	X	X