Imprecise Medicine? Measurement Error and Personalized Treatments

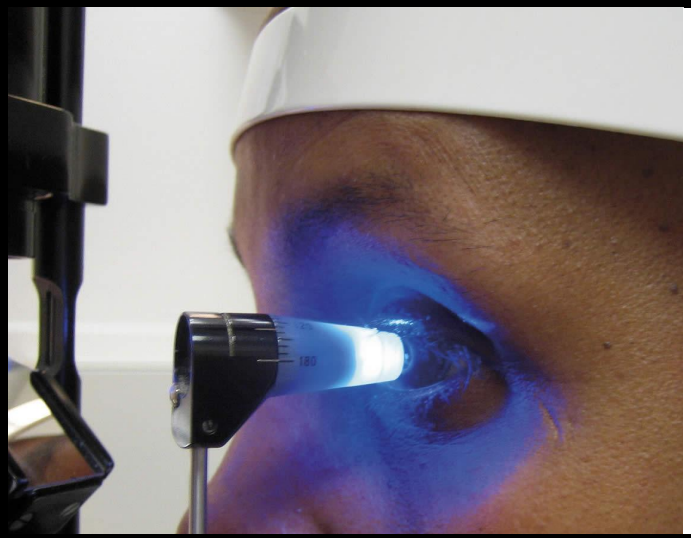
Michael Wallace, University of Waterloo

Slides available at: `mpwallace.github.io`

Glaucoma: One Disease, Many Treatments

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

IOP can be measured in various ways.



Glaucoma: One Disease, Many Treatments

Elevated IOP can cause vision loss, which can be measured through visual field tests.

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

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

Glaucoma: One Disease, Many Treatments

Treatment decisions are made based on various factors:

- Current and past IOP.
- Current and past treatments.
- Concerns over side effects.
- Broader risk factors.
- Other characteristics (such as age).

Glaucoma: One Disease, Many Treatments

Treatment decisions are made based on various factors:

- Current and past IOP.
- Current and past treatments.
- Concerns over side effects.
- Broader risk factors.
- Other characteristics (such as age).

Precision Medicine: tailoring treatment decisions to patient-level characteristics.

Dynamic treatment regimes

- Dynamic treatment regimes (DTRs) 'formalize' personalized treatment:

Dynamic treatment regimes

- Dynamic treatment regimes (DTRs) 'formalize' personalized treatment:

- "*Patient is currently taking Azarga eye drops. If current IOP is 15 or higher, add Alphagan eye drops, otherwise continue with only Azarga.*"

Dynamic treatment regimes

- Dynamic treatment regimes (DTRs) 'formalize' personalized treatment:

  - "*Patient is currently taking Azarga eye drops. If current IOP is 15 or higher, add Alphagan eye drops, otherwise continue with only Azarga.*"
- How do we choose the best DTR?

  Should our IOP cut-off be 13, 15, 20?
- What makes this difficult?

The Data

We typically work with data from observational studies.

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

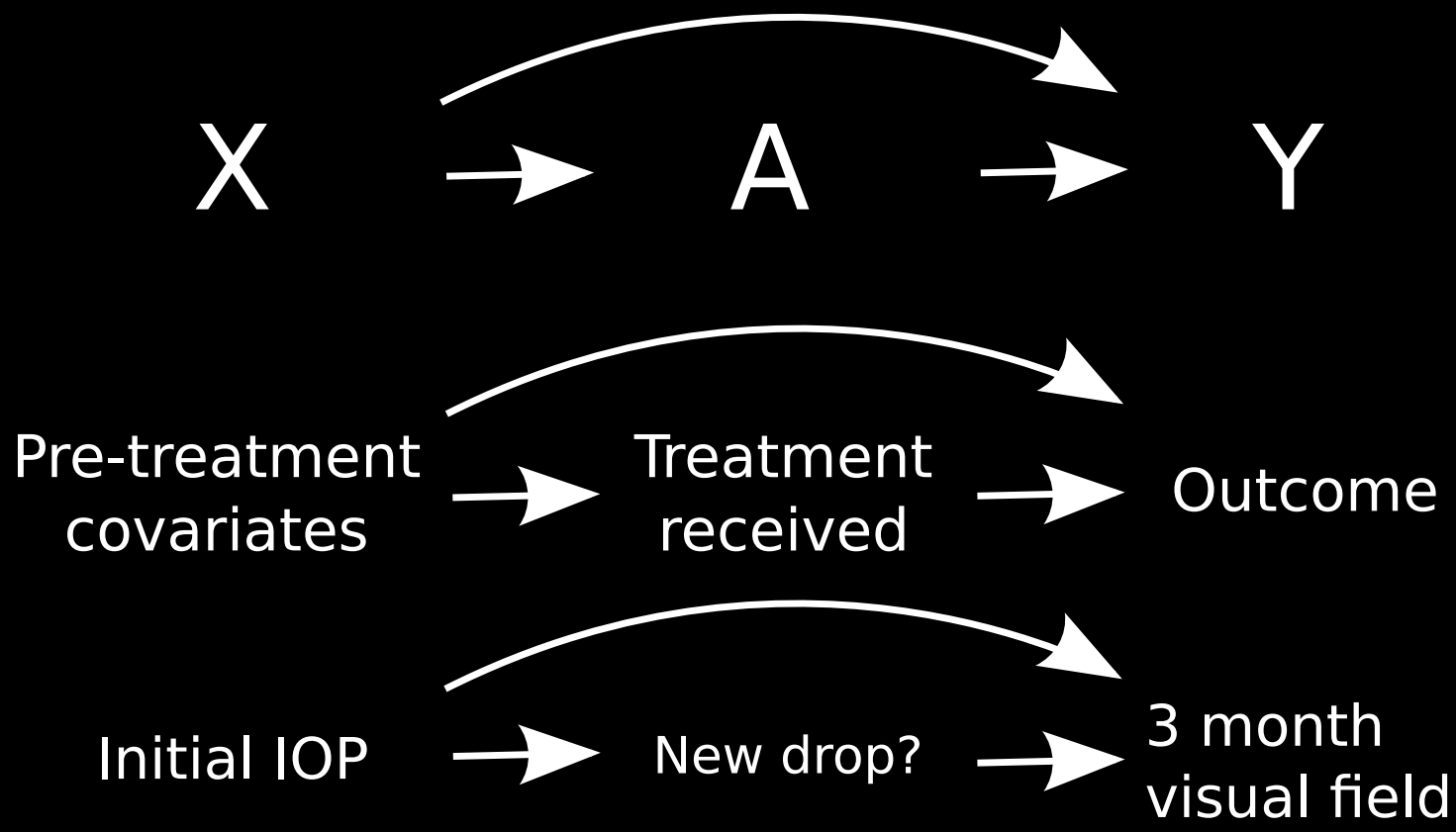
VFP = Visual Field Percentage

We typically work with data from observational studies.

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

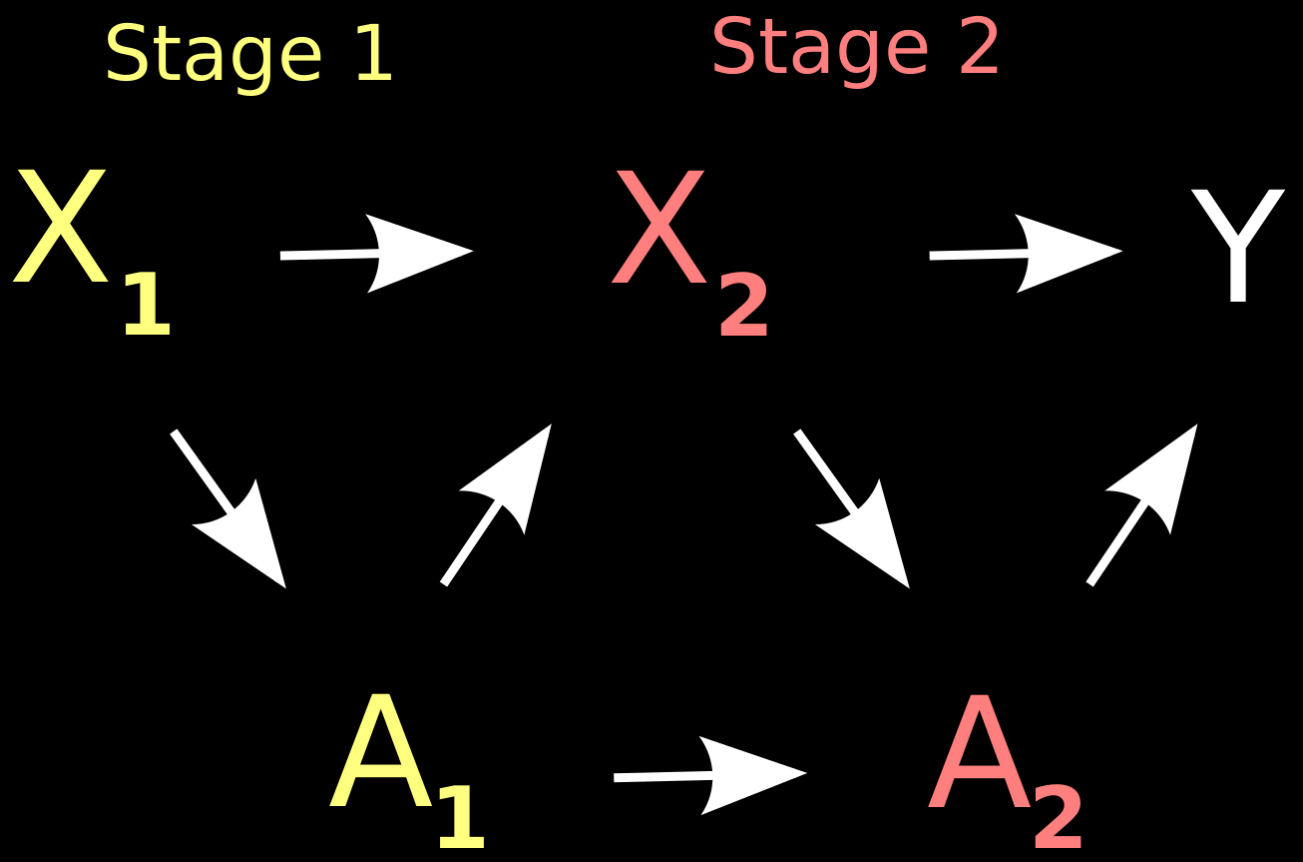
VFP = Visual Field Percentage

Notation



DTR: treatment A^{opt} that optimizes $E[Y|X, A^{opt}]$

Identifying the best treatment regime: multi-stage



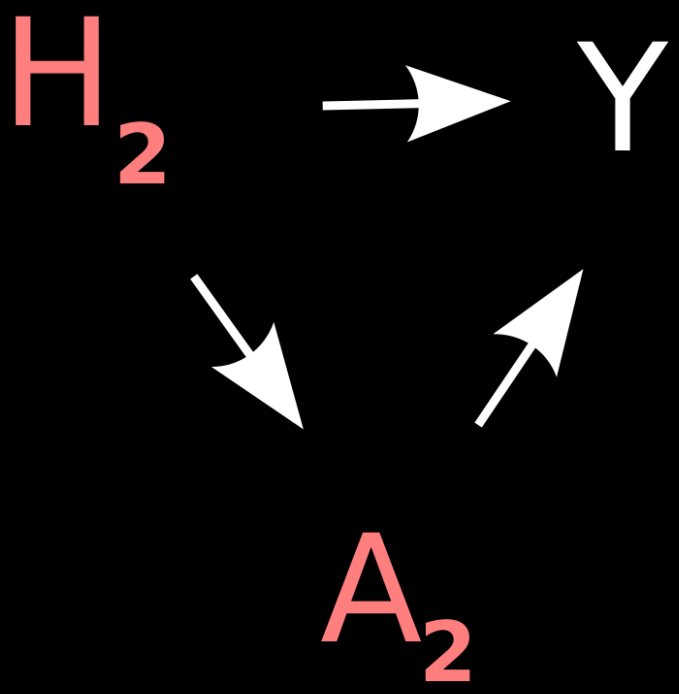
More often, we work in stages.

DTR: treatment sequence A_1^{opt}, A_2^{opt}

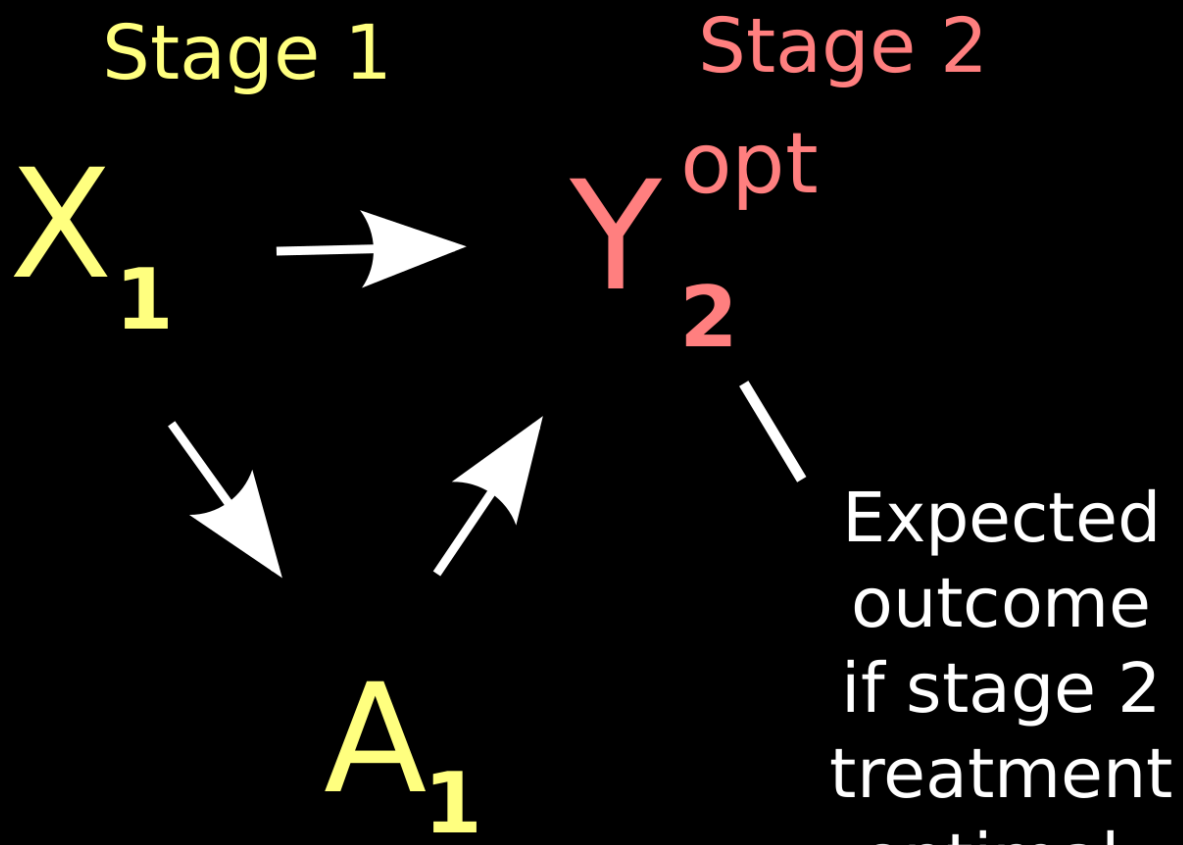
Identifying the best treatment regime: multi-stage

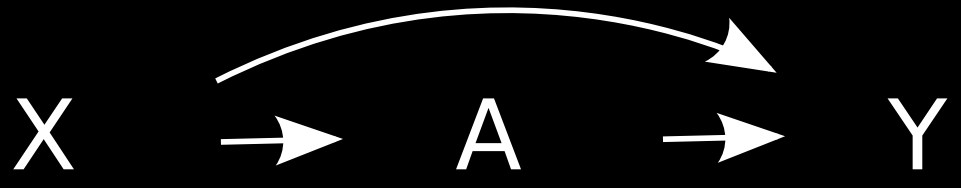
$$H_2 = (X_1, A_1, X_2)$$

Stage 2



Identifying the best treatment regime: multi-stage





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

Lots of methods available:

Q-learning

MSMs

G-estimation

IPTW

dWOLS

OWL

A-learning

etc...

Lots of methods available:

Q-learning

MSMs

G-estimation

IPTW

dWOLS

OWL

A-learning

etc...

Identifying the best treatment regime

- If only one treatment decision:

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

Identifying the best treatment regime

- If only one treatment decision:

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

- We might propose the following model

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

"$A^{opt} = 1$ if $\psi_0 + \psi_1 \text{IOP} > 0$"

Identifying the best treatment regime

- If only one treatment decision:

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

- We might propose the following model

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

"$A^{opt} = 1$ if $\psi_0 + \psi_1 \text{IOP} > 0$"

- More generally, split outcome into two components:

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

- Simplifies focus: find A^{opt} that maximizes $\gamma(X,A;\psi)$.

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:

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

Dynamic WOLS (dWOLS)

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

- Three models to specify:
  1. Blip model: $\gamma(X, A; \psi)$.
  2. Treatment-free model: $G(X; \beta)$.
  3. Treatment model: $P(A = 1|X; \alpha)$.

Dynamic WOLS (dWOLS)

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

- Three models to specify:
  1. Blip model: $\gamma(X, A; \psi)$.
  2. Treatment-free model: $G(X; \beta)$.
  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:

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

- A weighted regression 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:

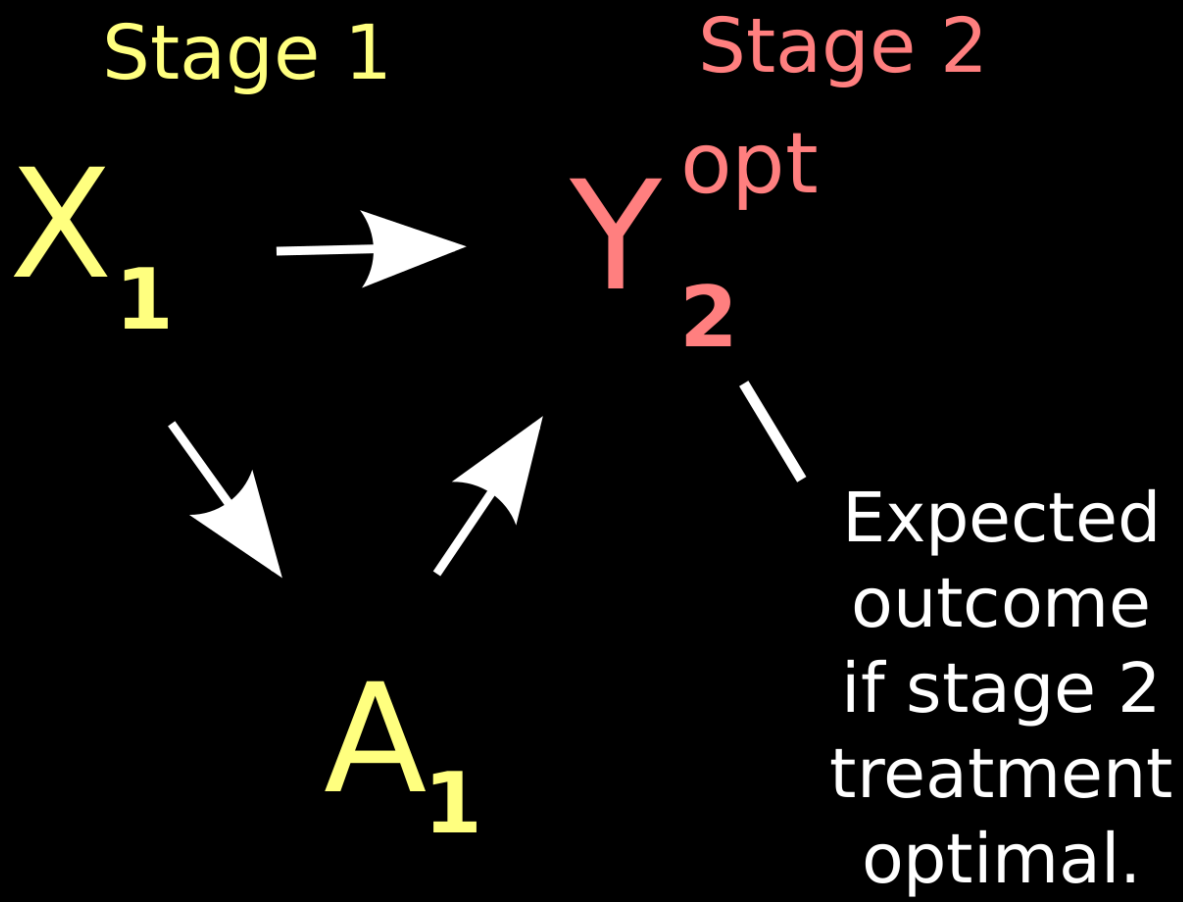
$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 \text{IOP} + \beta_2 \text{IOP}^2 + A(\psi_0 + \psi_1 \text{IOP})$$

- But we propose:

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

- A weighted regression with weights $w = |A - P(A = 1|X; \hat{\alpha})|$ will still yield consistent estimators of ψ_0, ψ_1.
- The estimators are "doubly robust": consistent if at least one of the treatment-free or treatment components is correctly specified.
- The blip must always be correct.

Identifying the best treatment regime: multi-stage



More formally, write $\widetilde{Y}_j$ for the stage j 'pseudo-outcome'.

$\widetilde{Y}_j$ is the expected outcome assuming optimal treatment from stage $j+1$ onwards.

Pseudo-outcome = observed outcome + estimated 'loss' of receiving non-optimal treatments

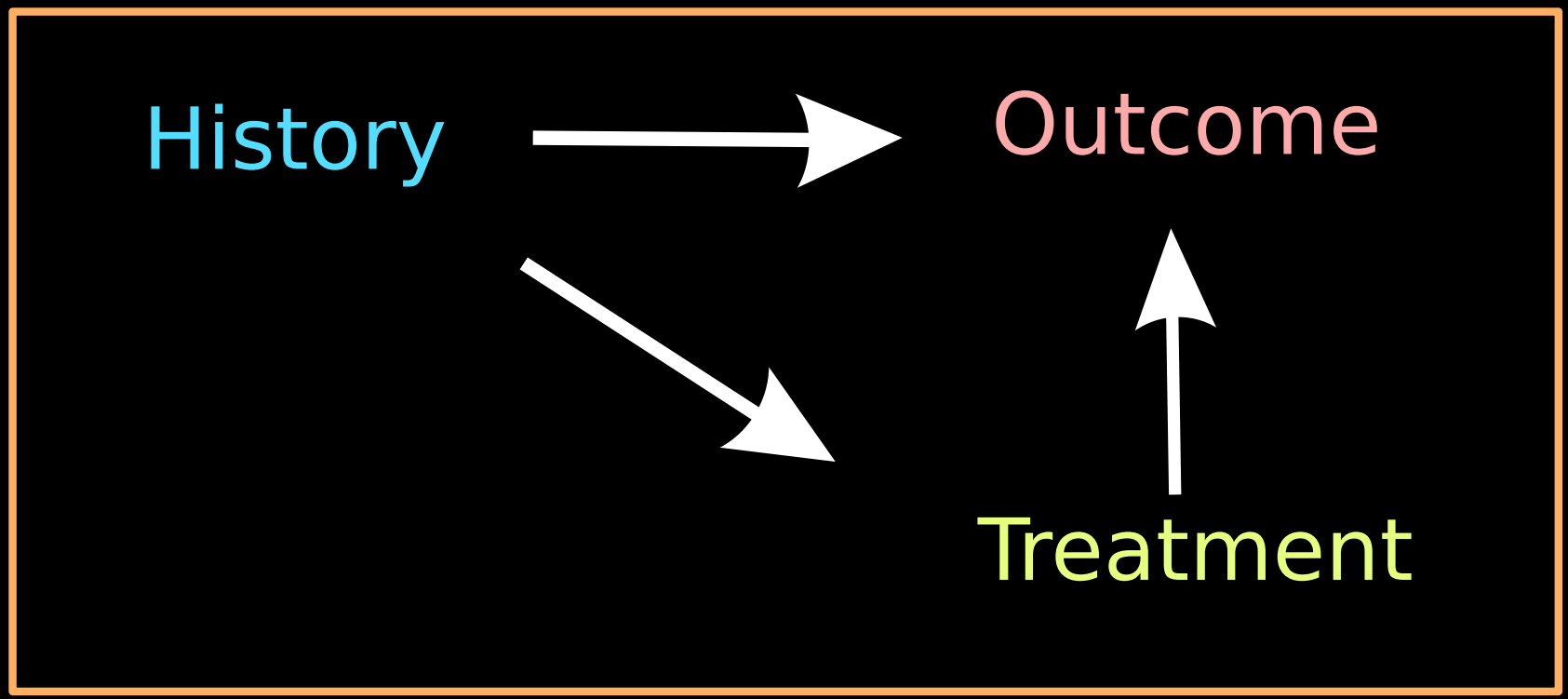
More formally, write $\widetilde{Y}_j$ for the stage j 'pseudo-outcome'.

$\widetilde{Y}_j$ is the expected outcome assuming optimal treatment from stage $j+1$ onwards.

Pseudo-outcome = observed outcome + estimated 'loss' of receiving non-optimal treatments

$$\widetilde{Y}_j = Y + \sum_{k=j+1}^{J} [\gamma_k(X_k, A_k^{opt}; \hat{\psi}_k) - \gamma_k(X_k, A_k; \hat{\psi}_k)]$$

We plug $\widetilde{Y}_j$ into our dWOLS procedure and proceed similarly.

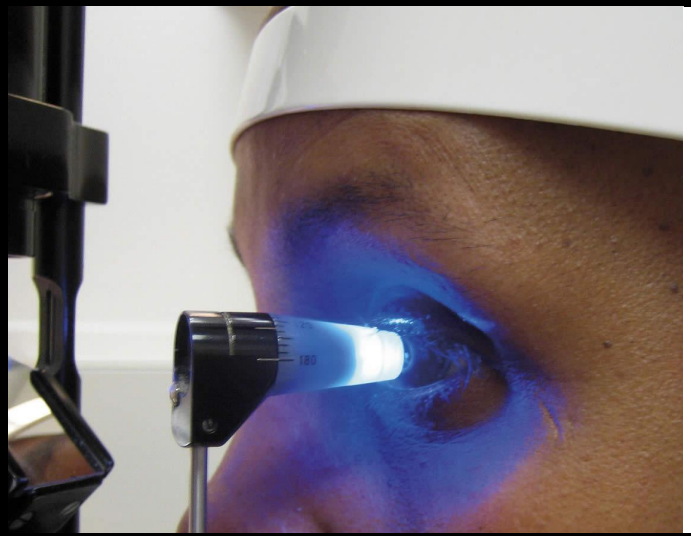

History
Outcome
Treatment

History

Target measurement: 'average' IOP.

Observed measurement: 1-3 in-clinic readings within < 5 minutes.

Some patients have access to more regular at-home tonometry.

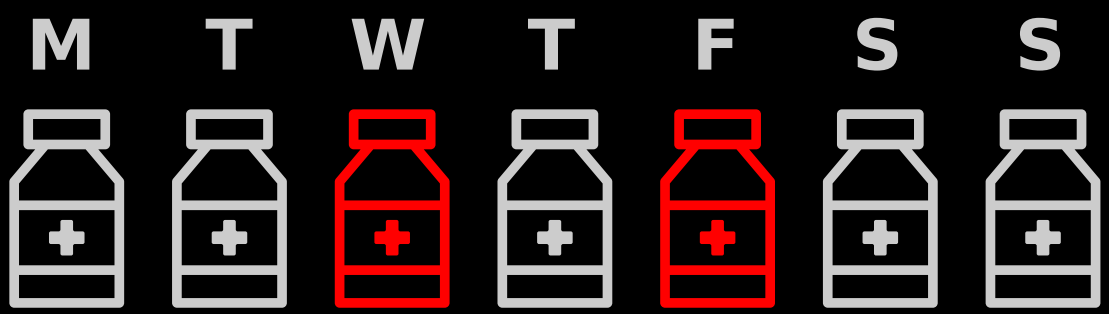


Treatment

Target measurement: adherence with prescribed dosing regimen.

Observed measurement: prescribed treatment or patient-reported adherence.

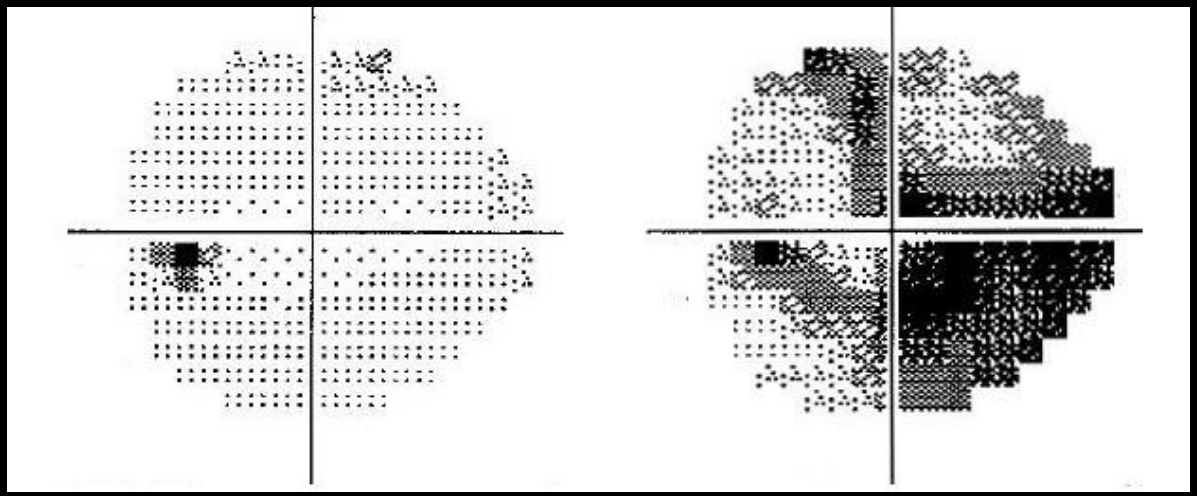
Full adherence with therapies reported in 10% of patients.

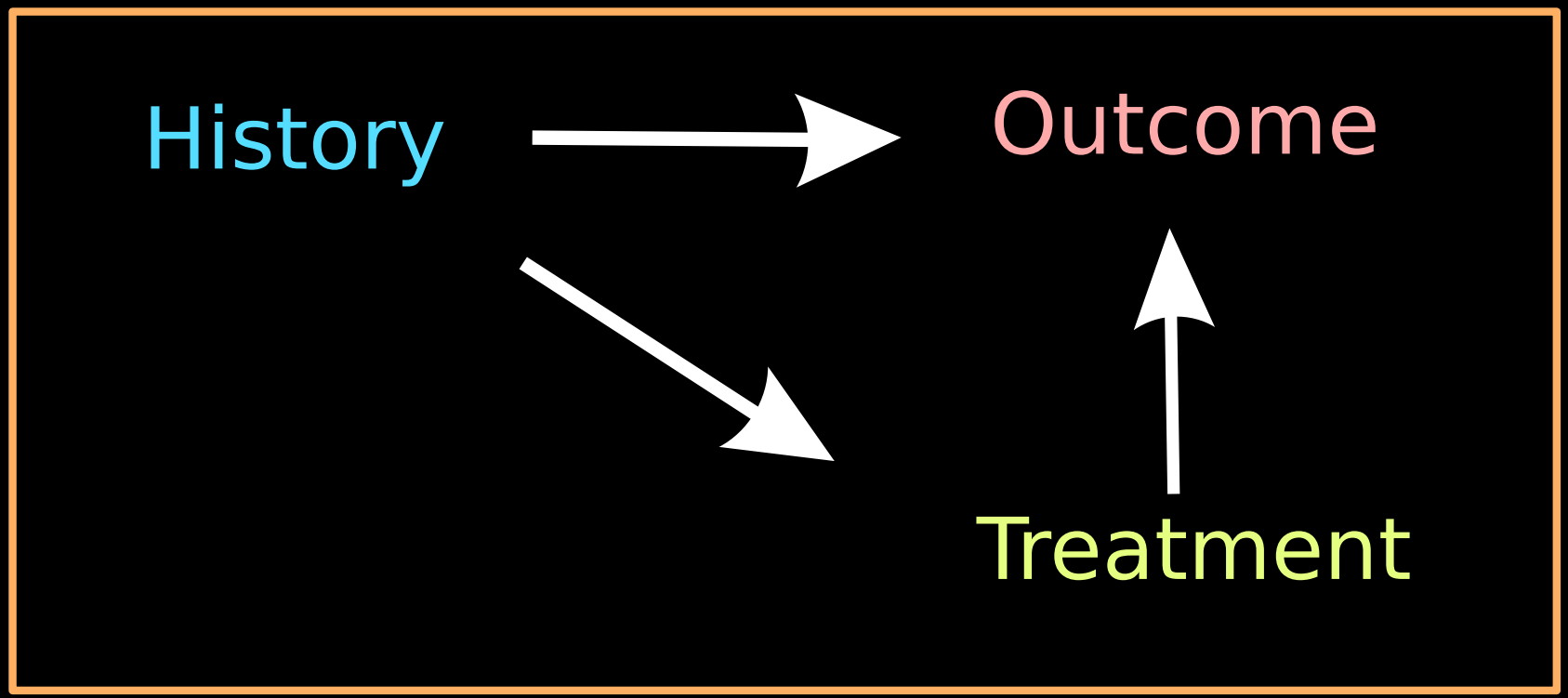


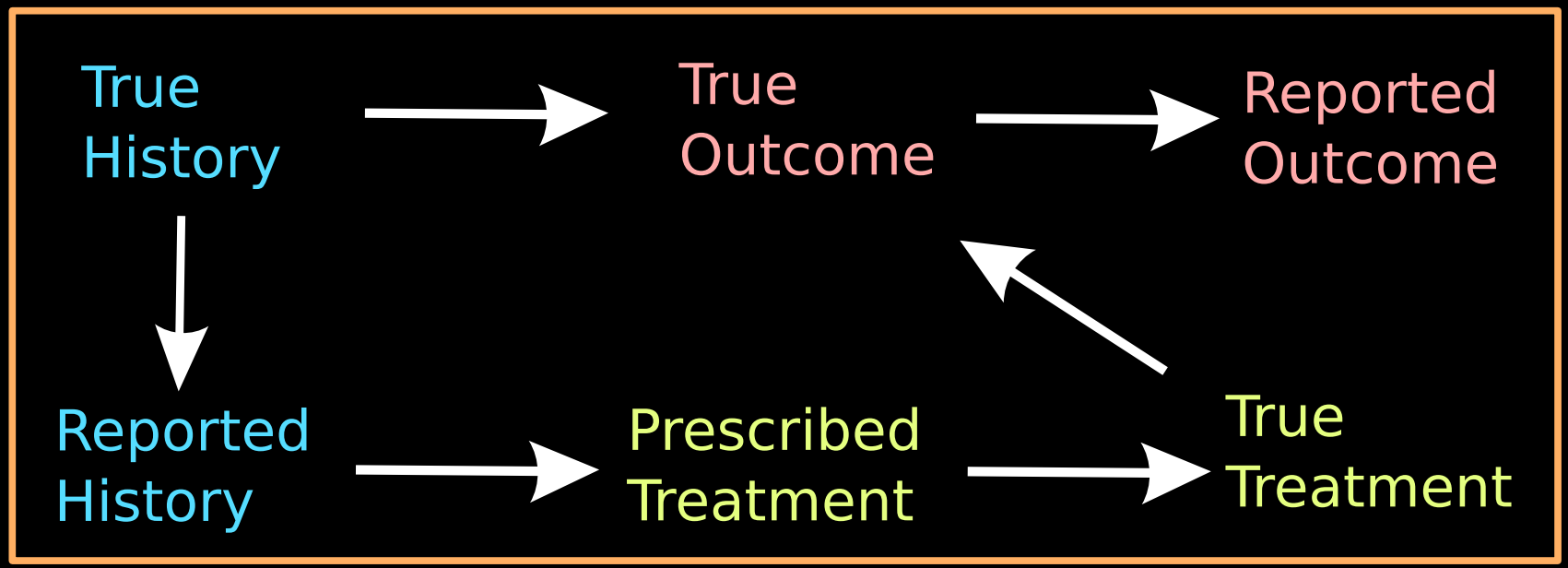
Outcome

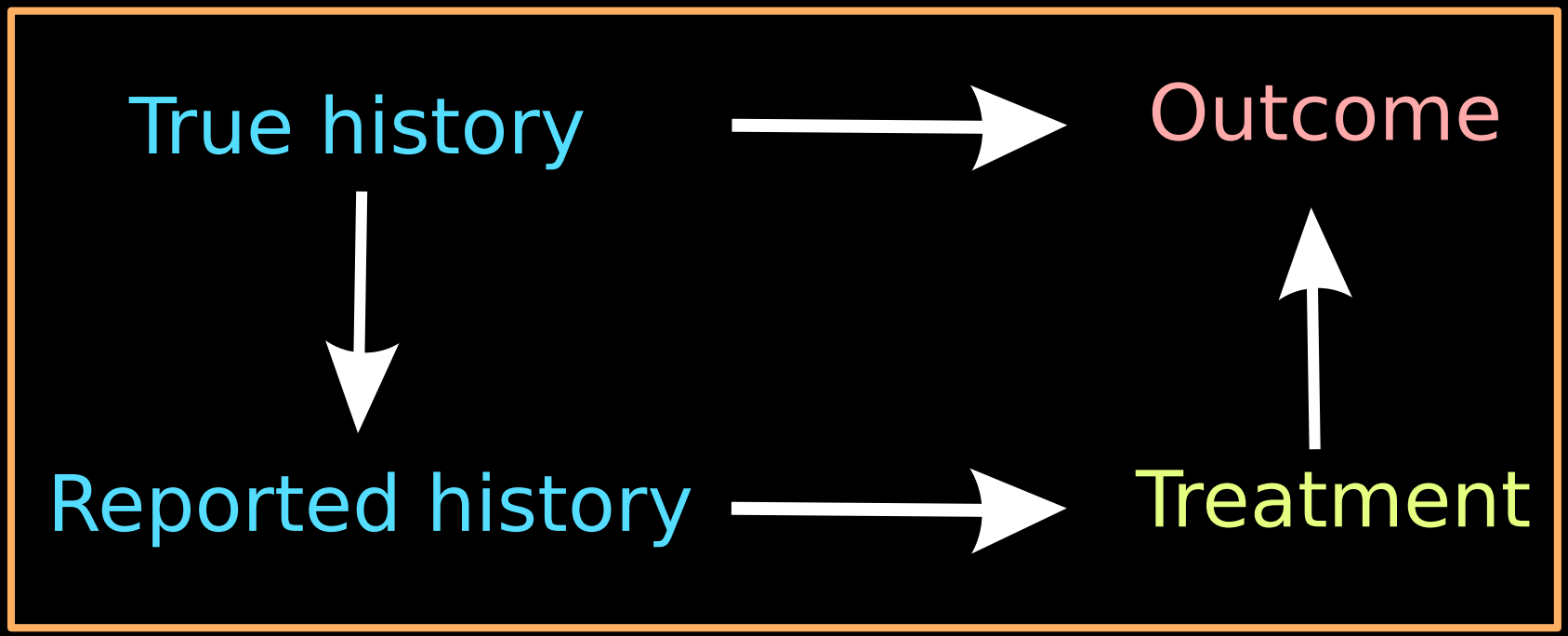
Target measurement: % of remaining vision.

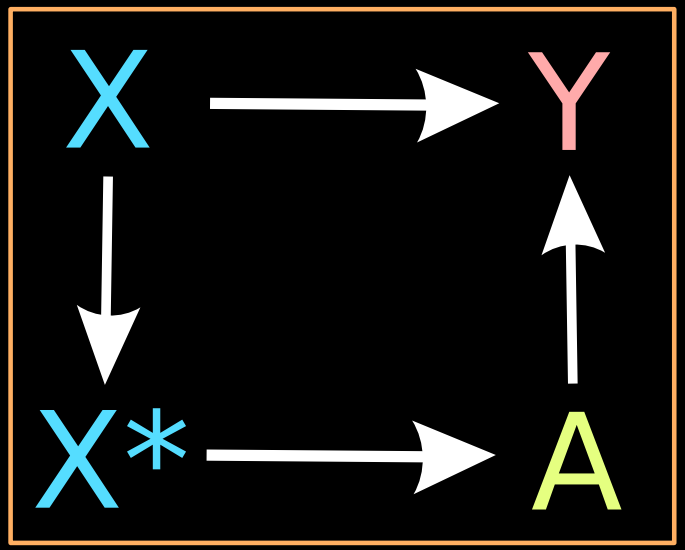
Observed measurement: visual field test.




History
Outcome
Treatment


True
History
True
Outcome
Reported
Outcome
Reported
History
Prescribed
Treatment
True
Treatment


True history
Outcome
Reported history
Treatment



Estimation: suppose the true outcome model is:

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

If we only observe X^*. What happens? What can we do about it?

Assume: classical additive measurement error:

$$\text{Observed} = \text{True} + \text{Error}$$

$$X^* = X + U$$

$U \sim N(0, \sigma_u^2)$; $Y \perp X^* | X$

Assume: classical additive measurement error:

$$\text{Observed} = \text{True} + \text{Error}$$
$$X^* = X + U$$

$U \sim N(0, \sigma_u^2)$; $Y \perp X^* | X$

Assume: replicate measurements available on at least some patients.

Patient	First IOP measurement	Second IOP measurement
1	16	15
2	20	16
3	21	17
4	16	16
5	15	18
...	...	...

Simple correction method: Regression Calibration.

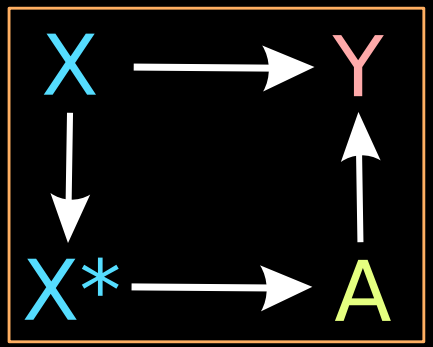
Principle:

1. Use additional data to estimate $E[X|X^*, A] = X_{rc}$.
2. Replace X with X_{rc} and carry out a standard analysis.
3. Adjust the resulting standard errors to account for the estimation in step 1.

Identifying the best treatment regime

- Suppose the true outcome model is:

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



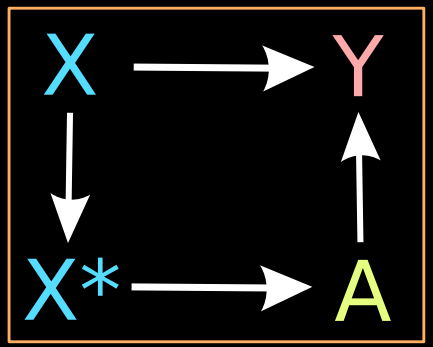
Identifying the best treatment regime

- Suppose the true outcome model is:

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

- If we have RC estimates X_{rc} then we could fit

$$E[Y|\cdot] = \beta_0 + \beta_1 X_{rc} + \beta_2 X_{rc}^2 + A(\psi_0 + \psi_1 X_{rc})$$



Identifying the best treatment regime

- Suppose the true outcome model is:

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

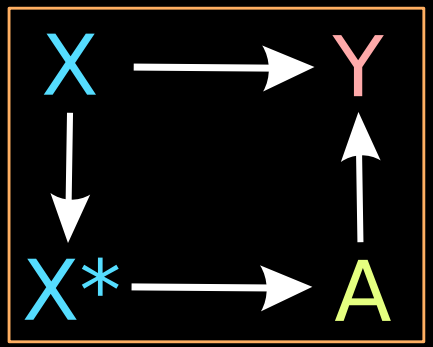
- If we have RC estimates X_{rc} then we could fit

$$E[Y|\cdot] = \beta_0 + \beta_1 X_{rc} + \beta_2 X_{rc}^2 + A(\psi_0 + \psi_1 X_{rc})$$

- But we might mis-specify the model as

$$E[Y|\cdot] = \beta_0 + \beta_1 X_{rc} + A(\psi_0 + \psi_1 X_{rc})$$

where A depends on X^*.



Identifying the best treatment regime

- Suppose the true outcome model is:

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

- If we have RC estimates X_{rc} then we could fit

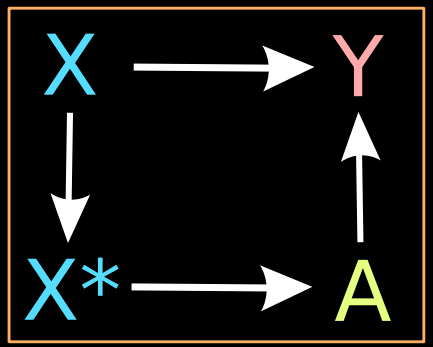
$$E[Y|\cdot] = \beta_0 + \beta_1 X_{rc} + \beta_2 X_{rc}^2 + A(\psi_0 + \psi_1 X_{rc})$$

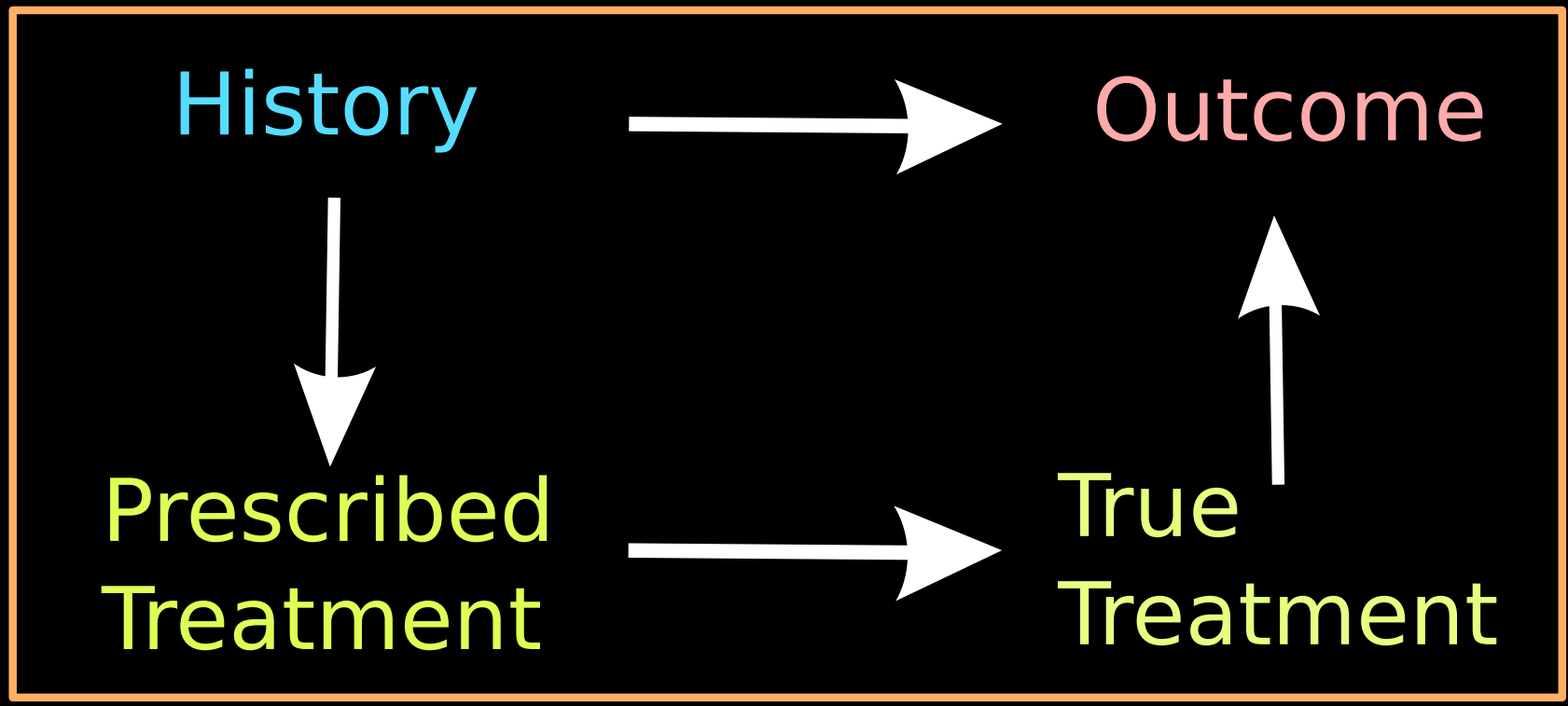
- But we might mis-specify the model as

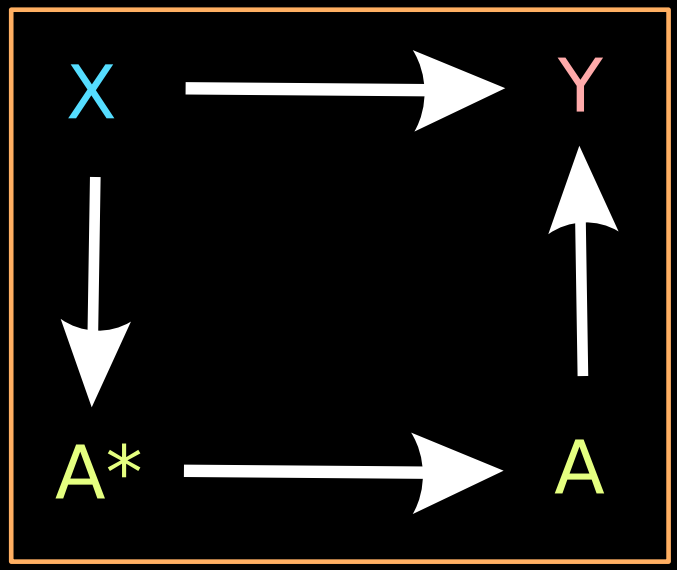
$$E[Y|\cdot] = \beta_0 + \beta_1 X_{rc} + A(\psi_0 + \psi_1 X_{rc})$$

  where A depends on X^*.

- Solution: dWOLS using $P(A = 1|X_{rc})$ estimates.

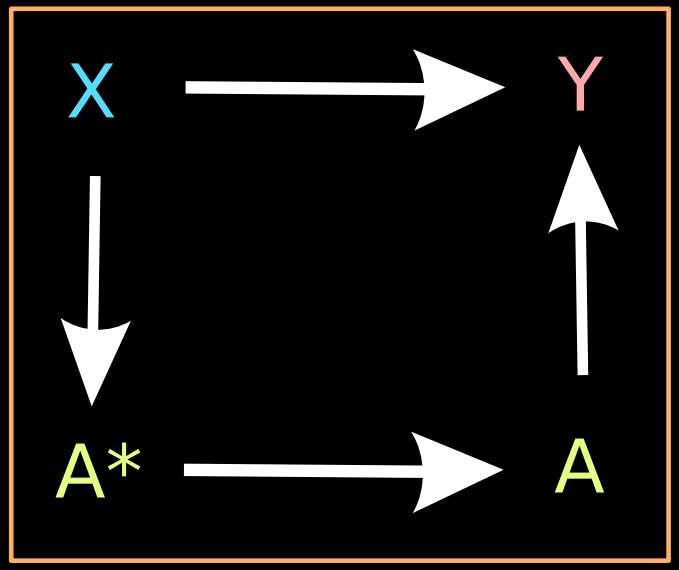



History
Outcome
Prescribed Treatment
True Treatment



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

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



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 only observe A^*.

Key question: do the misclassification probabilities depend on X?

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

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

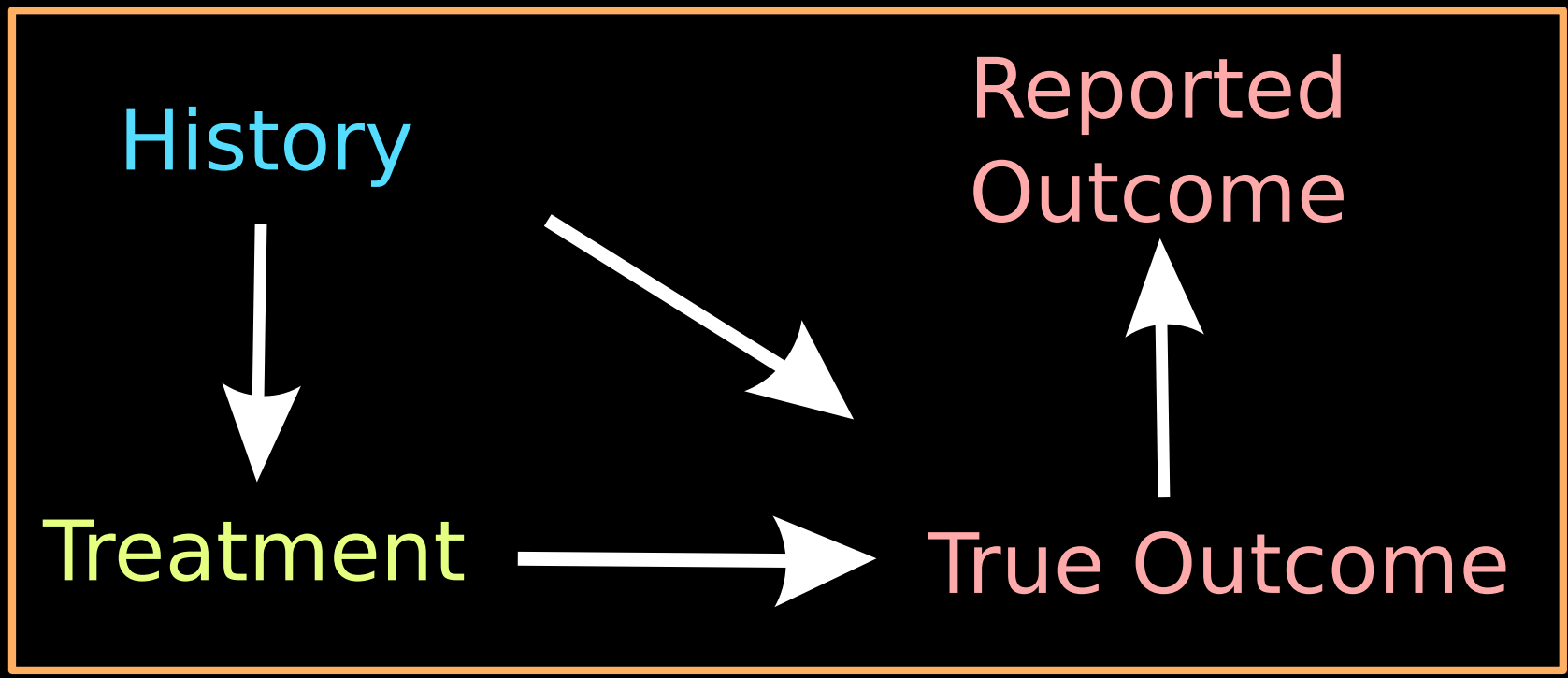
$$A^{opt} = 1 \text{ if } \psi_0 + \psi_1 X > 0$$

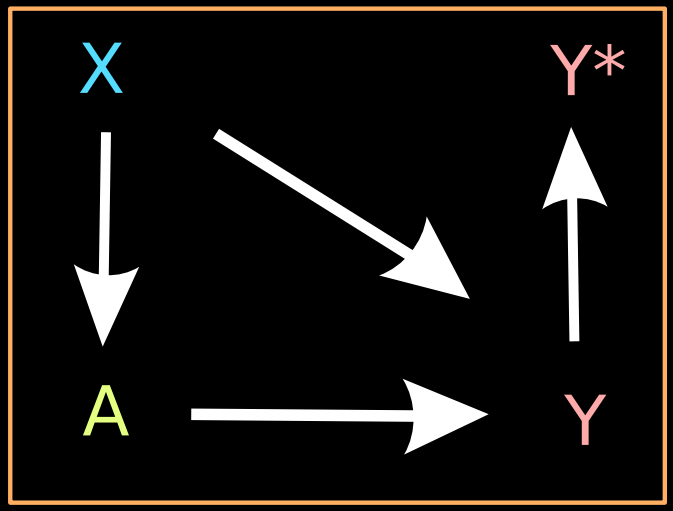
which is unaffected if ψ_0, ψ_1 are biased by the same factor.

If misclassification depends on X, then corrective action is required.

Upcoming work modifies G-estimation to account for treatment misclassification.

Further questions exist related to intention to treat analyses, and implications of (non-) adherence for identifying optimal treatment rules.


History
Reported
Outcome
Treatment
True Outcome



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 only observe

$$Y^* = Y + U \qquad U \sim N(\mu_u, \sigma_u^2)$$

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

- Unbiased error: parameter estimates also unbiased.

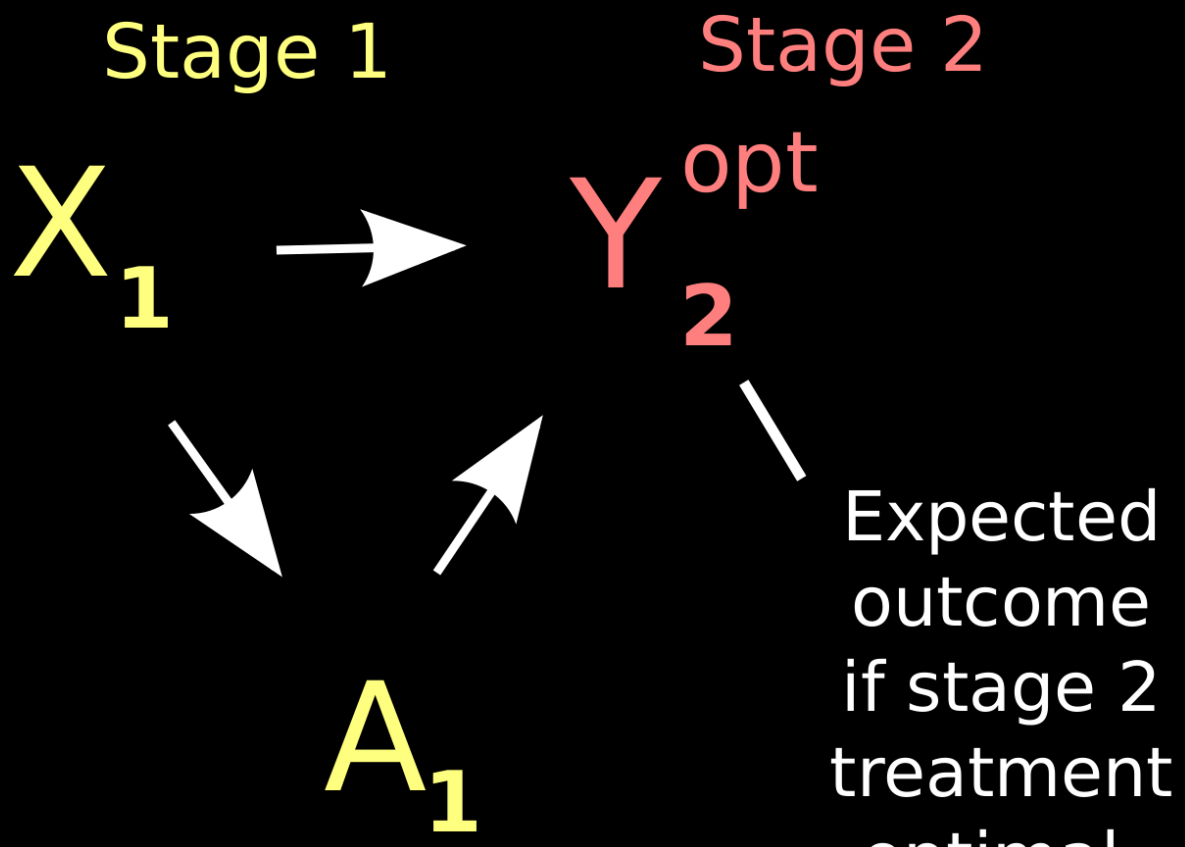
$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(\psi_0 + \psi_1 X)$$
$$Y^* = Y + U \qquad U \sim N(\mu_u, \sigma_u^2)$$

- Unbiased error: parameter estimates also unbiased.
- Biased error, independent of A: ψ estimators still consistent.

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

- Unbiased error: parameter estimates also unbiased.
- Biased error, independent of A: ψ estimators still consistent.
- Biased error, not independent of A: ψ estimators no longer reliable.

Measurement Error and Pseudo-outcomes



Recall the multi-stage case requires the computation of pseudo-outcomes:

$$\tilde{Y}_j = Y + \sum_{k=j+1}^{J} [\gamma_k(X_k, A_k^{opt}; \hat{\psi}_k) - \gamma_k(X_k, A_k; \hat{\psi}_k)].$$

Errors in X, A, or Y create additional problems.

Suppose we conclude that our treatment rule should be:

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

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?

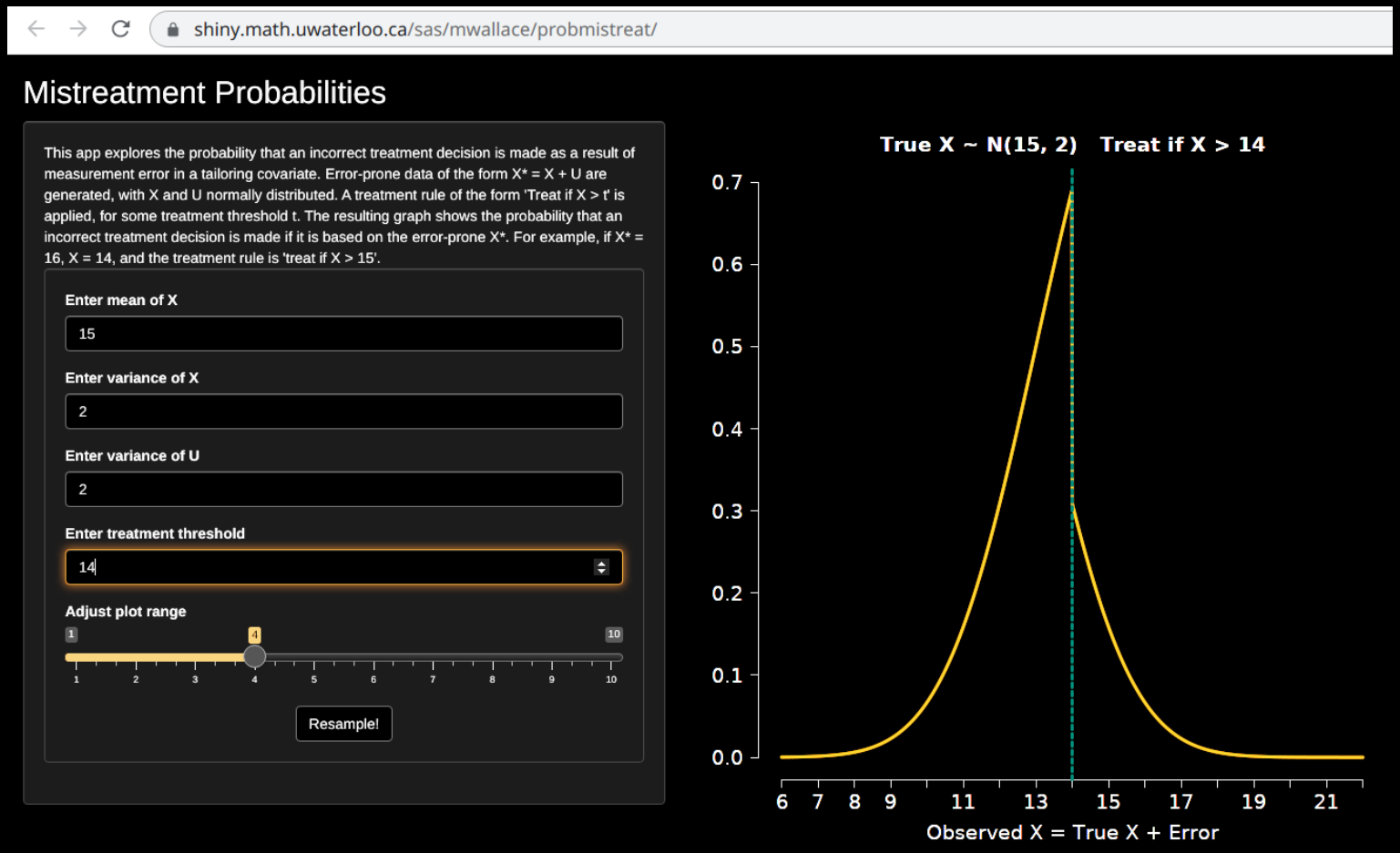
Suppose we conclude that our treatment rule should be:

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

We can explore such probabilities through computation/simulation:



https://shiny.math.uwaterloo.ca/sas/mwallace/probmistreat/

So where are we now?

- DTRs an important tool in precision medicine.
- Measurement error an important consideration in patient history, treatment, outcome, and future decision making.
- There are some special cases where errors have limited impact, or may be corrected for with standard theory.
- But: many more cases to explore.

Acknowledgments

Dylan Spicker
dylan.spicker@uwaterloo.ca

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

- **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)** https://doi.org/10.1002/sim.8690

- **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. Expected publication 2022.

✉ michael.wallace@uwaterloo.ca  @statacake

mpwallace.github.io