Measurement error and precision medicine

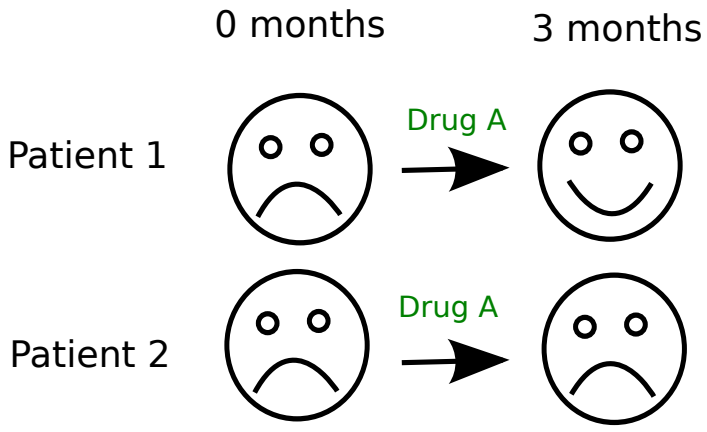
Michael Wallace, Dylan Spicker
Department of Statistics and Actuarial Science
University of Waterloo
michael.wallace@uwaterloo.ca

June 20, 2019

Last time I was in Manchester...

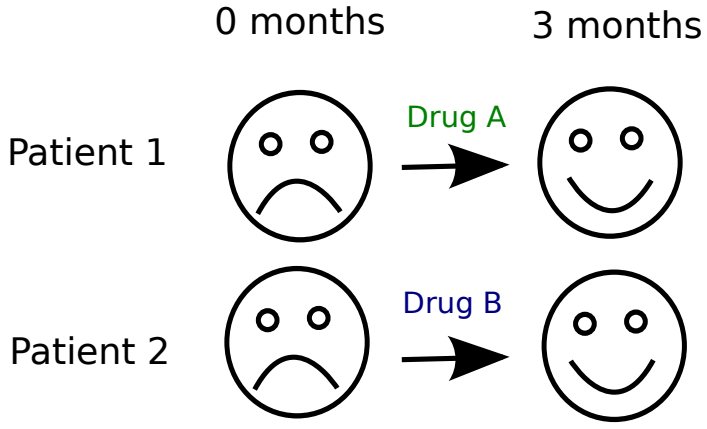
Treating the patient, not the diagnosis

Heterogeneity between patients:



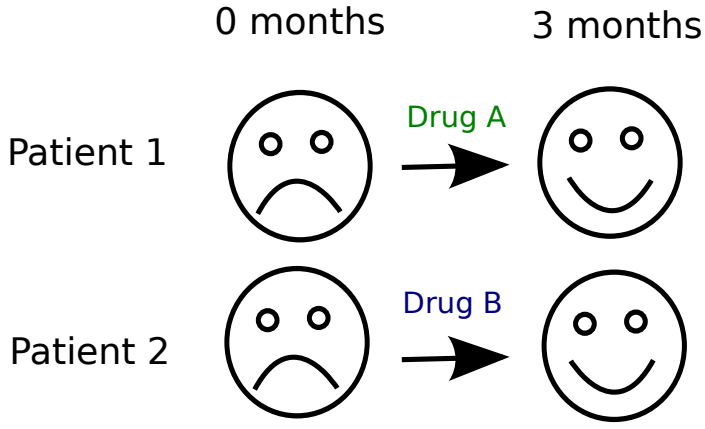
Treating the patient, not the diagnosis

Heterogeneity between patients:

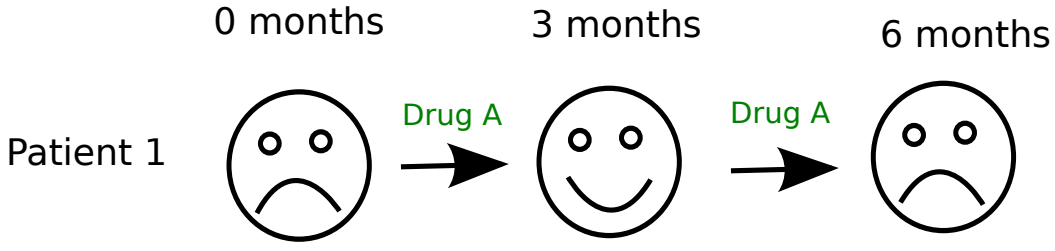


Treating the patient, not the diagnosis

Heterogeneity between patients:

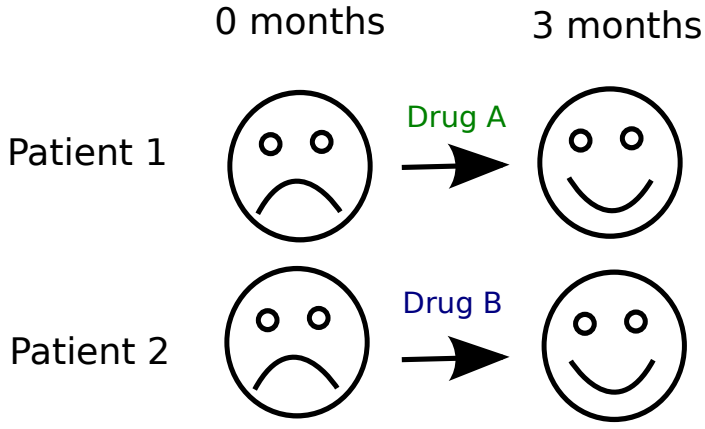


Heterogeneity within patients:

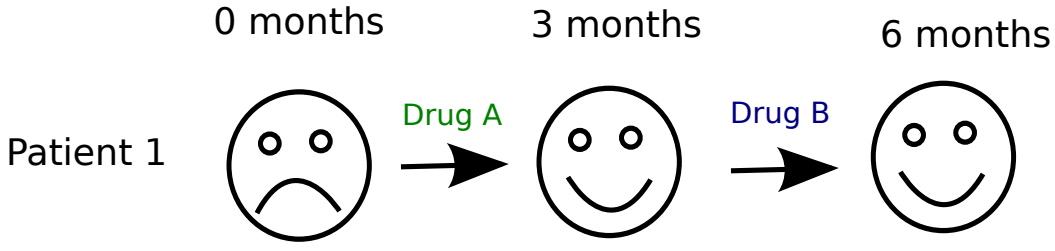


Treating the patient, not the diagnosis

Heterogeneity between patients:



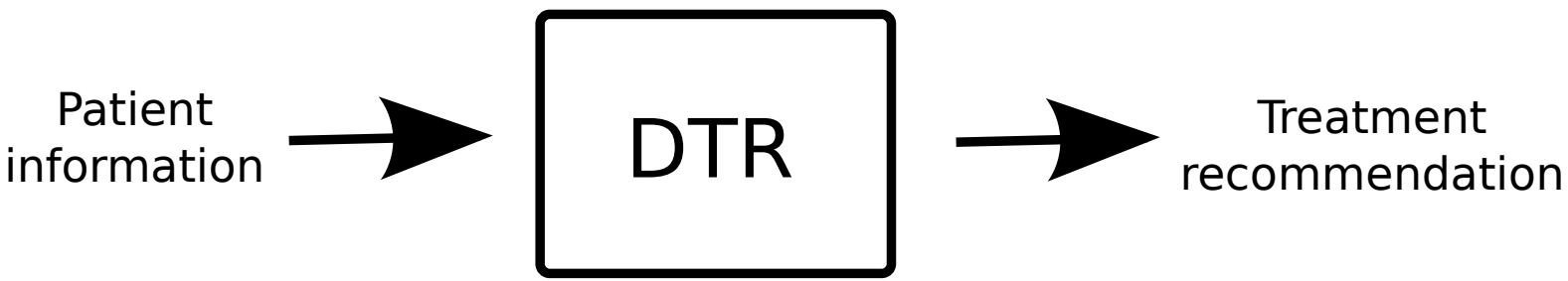
Heterogeneity within patients:



Dynamic treatment regimes

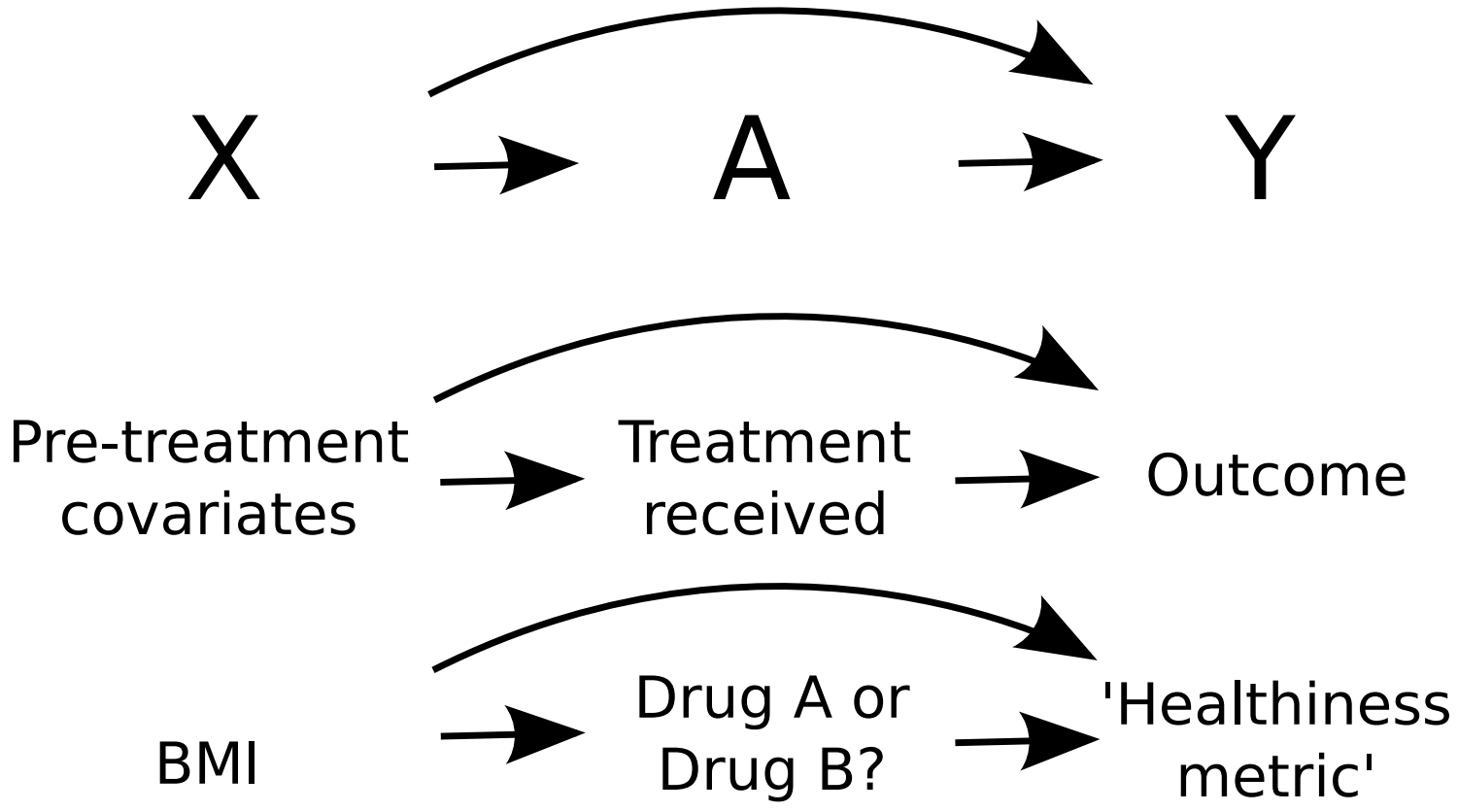
- Dynamic treatment regimes (DTRs) 'formalize' the process of precision medicine:

  "*If patient BMI over 30 prescribe therapy A, otherwise provide therapy B.*"



- DTRs can lead to improved results over standard 'one size fits all' approaches.

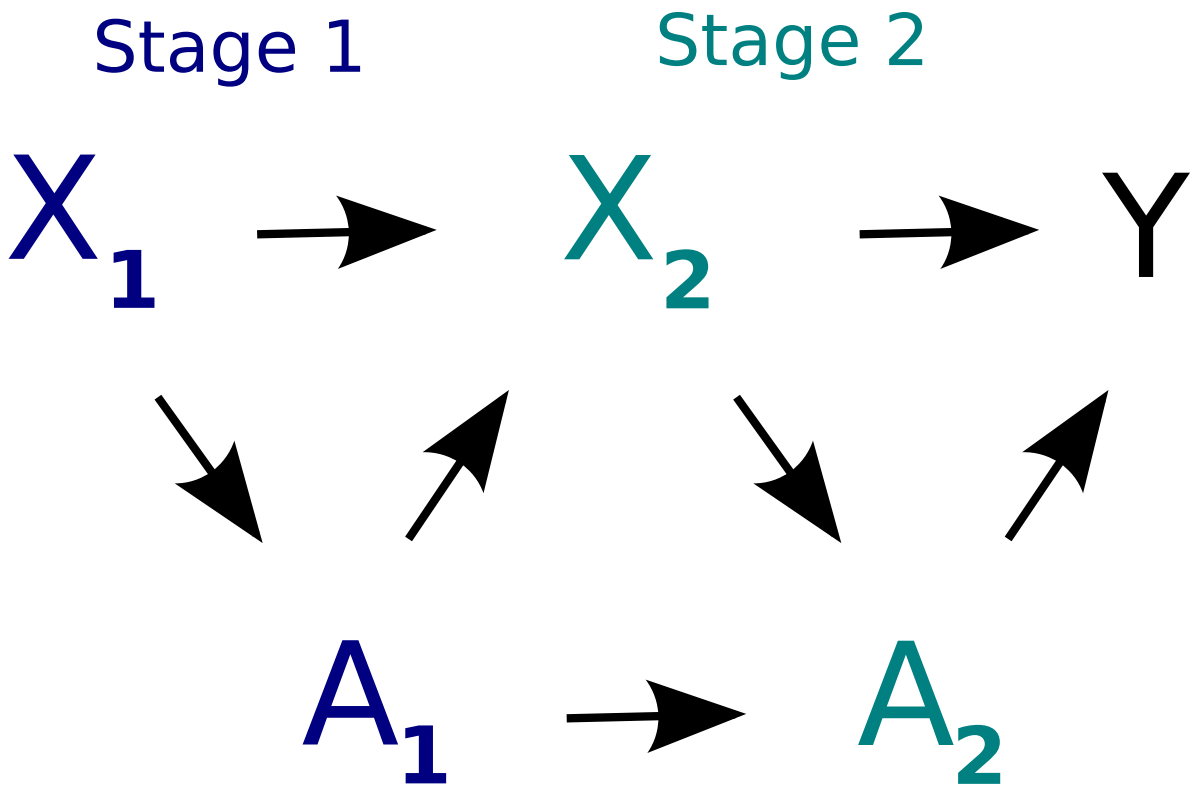
Notation



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

Identifying the best treatment regime: multi-stage

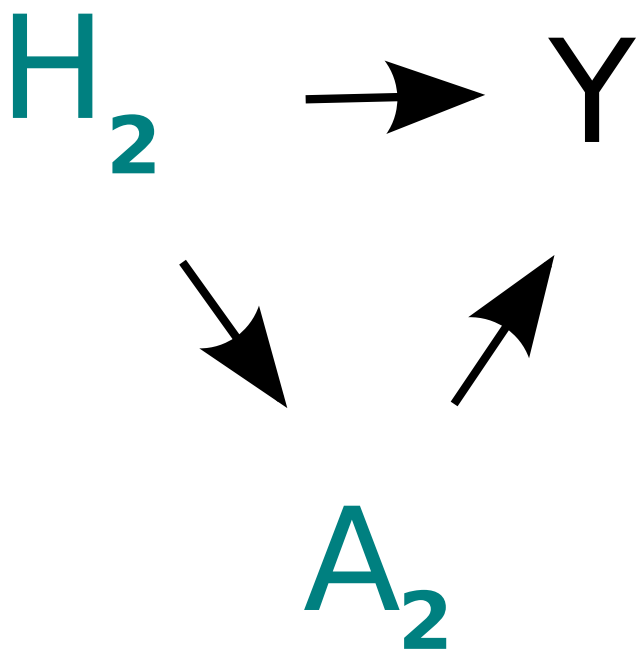
More often, we work in stages.



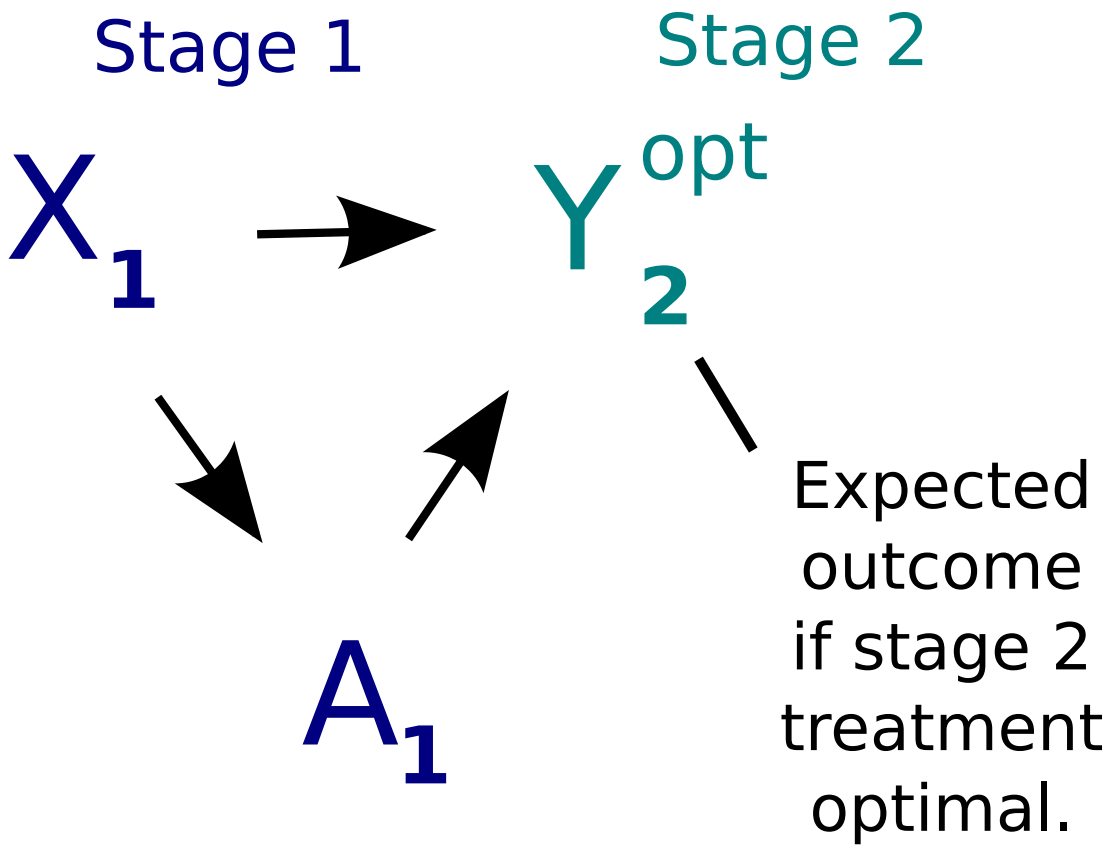
Identifying the best treatment regime: multi-stage

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

Stage 2



Identifying the best treatment regime: multi-stage



Lots of methods available:

Q-learning

MSMs

G-estimation

IPTW

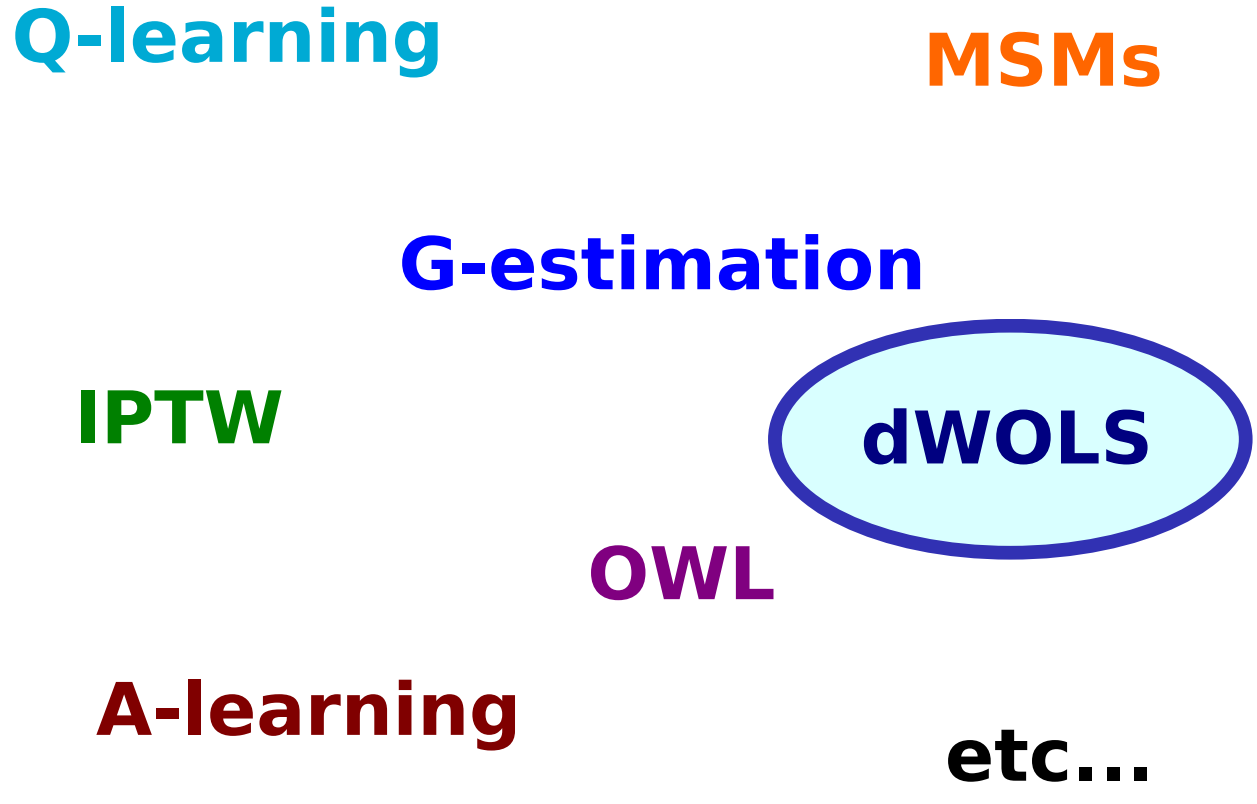
dWOLS

OWL

A-learning

etc...

Lots of methods available:



Identifying the best treatment regime

- If only one treatment decision:

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

- We might propose the following model

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

"Treat $(A = 1)$ if $A(\psi_0 + \psi_1 \text{BMI} > 0$"

Identifying the best treatment regime

- If only one treatment decision:

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

- We might propose the following model

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

"Treat $(A = 1)$ if $A(\psi_0 + \psi_1 \text{BMI} > 0$"

- 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:

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

- But we propose:

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

- Problem: what if A depends on BMI?

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:

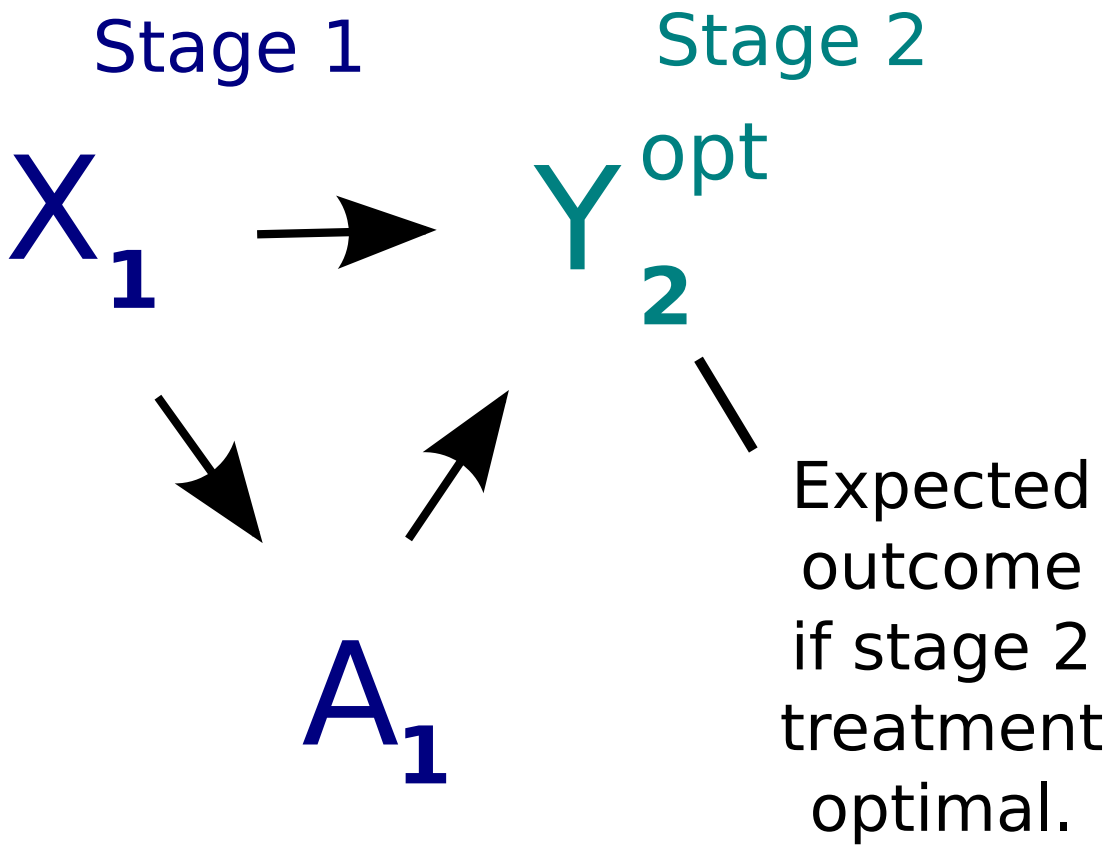
$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 \text{BMI} + \beta_2 \text{BMI}^2 + A(\psi_0 + \psi_1 \text{BMI})$$

- But we propose:

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

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

Multi-stage recursion



Multi-stage recursion

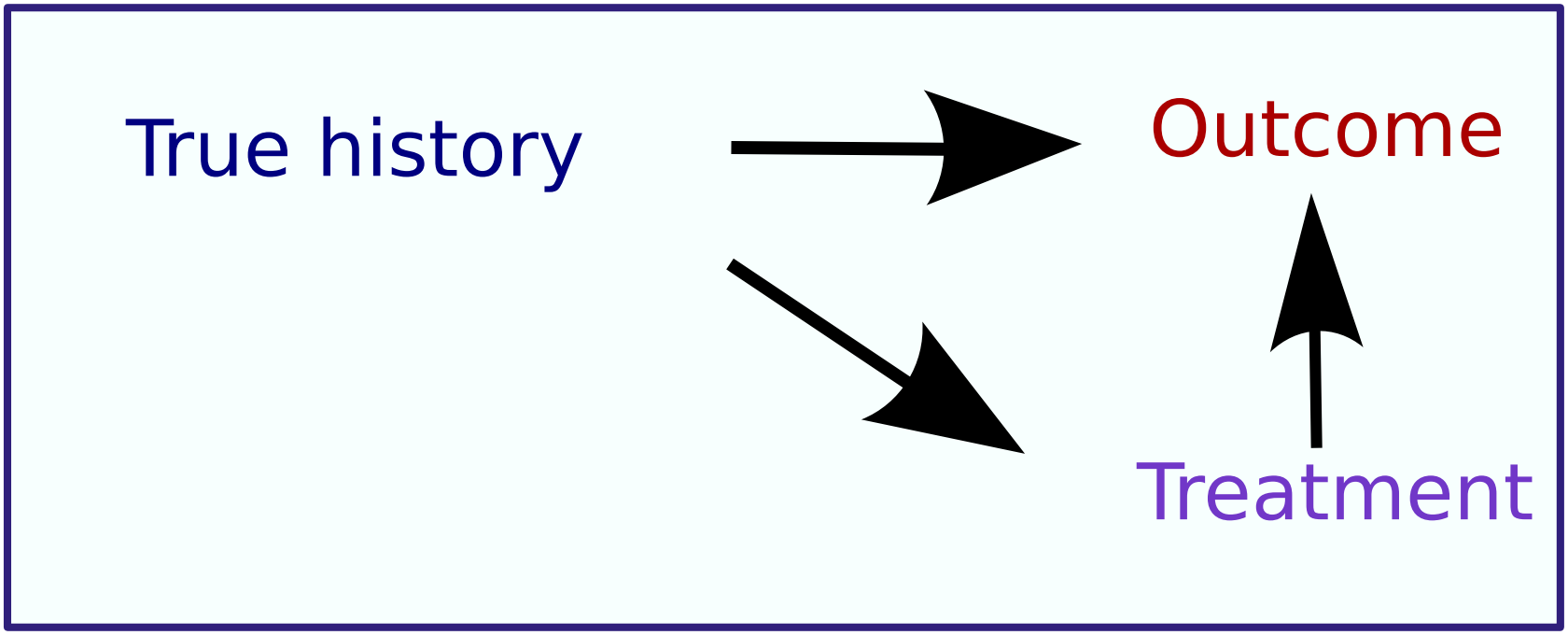
In the multi-stage setting we conduct a single-stage analysis at each stage by forming *pesudo-outcomes*:

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

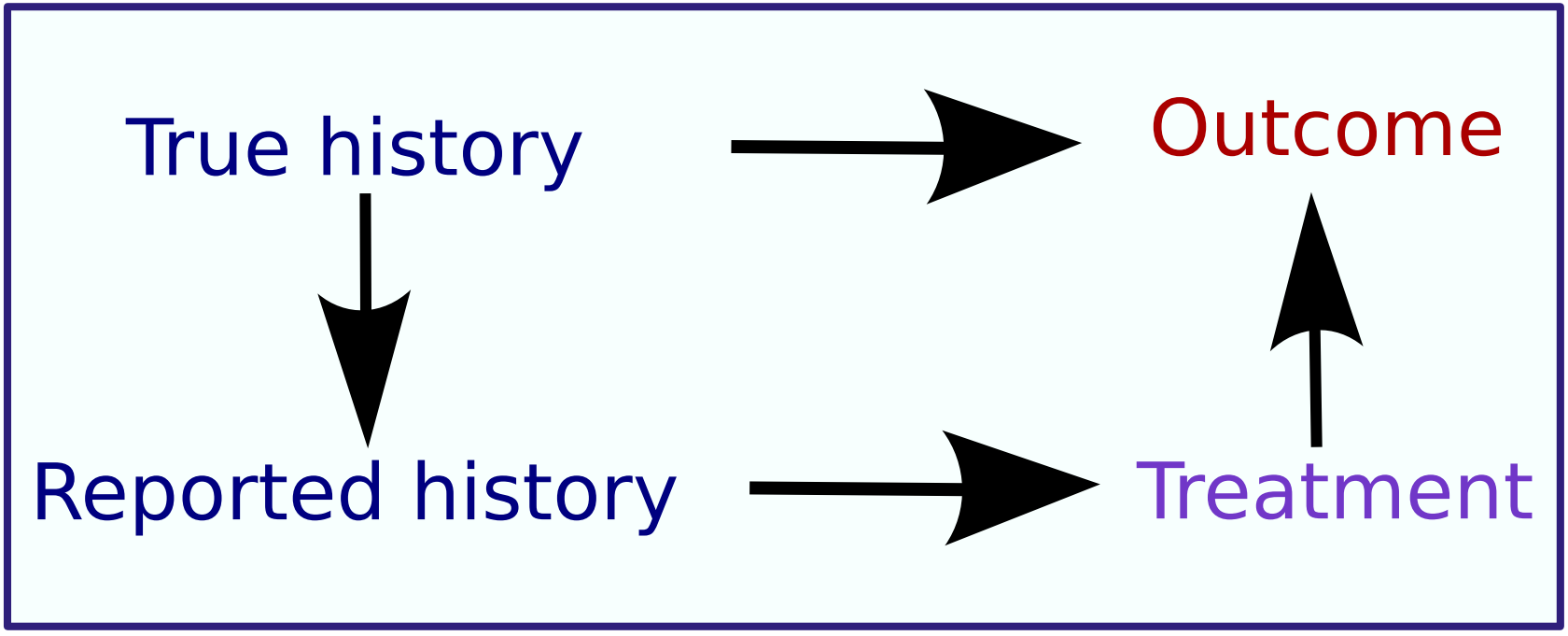
$\tilde{Y}_j$ is the expected outcome assuming optimal treatment from stage $j+1$ onwards.

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

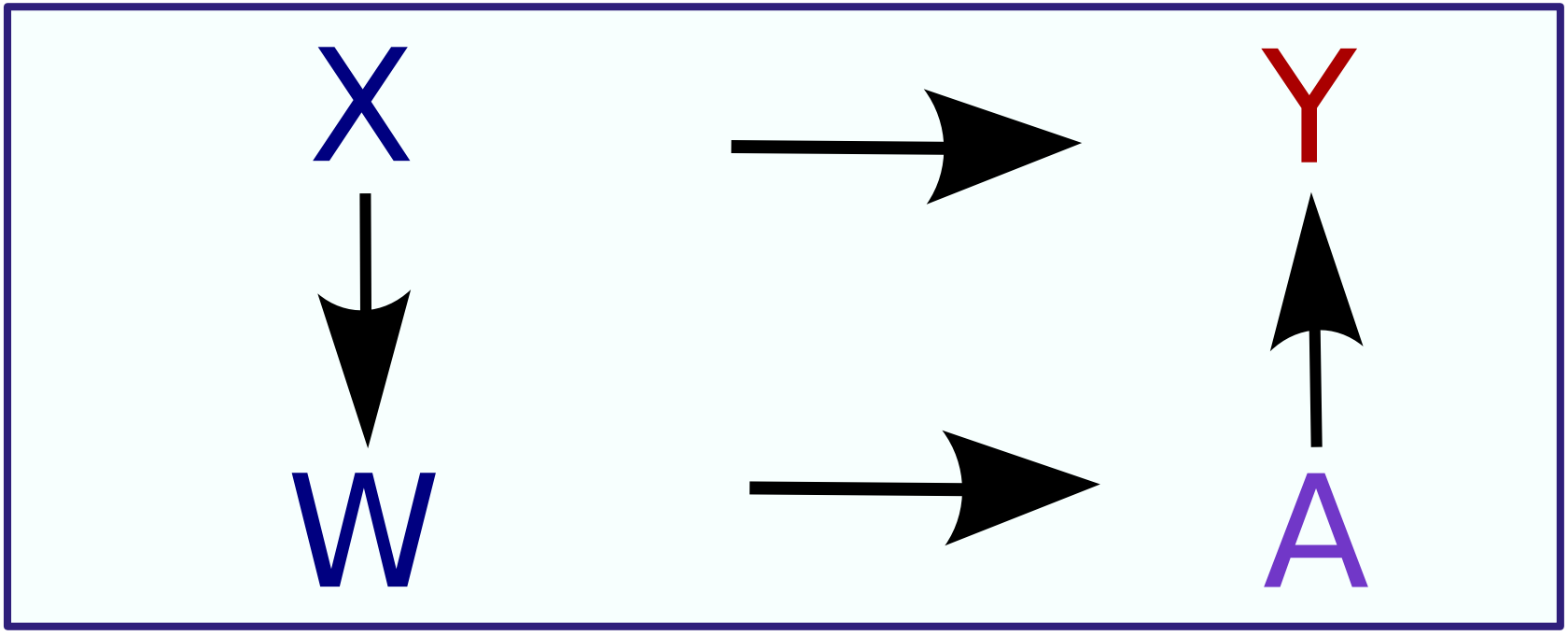
Measurement Error



Measurement Error



Measurement Error



Measurement Error

Assume: classical additive measurement error:

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

$$W = X + U$$

- $U \sim N(0, \sigma_u^2)$
- Non-differential: $Y \perp W|X$

Assume replicate measurements available on at least some patients.

Questions of interest

Three parts of the precision medicine puzzle:

- Estimation (and double robustness):
  $E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 \text{BMI} + \beta_2 \text{BMI}^2 + A(\psi_0 + \psi_1 \text{BMI})$
- Recursion: how do we form pseudo-outcomes?
- Future treatment: "Prescribe treatment if $\psi_0 + \psi_1 X > 0$"

Regression Calibration

Simple correction method: Regression Calibration.

Principle:

1. Use additional data to estimate $E[X|W, A] = X_{rc}$.
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

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

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

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

- A depends on W. Establish (approximate) covariate balance in X_{rc} by regressing A on X_{rc}.

Example (via simulation)

Outcome model

$$E[Y|X, A; \beta, \psi] = \beta_0 + \beta_1 X + \beta_2 X^2 + A(-1 + X)$$

So rule is "$A = 1$ is $-1 + X > 0$" (or $X > 1$).

A naive analysis returned the rule "$A = 1$ if $X > 2$"

“Isn’t this just a prediction problem?”

Various scenarios:

- Observed: W; Future: W
- Observed: W; Future: X
- Observed: X; Future: W
- Observed: X; Future: X

"Isn't this just a prediction problem?"

Various scenarios:

- Observed: W; Future: W
- Observed: W; Future: X
- Observed: X; Future: W
- Observed: X; Future: X $\leftarrow$ we've only studied this.

Question: is it worth obtaining replicates, validation data, etc. for future patients?

Future treatment

After estimating ψ, have "$A^{opt} = 1$ if $\hat{\psi}_0 + \hat{\psi}_1 X > 0$"

Suppose "$A^{opt} = 1$ if $X > \tau$" (or vice-versa) for 'threshold' τ

We observe $W = X + U$

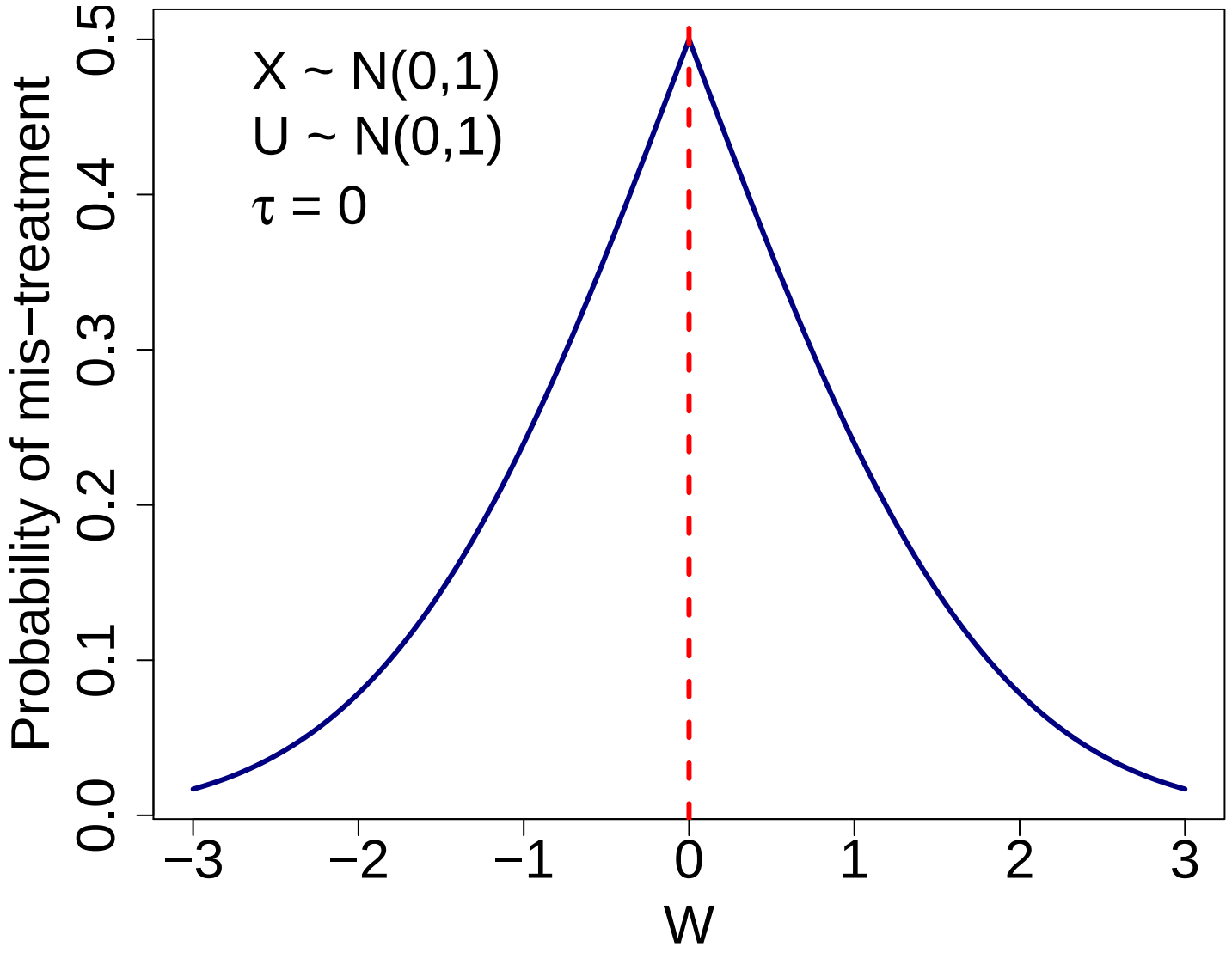
Questions of practical interest:

$$P(X < \tau | W = w > \tau) \quad P(X > \tau | W = w < \tau)$$

(e.g., if observed BMI $= 31$, the probability true BMI < 30)

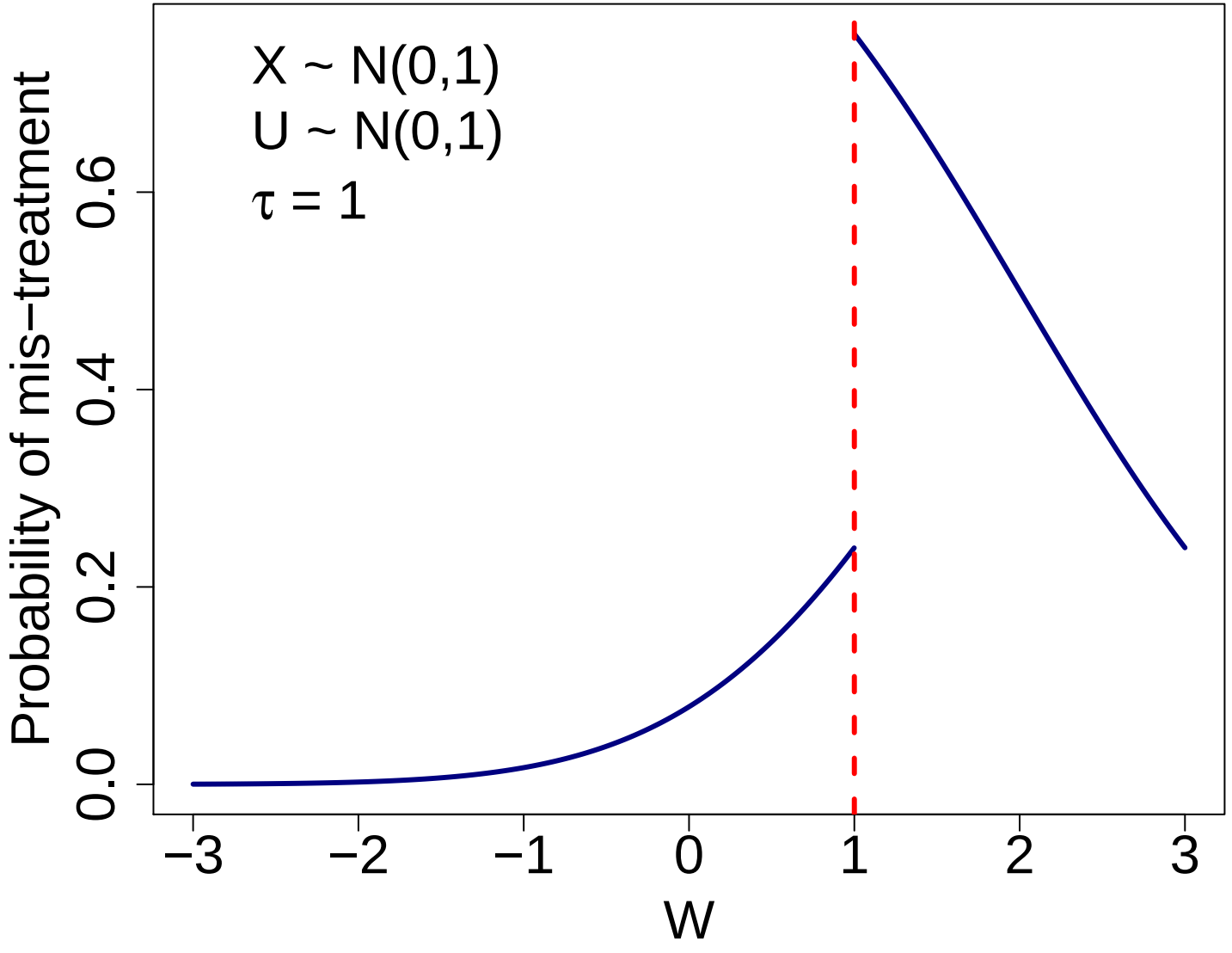
Future treatment

In some settings, results fairly intuitive:



Future treatment

In others, perhaps more of a surprise (to some):



Recursion

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

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

Question: What happens if we use W_k or X_{rc} instead of X_k? And what should we do about it?

Recursion

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

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

The problem: our pseudo-outcomes $\tilde{Y}_j$ will not be independent of future treatment.

Recursion

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

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

The problem: our pseudo-outcomes $\tilde{Y}_j$ will not be independent of future treatment.

One (possible) solution: if we have replicates W_{j1}, W_{j2}, with A_j based on W_{j1}, then form pseudo-outcoms based on W_{j2}.

Summary/Future Work

So far:

- Measurement error poses unique challenges in the precision medicine setting.
- Biased/incorrect treatment rules.
- Theoretical issues (double robustness, recursion).
- Consequences for future treatments tailored on error-prone observations.

Summary/Future Work

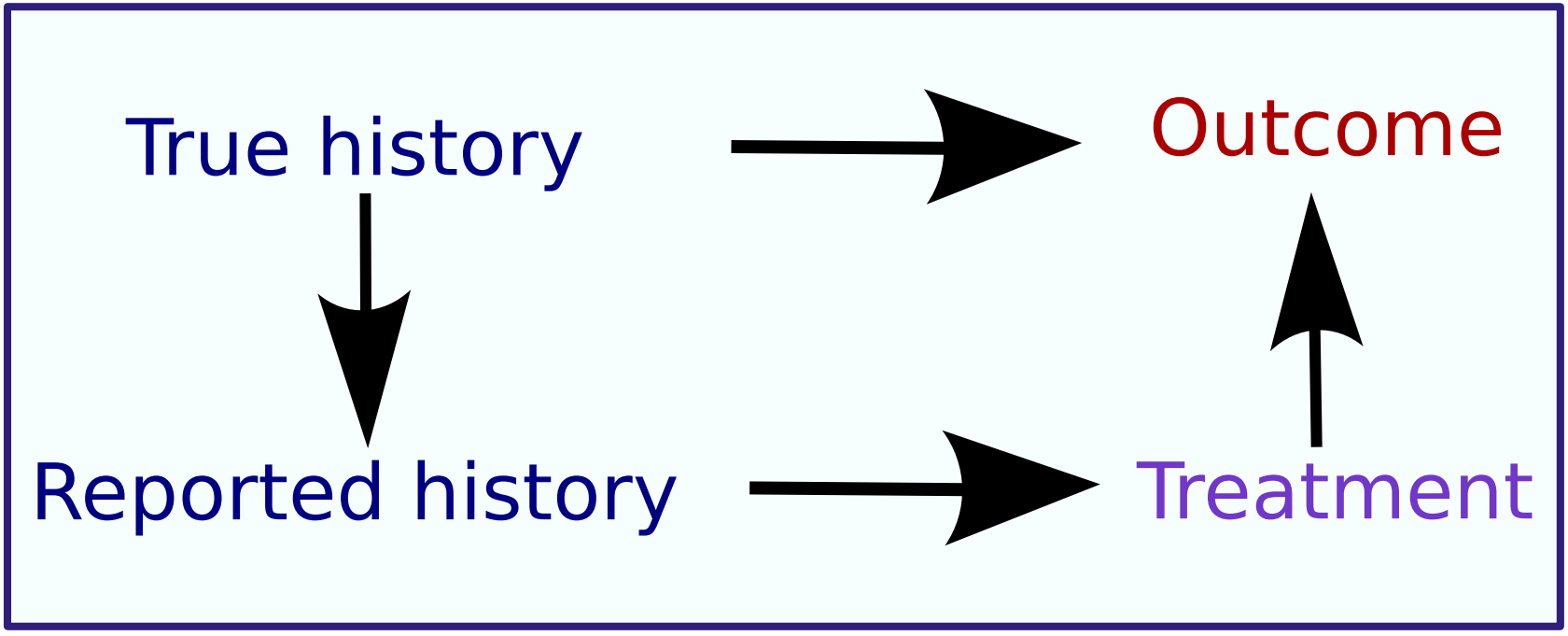
So far:

- Measurement error poses unique challenges in the precision medicine setting.
- Biased/incorrect treatment rules.
- Theoretical issues (double robustness, recursion).
- Consequences for future treatments tailored on error-prone observations.

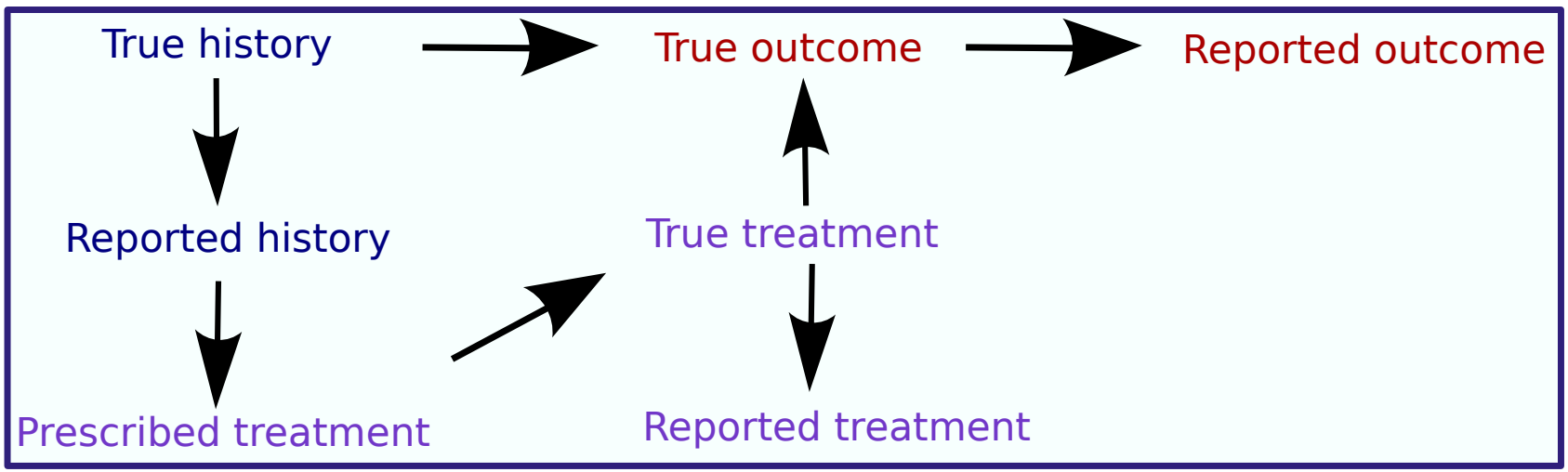
Moving forward:

- Other correction methods (SIMEX, conditional score, etc.).
- New methodological work specific to DTR/precision framework.
- Diagnostics for extant analyses/datasets.

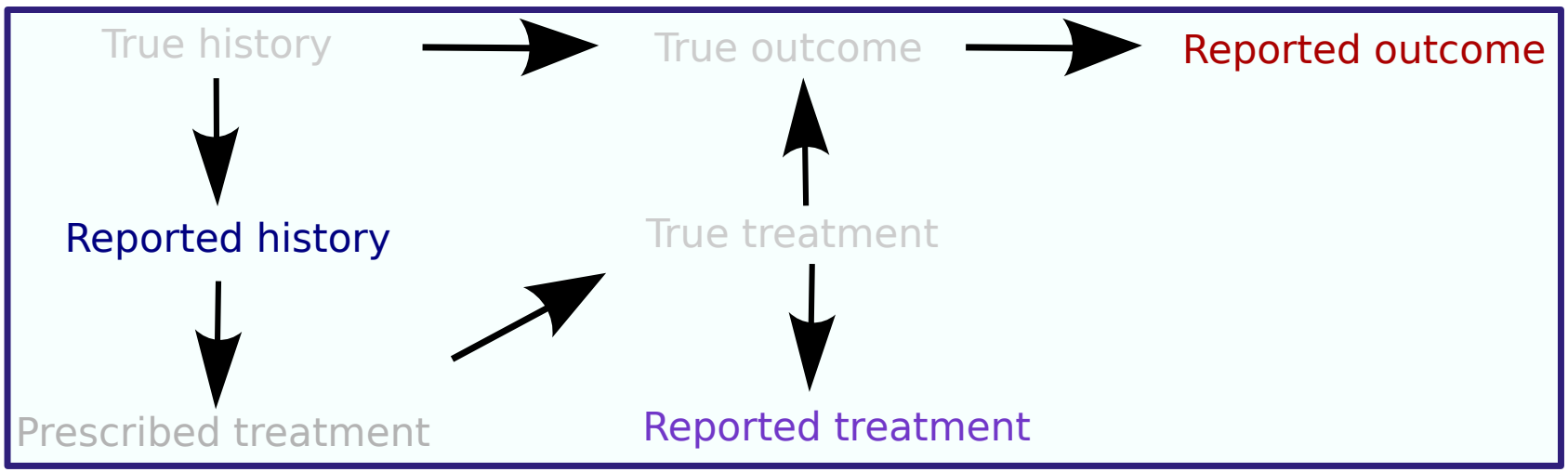
Measurement error



Measurement error



Measurement error



References/Acknowledgments

- A comprehensive guide to DTRs: B. Chakraborty and E. E. M. Moodie (2013). Statistical Methods for Dynamic Treatment Regimens. Springer: New York.
- 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.

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

NSERC CRSNG

UNIVERSITY OF WATERLOO