Penalized doubly robust regression-based estimation of adaptive treatment strategies

Erica E M Moodie

Biostatistics, McGill University
erica.moodie@mcgill.ca

Joint with Zeyu (Chris) Bian, Sahir Bhatnagar, Susan Shortreed, & Sylvie Lambert

Funding: NIMH, CIHR, & FRQS

Road-map

1. Precision medicine in the statistical literature
2. dWOLS: a regression-based method of estimation
3. Penalization and confounder selection
4. A case study

Evidence-based medicine: The statisticians' role

> [T]he medical statistician recognizes, and is familiar with the pros and cons of, that difficult question – should a fixed dose be given to all patients in a trial or should it be allowed to vary with the apparent needs of each patient as judged by the clinician?
>
> *Sir Austin Bradford Hill (1962)*

Precision medicine

- **Precision:** Refers to the tailoring of medical treatment to the individual characteristics of each patient.

Precision medicine

- **Precision:** Refers to the tailoring of medical treatment to the individual characteristics of each patient.
  - Does not mean selection or creation of drugs that are unique to a patient.
  - Does mean the ability to *classify individuals* into subpopulations that differ in their response to a specific treatments – perhaps repeatedly over time.

Precision medicine

- **Precision:** Refers to the tailoring of medical treatment to the individual characteristics of each patient.
  - Does not mean selection or creation of drugs that are unique to a patient.
  - Does mean the ability to *classify individuals* into subpopulations that differ in their response to a specific treatments – perhaps repeatedly over time.

Precision medicine

- The context:
  - several treatments or doses available;
  - patients may switch from one treatment to another;
  - often must account for clinician decisions about the treatment during patient monitoring.

Precision medicine

- The context:
  - several treatments or doses available;
  - patients may switch from one treatment to another;
  - often must account for clinician decisions about the treatment during patient monitoring.

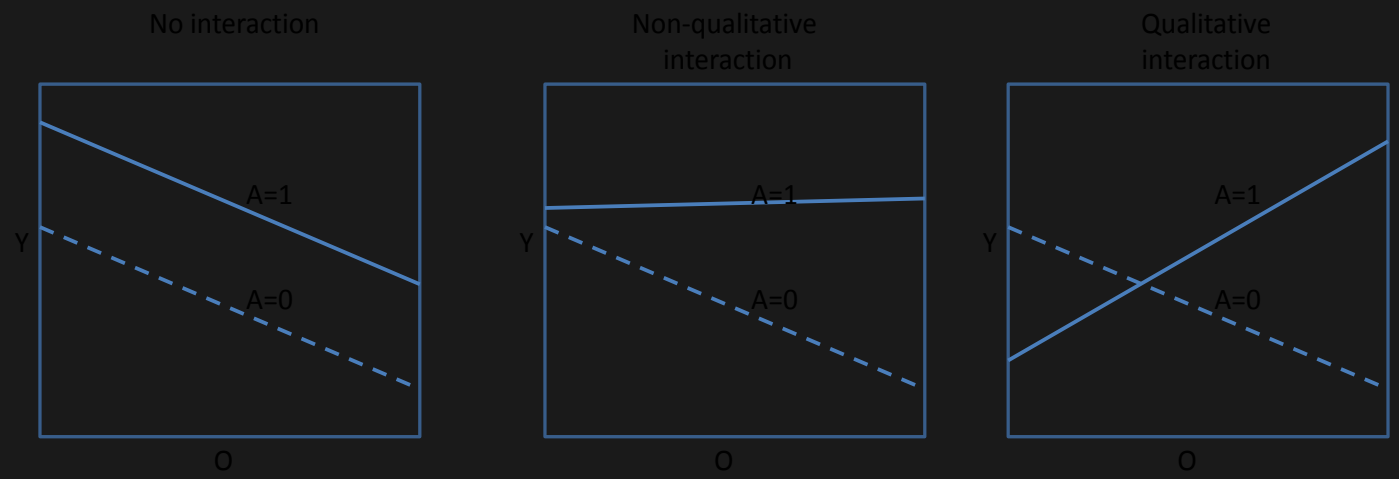
- Make use of information on patient characteristics such as
  - demographics, genetics, genomics;
  - physiologic or clinical measures;
  - medical history, etc.

  in order to determinate *which treatment* the patient should take *and when*.

When would we want treatment to be adaptive?

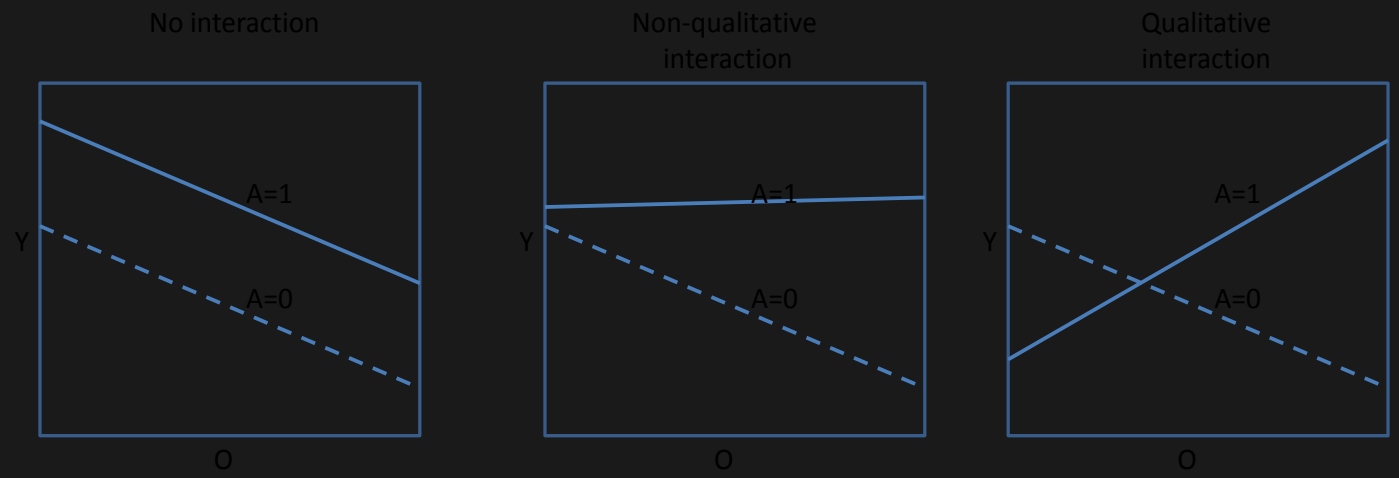
- Treatment tailoring is better because (when):
  - there is heterogeneity in patient response;
  - patient response may change over time;
  - response to treatment may inform future treatment choices;
  - patient compliance may be imperfect;
  - over-treating can lead to side-effects, treatment fatigue (poor compliance), and higher costs;
  - under-treating can lead to poorer patient outcomes.

When would we want treatment to be adaptive?



- Variables used to make treatment more targeted are called prescriptive or tailoring variables.

When would we want treatment to be adaptive?



- Variables used to make treatment more targeted are called prescriptive or tailoring variables.
- Resulting treatment algorithms known as adaptive treatment strategy (ATSs), dynamic treatment regimes, adaptive interventions, or policies.

Analytic methods

- Considerable interest in estimation and inference for ATSs in statistics and CS over the last 20 years.
- More common strategies include
  - Q-learning (sequential regression);
  - g-estimation;
  - dynamic weighted ordinary least squares (dWOLS);
  - weighted value search estimators, including weighted classifiers such as OWL and RWL, and (A)IPW.

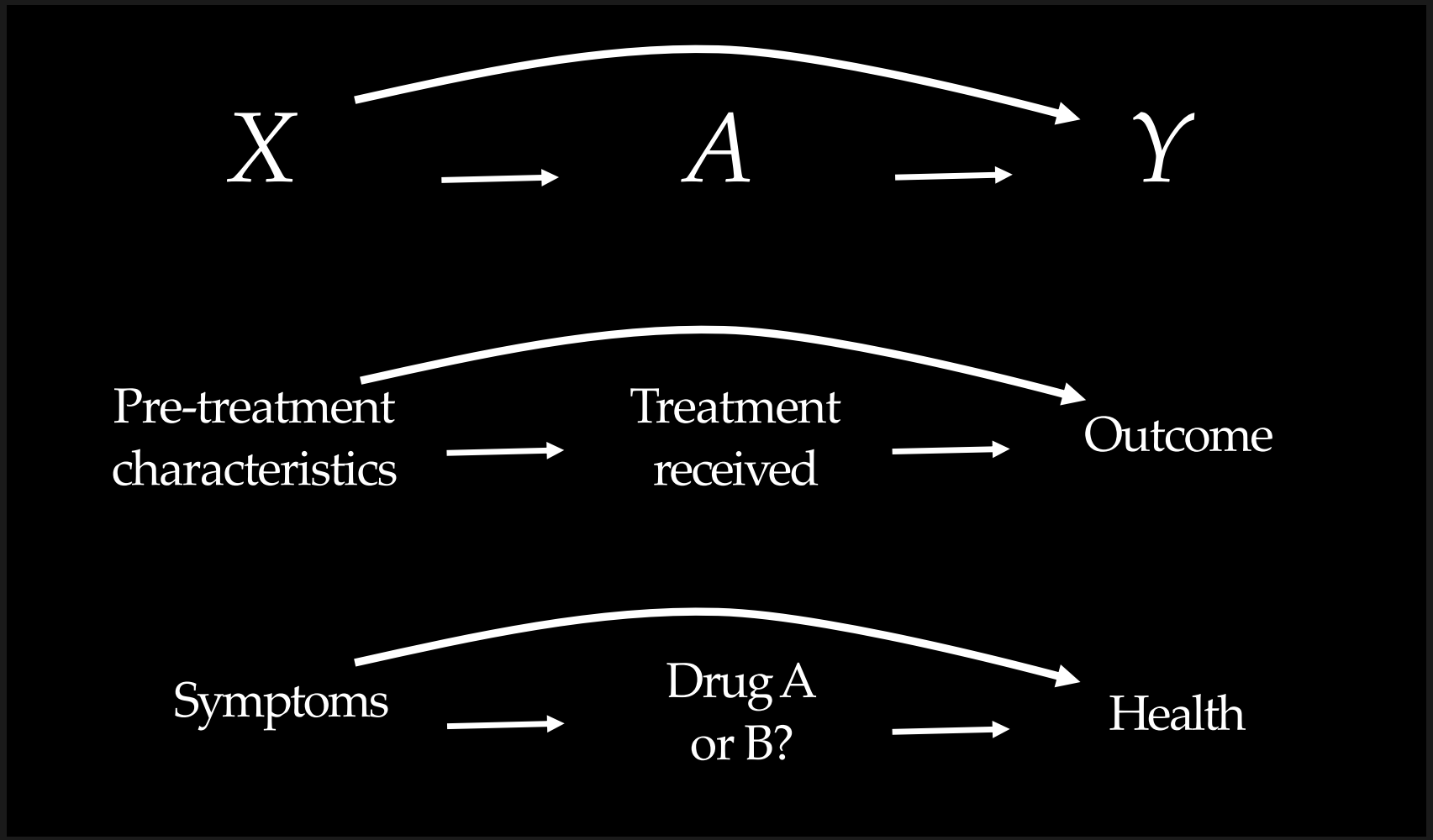
Aim

- Studies frequently collect many, many variables – but not all are useful for tailoring treatment, or for reducing confounding bias.
- Simplifying ATS can improve statistical efficiency, yield more practical treatment rules.

- Wanted to introduce methods of variable selection to choose most important confounders and critical tailoring variables in an ATS analysis within the doubly robust and highly user-friendly method of dWOLS.

The Single-stage Setting

Notation



ATSs: how do we find treatment A^{opt} that maximizes Y?

Identifying the best treatment regime

- If only one treatment decision: $\mathbb{E}[Y|X, A]$
- E.g., we might propose the following model

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

Identifying the best treatment regime

- If only one treatment decision: $\mathbb{E}[Y|X, A]$
- E.g., we might propose the following model

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

Impact of patient history in the absence of treatment

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

Identifying the best treatment regime

- If only one treatment decision: $\mathbb{E}[Y|X, A]$
- E.g., we might propose the following model

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

Impact of patient history in the absence of treatment

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

- In practice, we specify a model for the contrast, e.g.:

$$\gamma(x, a; \psi) = a(\psi_0 + \psi_1 x_1 + \psi_2 x_2),$$

then if $a \in \{0, 1\}$, $a^{\text{opt}}(x) = \mathbb{I}[\psi_0 + \psi_1 x_1 + \psi_2 x_2 > 0]$.

Identifying the best treatment regime

- If only one treatment decision: $\mathbb{E}[Y|X, A]$
- E.g., we might propose the following model

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

Impact of patient history in the absence of treatment

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

- In practice, we specify a model for the contrast, e.g.:

$$\gamma(x, a; \psi) = a(\psi_0 + \psi_1 x_1 + \psi_2 x_2),$$

  then if $a \in \{0, 1\}$, $a^{\text{opt}}(x) = \mathbb{I}[\psi_0 + \psi_1 x_1 + \psi_2 x_2 > 0]$.
- Contrasts specified in this way $\rightarrow$ *linear* decision rules.
- All methods considered assume $\gamma(x, a; \psi)$ well-specified.

Dynamic Weighted OLS (dWOLS)

$$\mathbb{E}[Y|X = x, A = a; \psi, \beta] = G(x; \beta) + \gamma(x, a; \psi)$$

- Three models to specify:
  1. Contrast model: $\gamma(x, a; \psi)$.
  2. Treatment-free model: $G(x; \beta)$.
  3. Treatment model: $\mathbb{E}[A|X = x; \alpha]$.
- Estimate ψ via WOLS with weights satisfying

$$\pi(x)w(1, x; \alpha) = (1 - \pi(x))w(0, x; \alpha),$$

for $\pi(x) = \mathbb{E}[A|X = x; \alpha]$, e.g. $w = |A - \mathbb{E}[A|x; \hat{\alpha}]|$, inverse probability of treatment weighting, etc.

dWOLS: editorial comments

- Appealing:
  - doubly robust;
  - (quite) easy to explain to clinical collaborators;
  - has been generalized to accommodate multiple treatments, and to continuous doses (GdWOLS);
  - extended to handle censored outcomes using accelerated failure time models (dwSurv);
  - many useful tools including residual diagnostics, model-validation based on double-robustness;
  - `DTRreg` implements many useful cases: binary or normally-distributed treatments for continuous outcomes, binary treatments for time-to-event outcomes.

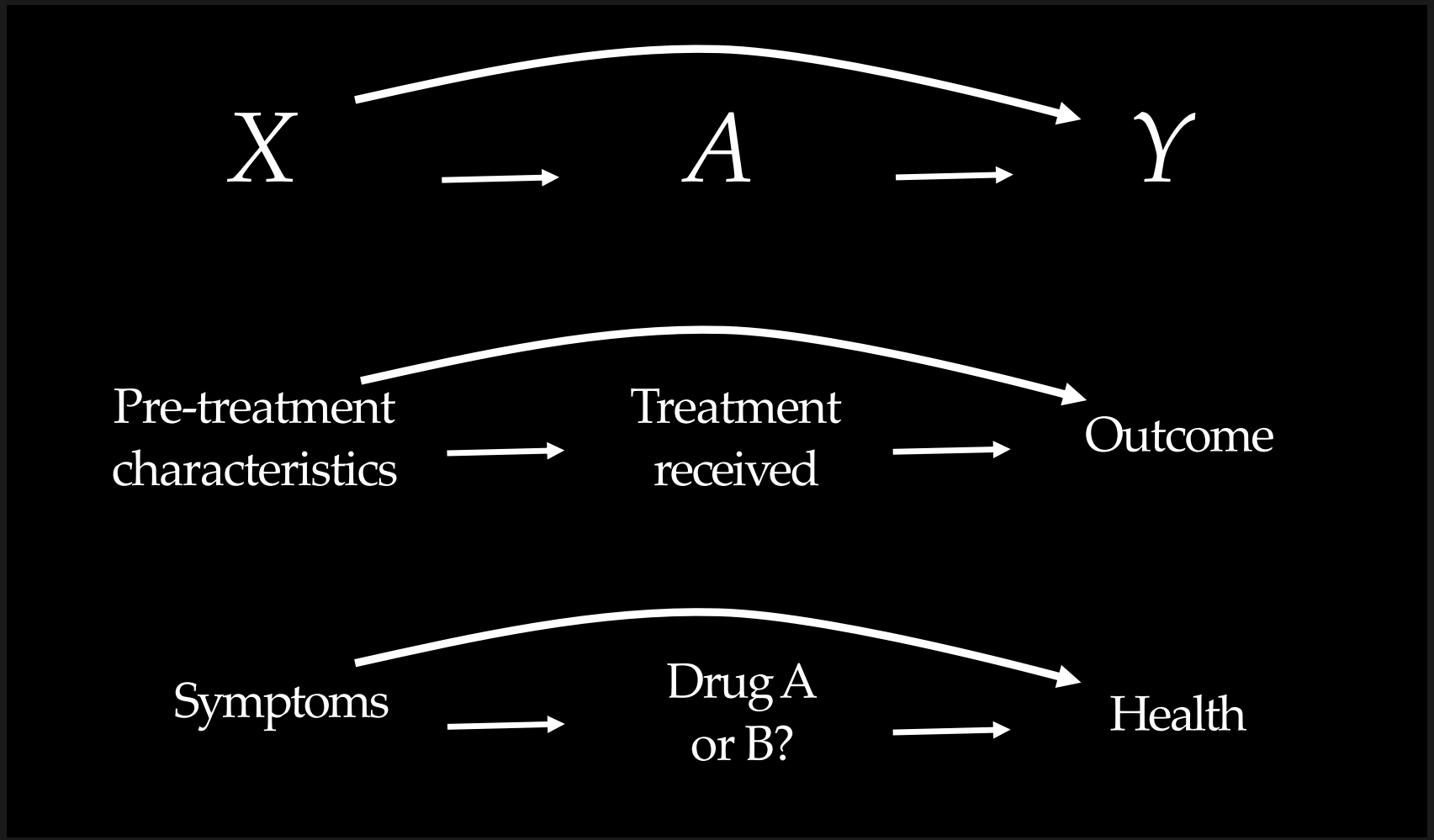
dWOLS: editorial comments

- Appealing:
  - doubly robust;
  - (quite) easy to explain to clinical collaborators;
  - has been generalized to accommodate multiple treatments, and to continuous doses (GdWOLS);
  - extended to handle censored outcomes using accelerated failure time models (dwSurv);
  - many useful tools including residual diagnostics, model-validation based on double-robustness;
  - `DTRreg` implements many useful cases: binary or normally-distributed treatments for continuous outcomes, binary treatments for time-to-event outcomes.
- Challenge/limitation:
  - limited results for discrete outcomes;
  - to date, no data-driven variable selection methods.

The Multi-stage Setting

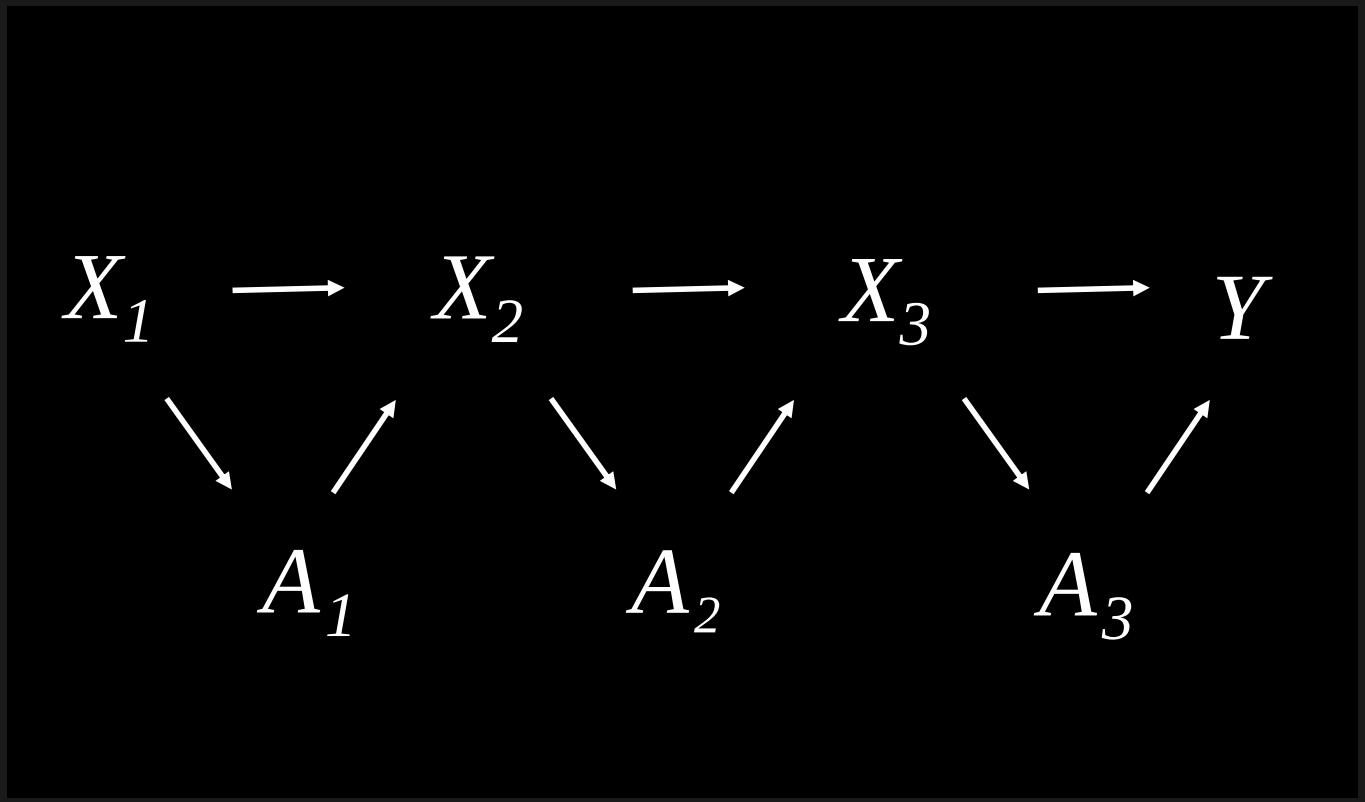
Identifying the best treatment regime: multi-stage

Recall data set up in the ITR setting:



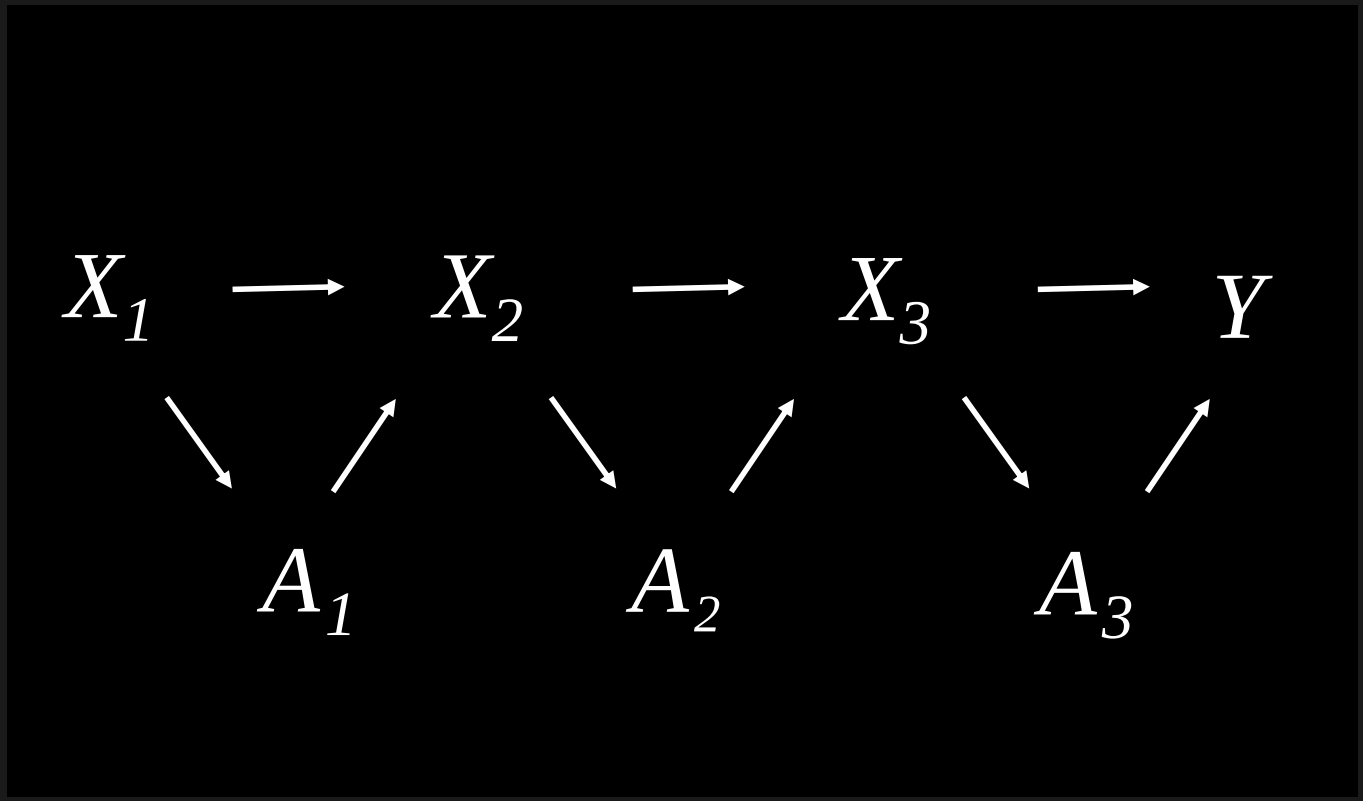
Identifying the best treatment regime: multi-stage

The multi-stage case is more complex:
Want a sequence $(a_1^{\text{opt}}, a_2^{\text{opt}}, a_3^{\text{opt}})$ that maximizes Y, but the choice of A_j affects future decisions.



Identifying the best treatment regime: multi-stage

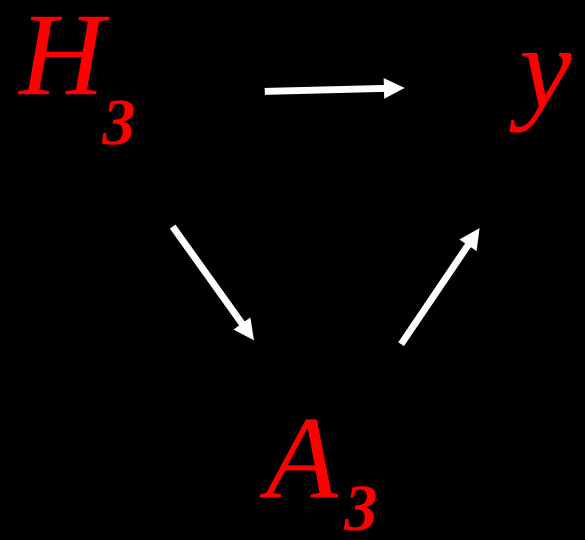
The multi-stage case is more complex:
Want a sequence $(a_1^{\text{opt}}, a_2^{\text{opt}}, a_3^{\text{opt}})$ that maximizes Y, but the choice of A_j affects future decisions.



Recursive implementation reduces estimation to a series of one-stage problems.

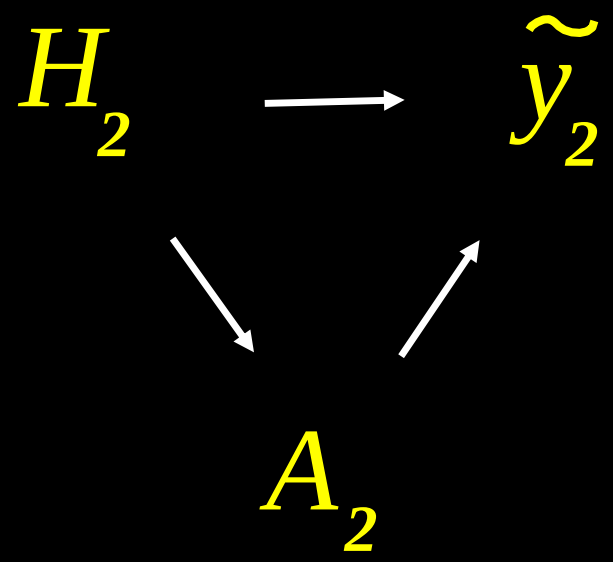
Identifying the best treatment regime: multi-stage

Letting $H_3 = (X_1, A_1, X_2, A_2, X_3)$ reduces finding A_3^{opt} to a single-stage problem.



Identifying the best treatment regime: multi-stage

Writing $H_2 = (X_1, A_1, X_2)$ reduces finding A_2^{opt} to a single-stage problem, where $\tilde{Y}_2$ is taken to be the "best" possible stage 2 outcome for someone with H_2.



Identifying the best treatment regime: multi-stage

- The $\tilde{Y}$ term is key to the backwards induction, sequential regression approach.
- It is called a pseudo-outcome.
- $\tilde{Y}$ in dWOLS *equals* the observed outcome for individuals who were treated optimally; otherwise, the observed outcome is "adjusted up" as predicted by the contrast function.

Assumptions

- Like many causal methods, dWOLS relies on some identifiability and estimation assumptions:
  - SUTVA
  - no unmeasured confounding, measurement error, or selective information
  - positivity
  - correct model specification (blip + one of treatment-free or propensity score)
  - hierarchy: any tailoring variable in the blip must appear in the treatment-free model.
- Of these, the last is key when considering any form of variable selection.

Selecting tailoring variables: pdWOLS

- Much like LASSO, we want to use an ℓ_1 penalty with the following objective function

$$Q(\boldsymbol{\theta}) = \mathcal{L}(Y; \beta, \psi) + \lambda(1-\alpha)\|\beta\|_1 + \lambda\alpha\|\boldsymbol{\psi}\|_1,$$

  for tuning parameters $\lambda > 0$ and $\alpha \in (0, 1)$, and squared error loss function

$$\mathcal{L}(Y; \beta, \psi) = \frac{1}{2n}\left\|\sqrt{W}\left(Y - \psi_0 A - \sum_{j=1}^{p} X_j\beta_j - \sum_{j=1}^{p}\psi_j(A \circ X_j)\right)\right\|_2^2.$$

- But this doesn't assure strong heredity.

Selecting tailoring variables: pdWOLS

- Instead, we reparameterize the squared error loss function:

$$\mathcal{L}(Y;\theta) = \frac{1}{2n}\left\|\sqrt{W}\left(Y - \psi_0 A - \sum_{j=1}^{p} X_j\beta_j - \sum_{j=1}^{p}\underbrace{\psi_0\tau_j\beta_j}_{\psi_j}(A \circ X_j)\right)\right\|_2^2$$

  and use objective function

$$Q(\theta) = \mathcal{L}(Y;\theta) + \lambda(1-\alpha)\|\beta\|_1 + \lambda\alpha\|\boldsymbol{\tau}\|_1$$

  for θ the collection of all parameters.
- Can be performed sequentially in a multi-stage ATS setting.

Selecting tailoring variables: pdWOLS

- Tuning parameter selection via a value information criterion (Shi et al., 2021): $VIC = n\widehat{V}(\psi) - \kappa_n \|\psi\|_0$ where

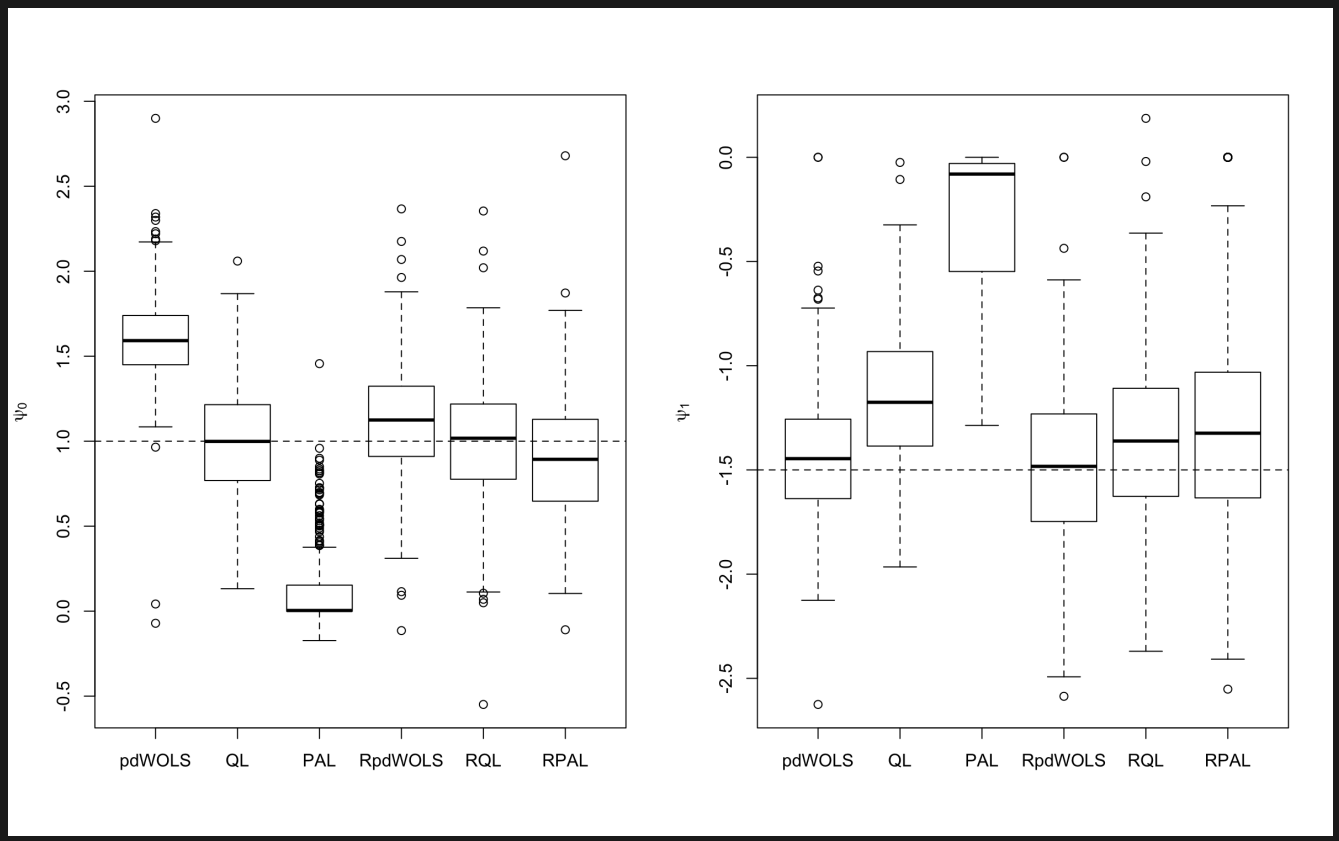
$$\widehat{V}(\psi) = \frac{1}{n}\sum_{i=1}^{n} Y_i \frac{A_i \widehat{a}^{opt}(x_i;\psi) + (1-A_i)(1-\widehat{a}^{opt}(x_i;\psi))}{A_i\widehat{\pi}(x_i) + (1-A_i)(1-\widehat{\pi}(x_i))}$$

  and κ_n is a positive sequence.
- We choose $\kappa_n = n^{1/3} log(p)^{2/3} log\ log(n)$ as this can achieve model selection consistency.

- Note that this will only accurately estimate the value function if the propensity score is correctly estimated.

Selecting tailoring variables

- Performance in simulation is excellent, even in difficult settings ($p = 400$, $n = 200$).
  - Excellent accuracy in selective, treatment recommendation.
  - Refitting after selection recommended to reduce bias.



Selecting a confounder set

- In a causal analysis, the propensity score (PS) serves as a means of breaking confounding.
- The goal of PS model building is thus to create *balance*, not to predict treatment allocation.
- In particular, instruments should not be included in a PS.
- Shortreed & Ertefaie (2017) proposed a method that considers both the treatment-covariate and the outcome-covariates relationships: outcome adaptive lasso.
  - Idea: Use initial estimate of coefficients in the outcome model as an adaptive lasso weight in the PS.
- Simulations show that combining OAL with pdWOLS improves accuracy of value function estimation, and slightly improves accuracy of tailoring variable selection.

Case study: Web-based stress reduction

- CVD, like many chronic diseases, have the potential to create significant stress and anxiety in patients.
- Lambert et al. conducted a pilot SMART to assess the feasibility and potential effect size of a stepped care approach using a web-based stress-management intervention.
- We focus on the first stage:
  - 50 patients with CVD were randomized into two arms (stratified by recruitment source and stress level.
  - Arms: website only group ($A = 0$), and website plus weekly telephone coaching group ($A = 1$).
- Outcome: Depression Anxiety Stress Scales
  - Goal is to minimize

Study sample

Characteristics of the adaptive web-based stress management study population

mean (SD) or *n* (%)	Website-only	Website+coach	SMD
n	25	25	
Age	61.0 (13.4)	61.3 (10.7)	0.03
DASS at baseline	19.2 (6.7)	18.9 (7.7)	0.03
Physical Component Score	42.9 (12.2)	45.4 (11.6)	0.20
Mental Component Score	40.4 (8.6)	40.9 (10.9)	0.04
Sex = Male	13 (52.0)	9 (36.0)	0.33
Marital = Married / Common law	14 (56.0)	15 (60.0)	0.08
Education = University degree	13 (52.0)	16 (64.0)	0.25
Employment = Full time	4 (16.0)	9 (36.0)	0.47
Chronic cardiac condition	18 (72.0)	16 (64.0)	0.17
Chronic hypertension	12 (48.0)	10 (40.0)	0.16
Chronic stomach condition	10 (40.0)	12 (48.0)	0.16
Chronic vision condition	11 (44.0)	7 (28.0)	0.34
Chronic back pain	10 (40.0)	14 (56.0)	0.32
Chronic cholesterol condition	9 (36.0)	11 (44.0)	0.16
Chronic obesity	7 (28.0)	13 (52.0)	0.50
Chronic osteoarthritis	9 (36.0)	8 (32.0)	0.09

Study sample

Characteristics of the adaptive web-based stress management study population

mean (SD) or *n* (%)	Website-only	Website+coach	SMD
n	25	25	
Age	61.0 (13.4)	61.3 (10.7)	0.03
DASS at baseline	19.2 (6.7)	18.9 (7.7)	0.03
Physical Component Score	42.9 (12.2)	45.4 (11.6)	0.20
Mental Component Score	40.4 (8.6)	40.9 (10.9)	0.04
Sex = Male	13 (52.0)	9 (36.0)	0.33
Marital = Married / Common law	14 (56.0)	15 (60.0)	0.08
Education = University degree	13 (52.0)	16 (64.0)	0.25
Employment = Full time	4 (16.0)	9 (36.0)	0.47
Chronic cardiac condition	18 (72.0)	16 (64.0)	0.17
Chronic hypertension	12 (48.0)	10 (40.0)	0.16
Chronic stomach condition	10 (40.0)	12 (48.0)	0.16
Chronic vision condition	11 (44.0)	7 (28.0)	0.34
Chronic back pain	10 (40.0)	14 (56.0)	0.32
Chronic cholesterol condition	9 (36.0)	11 (44.0)	0.16
Chronic obesity	7 (28.0)	13 (52.0)	0.50
Chronic osteoarthritis	9 (36.0)	8 (32.0)	0.09

Results

- Of the 16 potential tailoring variables, the following 7 were selected:

  DASS at baseline, age, mental component score, physical component score, vision conditions, sex, & employment status

- Optimal rule:

  $\widehat{a}^{opt} = I\{16.7 + 0.4\text{age} - 0.6\text{DASS} - 0.1\text{PCS} - 0.5\text{MCS} - 4.7I(\text{vision=yes}) - 5.2I(\text{male}) - 7.1I(\text{employed}) > 0\}$.

- Estimated DASS (95% CI) under this rule: $14.8(11.8, 17.9)$.

- If covariate imbalance was ignored, a slightly more complex optimal rule is obtained with 8 tailoring variables, which yields a slightly lower (better) predicted DASS at 13.8.

Observations

- ATSs have enormous potential to improve clinical practice in an evidence-based way.
- Lots of data (big n) needed to learn about ATS.
- Observational data can be particularly attractive:
  - Less expense than an RCT;
  - More information (large sample size);
  - Generalizable or 'characterizable' population;
  - But many, many variables we don't need.
- Trials ensure randomization, but balance may still be imperfect and again... often more measurements than 'useful.'

Concluding remarks

- Many methodological limitations remain.
- Estimation approaches that accommodate 'fixes' for the many challenges help.
- Further investigation of the VIC also critical: a 'singly robust' criterion for a doubly robust approach?
- An important avenue to pursue will is the extension of these methods to the discrete outcome setting.

Thanks




NIH National Institute of Mental Health

Fonds de recherche Santé Québec

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