Covariate Adjustment in Randomized Controlled Trials

Kelly Van Lancker

UNIVERSITEIT GENT

FACULTY OF SCIENCES

Big thank you to

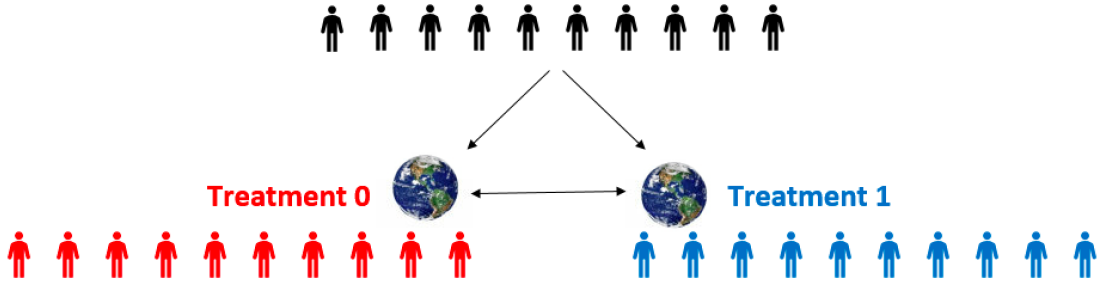
- Marlena S. Bannick (University of Washington)
- Josh Betz (Johns Hopkins University)
- Frank Bretz (Novartis)
- Iván Díaz (New York University)
- Oliver Dukes (Ghent University)
- Michael Rosenblum (Johns Hopkins University)
- Stijn Vansteelandt (Ghent University)
- Ting Ye (University of Washington)

Outline

Hypothetical World

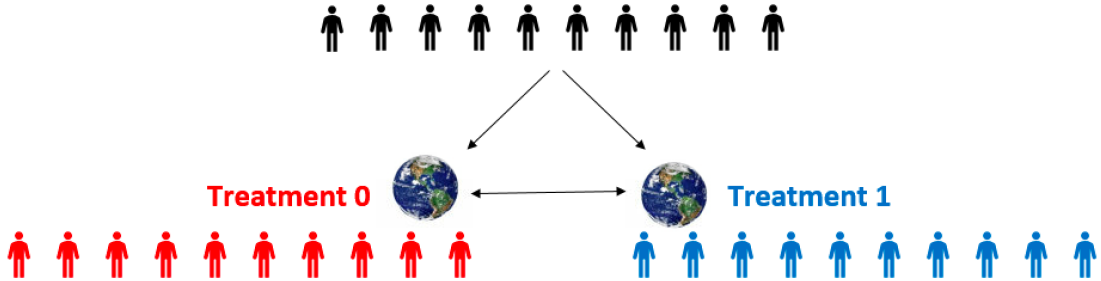
- Consider an eligible patient population.

Hypothetical World



- Consider an eligible patient population.
- Imagine two parallel worlds: one where everyone is assigned Treatment 0 and one where everyone is assigned Treatment 1.

Hypothetical World



- Consider an eligible patient population.
- Imagine two parallel worlds: one where everyone is assigned Treatment 0 and one where everyone is assigned Treatment 1.
  - Denote Y^0 and Y^1 the **potential/hypothetical outcomes** in the two parallel worlds.

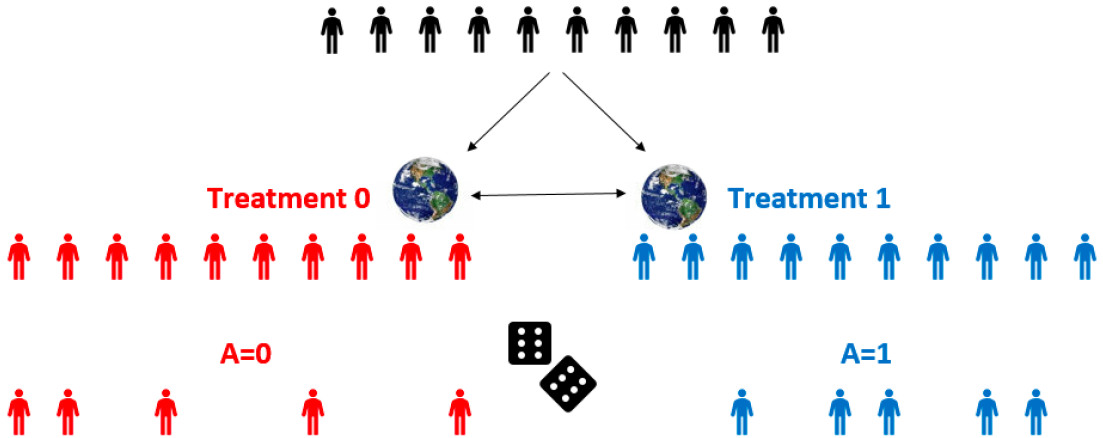
Marginal Causal Contrasts

- Causal contrasts of interest often reflect a contrast between the means of the distributions of Y^0 and Y^1: $E\left(Y^0\right)$ and $E\left(Y^1\right)$
  - Mean difference $E\left(Y^1\right) - E\left(Y^0\right)$
  - Mean ratio $E\left(Y^1\right) / E\left(Y^0\right)$
  - Odds ratio $\frac{E(Y^1)/\{1-E(Y^1)\}}{E(Y^0)/\{1-E(Y^0)\}}$
  - ...

Marginal Causal Contrasts

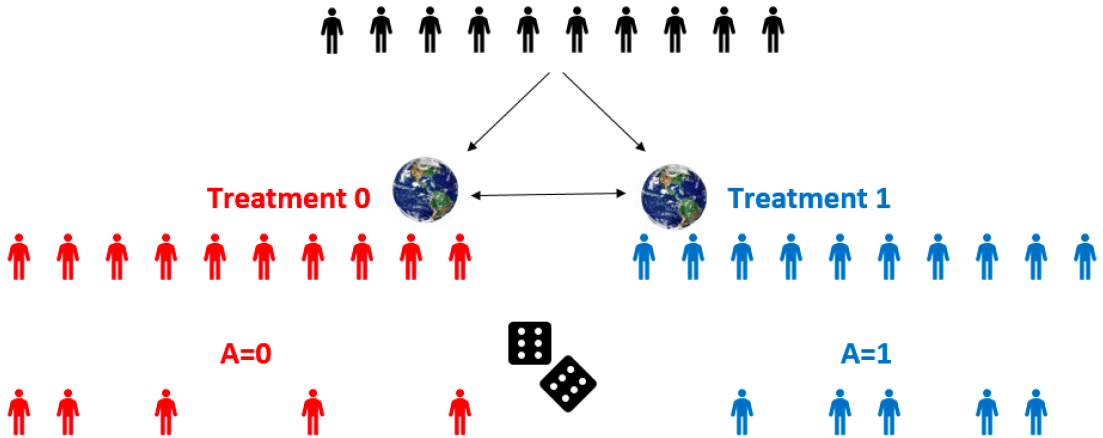
- Causal contrasts of interest often reflect a contrast between the means of the distributions of Y^0 and Y^1: $E\left(Y^0\right)$ and $E\left(Y^1\right)$
  - Mean difference $E\left(Y^1\right) - E\left(Y^0\right)$
  - Mean ratio $E\left(Y^1\right) / E\left(Y^0\right)$
  - Odds ratio $\frac{E(Y^1)/\{1-E(Y^1)\}}{E(Y^0)/\{1-E(Y^0)\}}$
  - ...
- These are **marginal/unconditional** causal contrasts.
- The (marginal) causal contrast can also be a contrast of other summaries of the distributions of Y^0 and Y^1; e.g., for time-to-event outcomes.

Real world: Randomization



- In real life, patients are randomized to only one group.
  - The randomized group is denoted A and the factual/observed outcome Y.

Real world: Randomization



- In real life, patients are randomized to only one group.
  - The randomized group is denoted A and the factual/observed outcome Y.
- Randomization ensures that causal contrasts correspond to statistical contrasts:
  - $E\left(Y^1\right) - E\left(Y^0\right) = E\left(Y|A=1\right) - E\left(Y|A=0\right)$.

Conditional Causal Contrasts

- Causal contrasts of interest can also reflect a contrast between the means of the distributions of Y^0 and Y^1 **in a subset of patients** (e.g., females):
  - e.g., mean difference $E\left(Y^1|sex = f\right) - E\left(Y^0|sex = f\right)$

Conditional Causal Contrasts

- Causal contrasts of interest can also reflect a contrast between the means of the distributions of Y^0 and Y^1 **in a subset of patients** (e.g., females):
  - e.g., mean difference $E\left(Y^1|sex = f\right) - E\left(Y^0|sex = f\right)$
- These are **conditional** (i.e., within stratum of baseline variable(s)) estimands.

Conditional Causal Contrasts

- Causal contrasts of interest can also reflect a contrast between the means of the distributions of Y^0 and Y^1 **in a subset of patients** (e.g., females):
  - e.g., mean difference $E\left(Y^1|sex = f\right) - E\left(Y^0|sex = f\right)$
- These are **conditional** (i.e., within stratum of baseline variable(s)) estimands.
- Randomization ensures that

$$E\left(Y^1|sex = f\right) - E\left(Y^0|sex = f\right) = E\left(Y|A = 1, sex = f\right) - E\left(Y|A = 0, sex = f\right).$$

Conditional Causal Contrasts

- Causal contrasts of interest can also reflect a contrast between the means of the distributions of Y^0 and Y^1 **in a subset of patients** (e.g., females):
  - e.g., mean difference $E\left(Y^1|sex = f\right) - E\left(Y^0|sex = f\right)$
- These are **conditional** (i.e., within stratum of baseline variable(s)) estimands.
- Randomization ensures that

  $E\left(Y^1|sex = f\right) - E\left(Y^0|sex = f\right) = E\left(Y|A = 1, sex = f\right) - E\left(Y|A = 0, sex = f\right).$
- However, estimation typically requires **model assumptions** (such as logistic regression model) and the estimate is often **uninterpretable** under model misspecification.

Conditioning versus Adjusting

- **Confusing**: Traditionally '**conditional**' and '**adjusted**' have been used interchangeably, likewise for '**marginal**' and '**unadjusted**'.

Conditioning versus Adjusting

- **Confusing**: Traditionally '**conditional**' and '**adjusted**' have been used interchangeably, likewise for '**marginal**' and '**unadjusted**'.
- Better suggestion:
  - marginal/conditional are related to the '**estimand**' of interest
  - unadjusted/adjusted are related to the '**analysis**' performed

Conditioning versus Adjusting

- **Confusing**: Traditionally '**conditional**' and '**adjusted**' have been used interchangeably, likewise for '**marginal**' and '**unadjusted**'.
- Better suggestion:
  - marginal/conditional are related to the '**estimand**' of interest
  - unadjusted/adjusted are related to the '**analysis**' performed
- Perfectly possible to obtain an adjusted estimator of a marginal estimand.
  - **Adjusted estimators of marginal estimands are almost always more precise than unadjusted estimators.**

Conditioning versus Adjusting

- **Confusing**: Traditionally '**conditional**' and '**adjusted**' have been used interchangeably, likewise for '**marginal**' and '**unadjusted**'.
- Better suggestion:
  - marginal/conditional are related to the '**estimand**' of interest
  - unadjusted/adjusted are related to the '**analysis**' performed
- Perfectly possible to obtain an adjusted estimator of a marginal estimand.
  - **Adjusted estimators of marginal estimands are almost always more precise than unadjusted estimators.**
- Recent FDA guidelines make a distinction between conditioning and adjusting (FDA, 2023).
  - Recommendations for covariate adjustment.
  - Advice on both conditional, and marginal estimands.

Outline

Conditional Causal Contrasts

- Until now, our estimand definitions have been completely **model-free**.
- In practice, treatment effects are usually encoded as **parameters in a generalised linear model**; e.g.

$$g\{E(Y|A,X)\} = \beta_0 + \beta_1 A + \beta_2 X$$

where $g(\cdot)$ is a pre-specified link function.

Conditional Contrasts: Continuous Outcome

- Let's consider a continuous outcome Y (e.g., bloodpressure).

Conditional Contrasts: Continuous Outcome

- Let's consider a continuous outcome Y (e.g., bloodpressure).
- One may then fit the model

$$E(Y|A, X) = \beta_0 + \beta_1 A + \beta_2 X.$$

Conditional Contrasts: Continuous Outcome

- Let's consider a continuous outcome Y (e.g., bloodpressure).
- One may then fit the model

$$E(Y|A, X) = \beta_0 + \beta_1 A + \beta_2 X.$$

- **Statistical modelling assumption**: no interaction between A and X on the linear scale
  - **Not implied by randomisation**.

Conditional Contrasts: Continuous Outcome

- Let's consider a continuous outcome Y (e.g., bloodpressure).
- One may then fit the model

$$E(Y|A,X) = \beta_0 + \beta_1 A + \beta_2 X.$$

- **Statistical modelling assumption**: no interaction between A and X on the linear scale
  - **Not implied by randomisation**.
- If assumption holds:
  - β_1 carries an interpretation as *both* an **age-specific** (i.e., conditional) causal effect and a **marginal** causal effect.
  - This is not necessarily true if one includes interactions! (Ye et al., 2022)

Conditional Causal Contrasts: Other Outcomes

- For a binary outcome Y, it is more common to choose the logistic regression model

$$logit\{E(Y|A,X)\} = \beta_0 + \beta_1 A + \beta_2 X.$$

- If the model reflects the truth, then the effect of treatment (β_1) does not differ for different values of X.

Conditional Causal Contrasts: Other Outcomes

- For a binary outcome Y, it is more common to choose the logistic regression model

$$logit\{E(Y|A,X)\} = \beta_0 + \beta_1 A + \beta_2 X.$$

- If the model reflects the truth, then the effect of treatment (β_1) does not differ for different values of X.

- Unlike in the linear case, $exp(\beta_1)$ would *only* retain an interpretation as a **conditional effect**,

$$\frac{E(Y^1|X=x)/\{1-E(Y^1|X=x)\}}{E(Y^0|X=x)/\{1-E(Y^0|X=x)\}},$$

which may **differ from the marginal** causal odds ratio

$$\frac{E(Y^1)/\{1-E(Y^1)\}}{E(Y^0)/\{1-E(Y^0)\}}.$$

Conditional Causal Contrasts: Other Outcomes

- **Standard practice based on logistic regression does not typically target a marginal effect.**

Conditional Causal Contrasts: Other Outcomes

- **Standard practice based on logistic regression does not typically target a marginal effect.**
- This phenomenon occurs due to the **non-collapsibility** of the logistic link function; see Daniel et al. (2021).
  - Not unique to logistic regression; e.g., Cox proportional hazards models.

Conditional Causal Contrasts: Other Outcomes

- **Standard practice based on logistic regression does not typically target a marginal effect.**
- This phenomenon occurs due to the **non-collapsibility** of the logistic link function; see Daniel et al. (2021).
  - Not unique to logistic regression; e.g., Cox proportional hazards models.
- When the model is misspecified, the standard likelihood-based estimators of β_1 may not generally target either

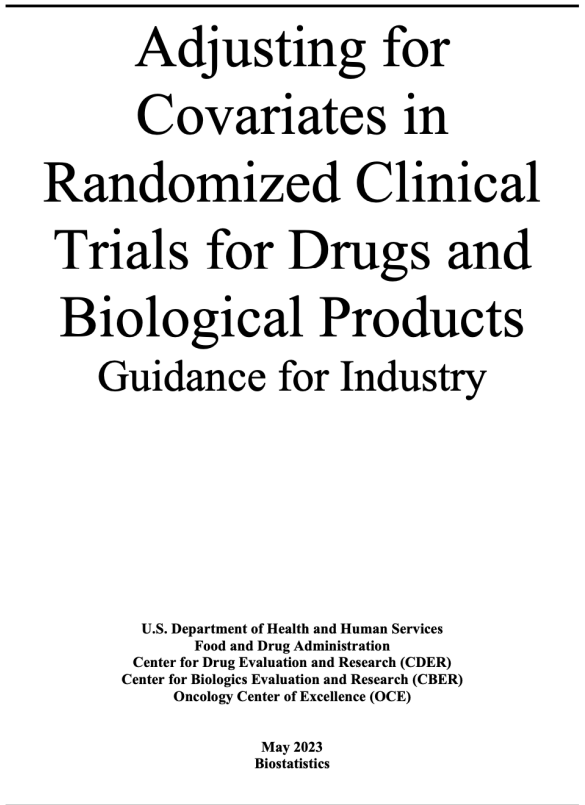
$$\frac{E(Y^1|X=x)/\{1-E(Y^1|X=x)\}}{E(Y^0|X=x)/\{1-E(Y^0|X=x)\}} \text{ or } \frac{E(Y^1)/\{1-E(Y^1)\}}{E(Y^0)/\{1-E(Y^0)\}}.$$

  - The concern for model misspecification for non-linear models is for example highlighted in the (EMA, 2015) guideline.

Outline

FDA guidance on covariate adjustment

"after suitably addressing the treatment effect definition, **covariate adjustment** using linear or nonlinear models can be used **to increase precision**."



Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products

Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

May 2023
Biostatistics

Covariate Adjustment for Marginal Estimands

- **Covariate adjustment** is a statistical analysis method with high potential to **improve precision** for many trials.
  - **Pre-planned** adjustment for baseline variables when estimating **average treatment effect**.
  - Estimand is same as when using unadjusted estimator (e.g., difference in means).
  - **Goal**: avoid making any model assumptions beyond what's assumed for unadjusted estimator (**robustness to model misspecification**).

  (e.g., Koch et al., 1998; Yang and Tsiatis, 2001; Rubin and van der Laan, 2008; Tsiatis et al., 2008; Moore and van der Laan, 2009b,a; Zhang, 2015; Jiang et al., 2018; Benkeser et al., 2020)

- In what follows, we focus on binary and continuous endpoints.

Example

- Suppose we aim to learn the treatment effect on a binary outcome Y (e.g., 'disease').

Age	A	Y	Y^1	Y^0
40	1	1	1	?
50	1	0	0	?
60	1	1	1	?
50	0	0	?	0
30	0	1	?	1
40	0	0	?	0

Example

- Suppose we aim to learn the treatment effect on a binary outcome Y (e.g., 'disease').

Age	A	Y	Y^1	Y^0
40	1	1	1	?
50	1	0	0	?
60	1	1	1	?
50	0	0	?	0
30	0	1	?	1
40	0	0	?	0

- By randomization: fine to compare outcomes of treated with outcomes of untreated.

Example

- Suppose we aim to learn the treatment effect on a binary outcome Y (e.g., 'disease').

Age	A	Y	Y^1	Y^0
40	1	1	1	?
50	1	0	0	?
60	1	1	1	?
50	0	0	?	0
30	0	1	?	1
40	0	0	?	0

- By randomization: fine to compare outcomes of treated with outcomes of untreated.
- Based on baseline covariates (e.g., age): guesses about what outcome would be for all participants if they were (un)treated.
  - **By using the models that were used to obtain conditional estimates.**

Covariate adjusted estimator: Model fitting

Step 1: Model fitting

Fit a logistic regression model for outcome Y given treatment allocation A and Age.

Covariate adjusted estimator: Model fitting

Step 1: Model fitting

Fit a logistic regression model for outcome Y given treatment allocation A and Age.

- Simple model [Focus challenge]:

$$P(Y = 1|A, Age) = logit^{-1}(\beta_0 + \beta_1 A + \beta_2 Age).$$

Covariate adjusted estimator: Model fitting

Step 1: Model fitting

Fit a logistic regression model for outcome Y given treatment allocation A and Age.

- Simple model [Focus challenge]:

$$P(Y = 1|A, Age) = logit^{-1}(\beta_0 + \beta_1 A + \beta_2 Age).$$

- Model with interactions:

$$P(Y = 1|A, Age) = logit^{-1}(\beta_0 + \beta_1 A + \beta_2 Age + \beta_3 A \cdot Age).$$

Covariate adjusted estimator: Model fitting

Step 1: Model fitting

Fit a logistic regression model for outcome Y given treatment allocation A and Age.

- Simple model [Focus challenge]:

$$P(Y = 1|A, Age) = logit^{-1}(\beta_0 + \beta_1 A + \beta_2 Age).$$

- Model with interactions:

$$P(Y = 1|A, Age) = logit^{-1}(\beta_0 + \beta_1 A + \beta_2 Age + \beta_3 A \cdot Age).$$

  - Two logistic regression models, one per arm.

Covariate adjusted estimator: Predicting

Step 2: Predicting

Use

$$P(Y = 1|A, Age) = logit^{-1}(\beta_0 + \beta_1 A + \beta_2 Age)$$

to impute outcome **under treatment ($A = 1$) and control ($A = 0$)** for **all** patients:

Age	A	Y	Y^1	$\hat{P}^1$	Y^0	$\hat{P}^0$
40	1	1	1	0.8	?	0.7
50	1	0	0	0.6	?	0.55
60	1	1	1	0.7	?	0.6
50	0	0	?	0.7	0	0.6
30	0	1	?	0.6	1	0.5
40	0	0	?	0.5	0	0.45

- $\hat{P}^1$ for patient i: $logit^{-1}(\widehat{\beta}_0 + \widehat{\beta}_1 + \widehat{\beta}_2 Age_i)$
- $\hat{P}^0$ for patient i: $logit^{-1}(\widehat{\beta}_0 + \widehat{\beta}_2 Age_i)$

Covariate adjusted estimator: Averaging

Step 3: **Averaging**

Take the average of imputed outcomes:

Covariate adjusted estimator: Averaging

Step 3: Averaging

Take the average of imputed outcomes:

- Compute standardized estimators for treatment specific means
  - $\widehat{E}\left(Y^1\right) = \frac{1}{n}\sum_{i=1}^{n} \hat{P}_i^1$
  - $\widehat{E}\left(Y^0\right) = \frac{1}{n}\sum_{i=1}^{n} \hat{P}_i^0$
- Calculate treatment effect of interest:
  - Mean difference $\hat{E}\left(Y^1\right) - \hat{E}\left(Y^0\right)$
  - Odds ratio $\frac{\hat{E}(Y^1)/\{1-\hat{E}(Y^1)\}}{\hat{E}(Y^0)/\{1-\hat{E}(Y^0)\}}$

Covariate adjusted estimator: Averaging

Step 3: Averaging

Take the average of imputed outcomes:

- Compute standardized estimators for treatment specific means
  - $\widehat{E}\left(Y^1\right) = \frac{1}{n}\sum_{i=1}^{n} \hat{P}_i^1$
  - $\widehat{E}\left(Y^0\right) = \frac{1}{n}\sum_{i=1}^{n} \hat{P}_i^0$
- Calculate treatment effect of interest:
  - Mean difference $\hat{E}\left(Y^1\right) - \hat{E}\left(Y^0\right)$
  - Odds ratio $\frac{\hat{E}(Y^1)/\{1-\hat{E}(Y^1)\}}{\hat{E}(Y^0)/\{1-\hat{E}(Y^0)\}}$

Mean of predictions based on glm's

Simulation Results

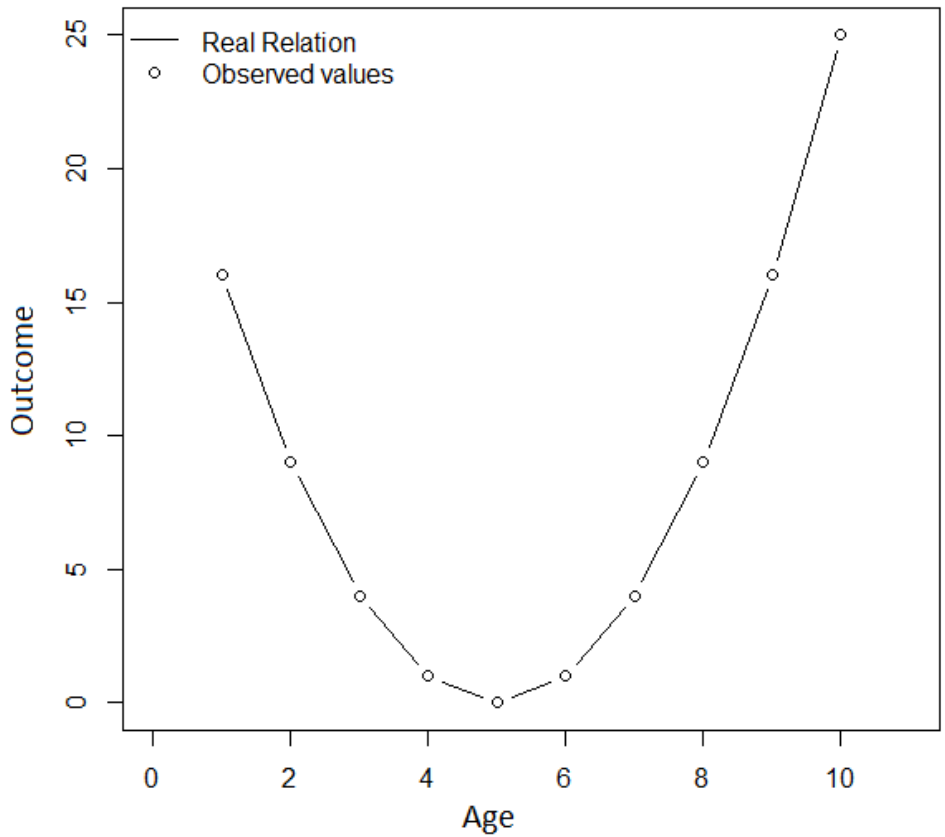
Results for binary outcome and risk difference under correctly specified models

n	Effect	Estimator type	Bias	Power	MSE	RE
100	-0.201	Unadj.	0.025	0.463	0.829	1.000
		Adj.	0.023	0.607	0.755	0.911
200	-0.201	Unadj.	0.010	0.821	0.864	1.000
		Adj.	-0.001	0.895	0.749	0.867
500	-0.126	Unadj.	-0.013	0.798	0.979	1.000
		Adj.	-0.007	0.862	0.850	0.868
1000	-0.091	Unadj.	0.012	0.837	0.898	1.000
		Adj.	0.020	0.892	0.817	0.910

Results from Benkeser et al. (2020) "Improving precision and power in randomized trials for COVID-19 treatments using covariate adjustment, for binary, ordinal, and time-to-event outcomes." Biometrics.

What if models are misspecified?

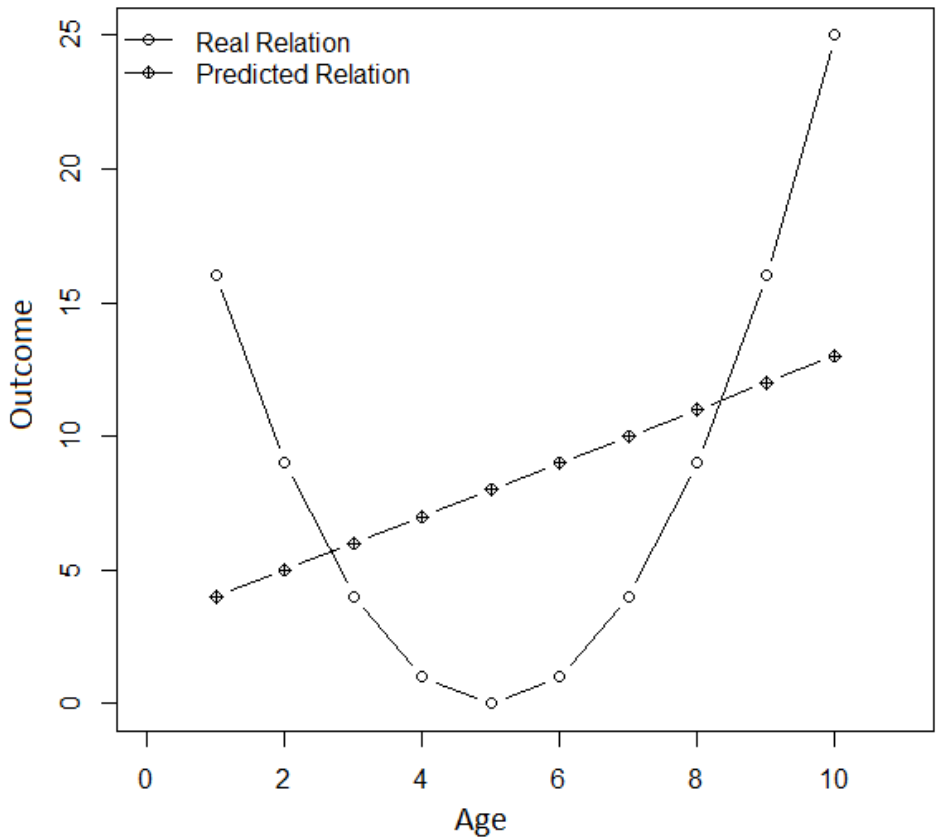
What if relationship between age and outcome in treated patients is not linear...



For simplicity, the outcome is continuous now

What if models are misspecified?

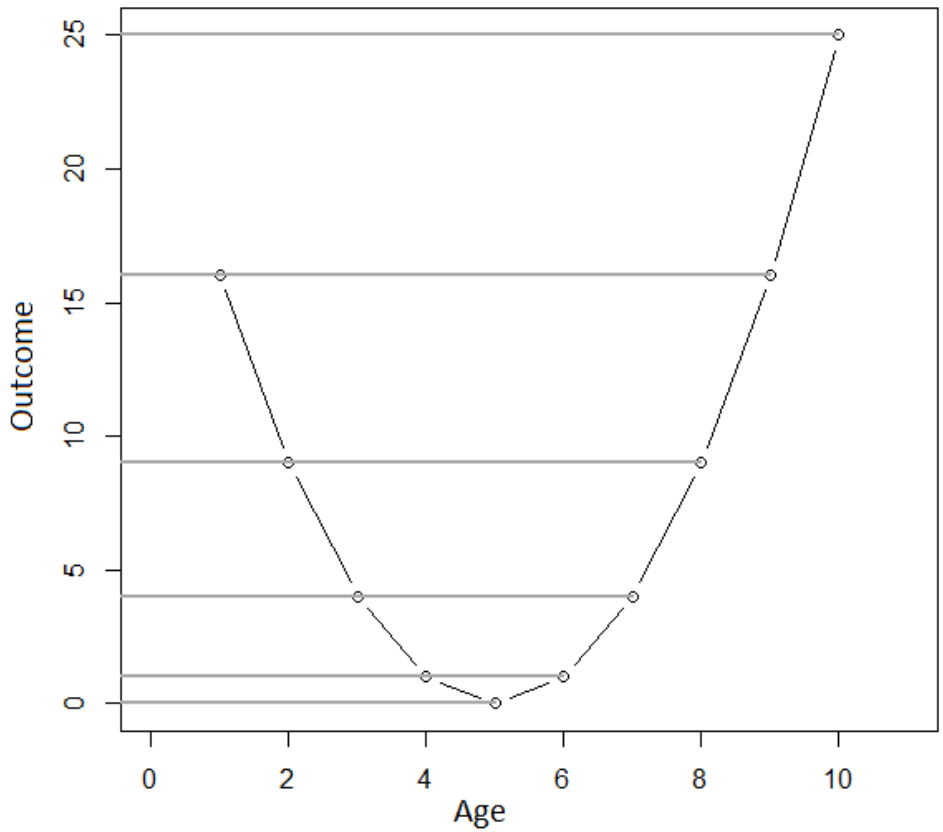
. . . , but we fit a misspecified model *outcome* $\sim$ *age*?



For simplicity, the outcome is continuous now

What if models are misspecified?

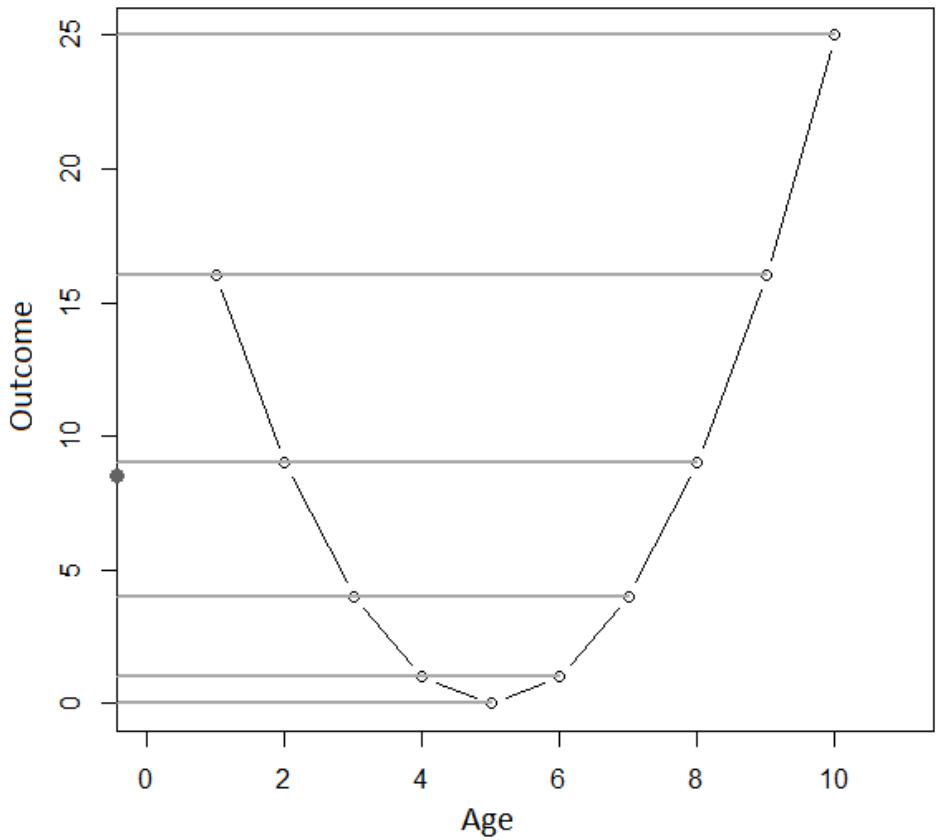
Projections of the observed outcomes on the y-axis,



For simplicity, the outcome is continuous now

What if models are misspecified?

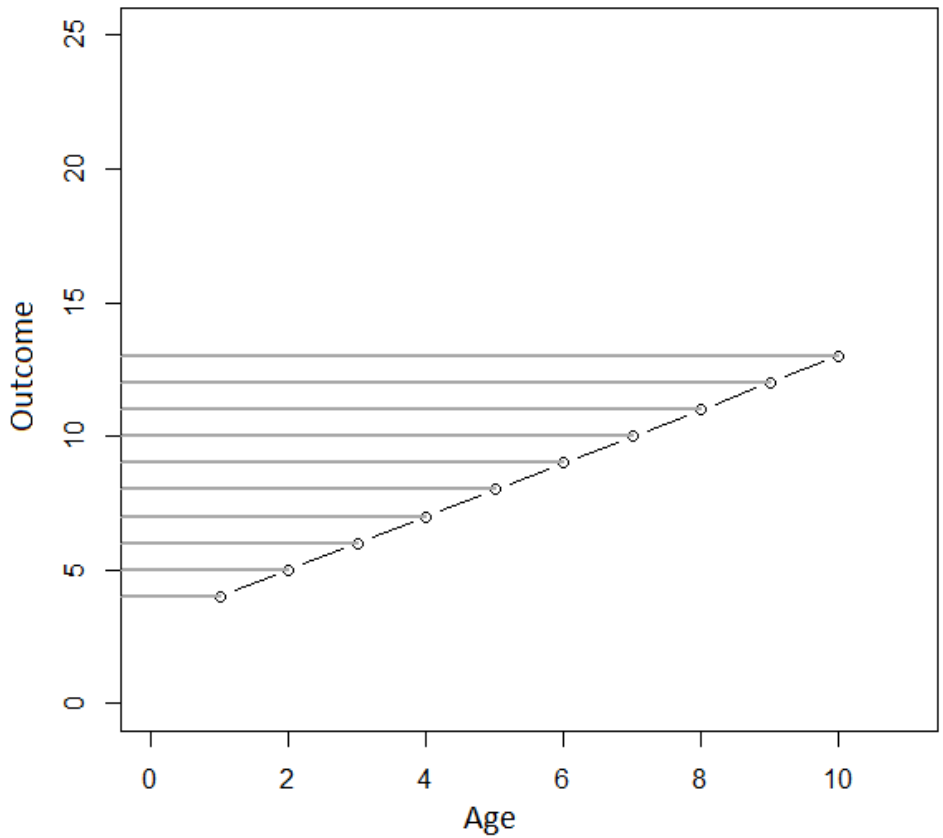
average to 8.5.



For simplicity, the outcome is continuous now

What if models are misspecified?

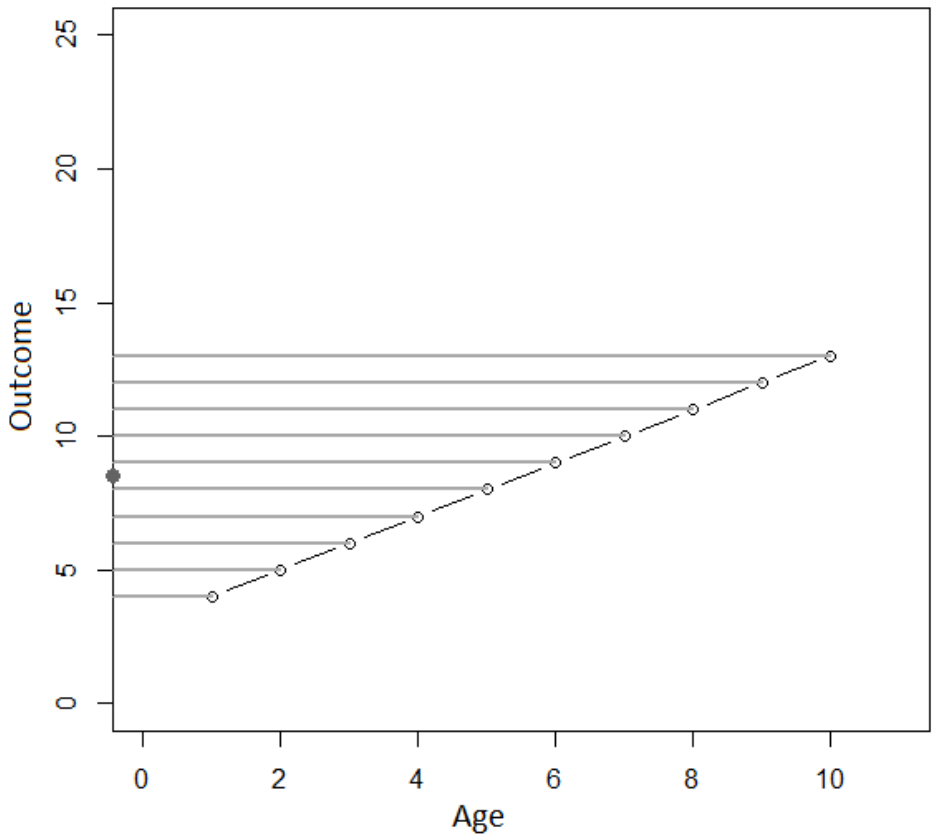
Projections of the predictions on the y-axis,



For simplicity, the outcome is continuous now

What if models are misspecified?

also average to 8.5.



For simplicity, the outcome is continuous now

What if models are misspecified?

- In treatment arm:
  mean of predictions (under treatment) = mean of observed outcomes,
  regardless of whether your model is correct or not.

What if models are misspecified?

- In treatment arm:
  mean of predictions (under treatment) = mean of observed outcomes,
  regardless of whether your model is correct or not.

- Under randomization, this **robustness against misspecification** also holds for mean of predictions (under treatment) for **all** patients

What if models are misspecified?

- In treatment arm:
  mean of predictions (under treatment) = mean of observed outcomes,
  regardless of whether your model is correct or not.

- Under randomization, this **robustness against misspecification** also holds for mean of predictions (under treatment) for **all** patients

  $\Rightarrow$ **Consistent estimator**, even when model is wrong.

What if models are misspecified?

- In treatment arm:
  mean of predictions (under treatment) = mean of observed outcomes,
  regardless of whether your model is correct or not.

- Under randomization, this **robustness against misspecification** also holds for mean of predictions (under treatment) for **all** patients

  $\Rightarrow$ **Consistent estimator**, even when model is wrong.

- Model misspecification may **reduce efficiency**.
  - Despite the precision loss, (almost) never outperformed by the standard analyses.

Inference

- Standard errors easy to calculate

1. **Robust standard errors** (Tsiatis et al., 2008; Rosenblum and Van Der Laan, 2009; Ye et al., 2023):
   - Similar to variance of sample mean

     sample variance of
     $2A(Y-\hat{P}^1)+\hat{P}^1-(2(1-A)(Y-\hat{P}^0)+\hat{P}^0)-(\hat{E}(Y^1)-\hat{E}(Y^0))$
     for a mean difference
   - **Takes into account uncertainty in imputations**

Inference

- Standard errors easy to calculate

1. **Robust standard errors** (Tsiatis et al., 2008; Rosenblum and Van Der Laan, 2009; Ye et al., 2023):
   - Similar to variance of sample mean

     sample variance of
     $2A(Y-\hat{P}^1)+\hat{P}^1-(2(1-A)(Y-\hat{P}^0)+\hat{P}^0)-(\hat{E}(Y^1)-\hat{E}(Y^0))$
     for a mean difference
   - **Takes into account uncertainty in imputations**

2. Non-parametric bootstrap

Inference

- Standard errors easy to calculate
  1. **Robust standard errors** (Tsiatis et al., 2008; Rosenblum and Van Der Laan, 2009; Ye et al., 2023):
     - Similar to variance of sample mean

       sample variance of
       $2A(Y-\hat{P}^1)+\hat{P}^1-(2(1-A)(Y-\hat{P}^0)+\hat{P}^0)-(\hat{E}(Y^1)-\hat{E}(Y^0))$
       for a mean difference
     - **Takes into account uncertainty in imputations**
  2. Non-parametric bootstrap
- and are **valid** even when the model is **misspecified**

Inference

- Standard errors easy to calculate
  1. **Robust standard errors** (Tsiatis et al., 2008; Rosenblum and Van Der Laan, 2009; Ye et al., 2023):
     - Similar to variance of sample mean

       sample variance of
       $2A(Y-\hat{P}^1)+\hat{P}^1-(2(1-A)(Y-\hat{P}^0)+\hat{P}^0)-(\hat{E}(Y^1)-\hat{E}(Y^0))$
       for a mean difference
     - **Takes into account uncertainty in imputations**
  2. Non-parametric bootstrap
- and are **valid** even when the model is **misspecified**
- or when **variable selection** is used (Avagyan and Vansteelandt, 2021).

Recommendations

- Important to use predictions based on **glm's with canonical link**.
  - Otherwise we need slightly **different approach** (AIPW, TMLE).

Recommendations

- Important to use predictions based on **glm's with canonical link**.
  - Otherwise we need slightly **different approach** (AIPW, TMLE).
- Use of baseline covariates raises concerns due to **missing data**
  - Easily addressed: **mean/mode imputation**.
  - Without inflating risk of bias.

Recommendations

- Important to use predictions based on **glm's with canonical link**.
  - Otherwise we need slightly **different approach** (AIPW, TMLE).
- Use of baseline covariates raises concerns due to **missing data**
  - Easily addressed: **mean/mode imputation**.
  - Without inflating risk of bias.
- I haven't covered all available methods
  - There are no other methods that have more power and have the same robustness.

Outline

Super-covariates

- Recently a new approach based on so-called **super-covariates** gained attention (PROCOVA):
  1. **Fitting** a prognostic model to predict outcomes under the control condition in a **historical dataset**.

Super-covariates

- Recently a new approach based on so-called **super-covariates** gained attention (PROCOVA):
  1. **Fitting** a prognostic model to predict outcomes under the control condition in a **historical dataset**.
  2. Use this to make a **prediction** for all patients in current trial under control (i.e., super-covariate).

Super-covariates

- Recently a new approach based on so-called **super-covariates** gained attention (PROCOVA):
  1. **Fitting** a prognostic model to predict outcomes under the control condition in a **historical dataset**.
  2. Use this to make a **prediction** for all patients in current trial under control (i.e., super-covariate).
  3. Estimating the treatment effect in the current study using a **linear model** while adjusting for the super-covariate only.

$$E(Y|A, X_{SC}) = \beta_0 + \beta_1 A + \beta_2 X_{SC}.$$

Concerns about PROCOVA

Some concerns

- **Limited to continuous outcomes**; no easy way to extend to ordinal, binary or time-to-event endpoints.

Concerns about PROCOVA

Some concerns

- **Limited to continuous outcomes**; no easy way to extend to ordinal, binary or time-to-event endpoints.
- Does **not allow for treatment effect heterogeneity** (i.e., treatment effect is same in all covariate strata).
  - The prognostic score will be **less predictive** of the treatment-arm outcomes (Schuler et al., 2022).

Concerns about PROCOVA

Some concerns

- **Limited to continuous outcomes**; no easy way to extend to ordinal, binary or time-to-event endpoints.
- Does **not allow for treatment effect heterogeneity** (i.e., treatment effect is same in all covariate strata).
  - The prognostic score will be **less predictive** of the treatment-arm outcomes (Schuler et al., 2022).
- Historical trial population will generally differ from that in the considered trial: **loss of power**.

Concerns about PROCOVA

Some concerns

- **Limited to continuous outcomes**; no easy way to extend to ordinal, binary or time-to-event endpoints.
- Does **not allow for treatment effect heterogeneity** (i.e., treatment effect is same in all covariate strata).
  - The prognostic score will be **less predictive** of the treatment-arm outcomes (Schuler et al., 2022).
- Historical trial population will generally differ from that in the considered trial: **loss of power**.
  - **Association between outcome and covariates may differ** between the historical and trial population.

Possibly safer option

- Also include (a few) **other baseline covariates** in addition to the super-covariate.
  - For example, including the **baseline measurement of the primary outcome** is usually recommended.

Possibly safer option

- Also include (a few) **other baseline covariates** in addition to the super-covariate.
  - For example, including the **baseline measurement of the primary outcome** is usually recommended.

Licensed Published by De Gruyter December 22, 2021

Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score

Alejandro Schuler, David Walsh, Diana Hall, Jon Walsh, Charles Fisher, for the Critical Path for Alzheimer's Disease, the Alzheimer's Disease Neuroimaging Initiative and the Alzheimer's Disease Cooperative Study

From the journal The International Journal of Biostatistics
https://doi.org/10.1515/ijb-2021-0072

Outline

How to determine covariates for adjustment?

- To gain power, it's important that the variables are **associated with the outcome** (i.e., prognostic).

How to determine covariates for adjustment?

- To gain power, it's important that the variables are **associated with the outcome** (i.e., prognostic).
- Recommended by EMA and FDA: **Pre-specify baseline variables and model**.
  - This is a **difficult**, if not impossible, task!

How to determine covariates for adjustment?

- To gain power, it's important that the variables are **associated with the outcome** (i.e., prognostic).
- Recommended by EMA and FDA: **Pre-specify baseline variables and model**.
  - This is a **difficult**, if not impossible, task!
- Another possibility is to use **data-adaptive methods**, which includes stepwise regression, Lasso, machine learning, ...

How to determine covariates for adjustment?

- To gain power, it's important that the variables are **associated with the outcome** (i.e., prognostic).
- Recommended by EMA and FDA: **Pre-specify baseline variables and model**.
  - This is a **difficult**, if not impossible, task!
- Another possibility is to use **data-adaptive methods**, which includes stepwise regression, Lasso, machine learning, ...
  - **No need for historical data at all**,
  - but one can make use of historical data when available.
    - E.g., including a **super-covariate** as one of the potential variables to select from can lead to **finite sample gains**.

How to determine covariates for adjustment?

- To gain power, it's important that the variables are **associated with the outcome** (i.e., prognostic).
- Recommended by EMA and FDA: **Pre-specify baseline variables and model**.
  - This is a **difficult**, if not impossible, task!
- Another possibility is to use **data-adaptive methods**, which includes stepwise regression, Lasso, machine learning, ...
  - **No need for historical data at all**,
  - but one can make use of historical data when available.
    - E.g., including a **super-covariate** as one of the potential variables to select from can lead to **finite sample gains**.

Important: Pre-specify variable / model selection algorithm as well as a list of candidate covariates!

Naïve use of machine learning is problematic

- Unfortunately, **naïve use of machine learning induces bias**.

Naïve use of machine learning is problematic

- Unfortunately, **naïve use of machine learning induces bias**.
- There is **regularization bias** because machine learning has been **optimized for prediction**,
  not for the evaluation of treatment effects.

Naïve use of machine learning is problematic

- Unfortunately, **naïve use of machine learning induces bias**.
- There is **regularization bias** because machine learning has been **optimized for prediction**,
  not for the evaluation of treatment effects.
- Obtaining correct standard errors is challenging, because the **uncertainty in machine learning predictions is unknown**.

Naïve use of machine learning is problematic

- Unfortunately, **naïve use of machine learning induces bias**.
- There is **regularization bias** because machine learning has been **optimized for prediction**,
  not for the evaluation of treatment effects.
- Obtaining correct standard errors is challenging, because the **uncertainty in machine learning predictions is unknown**.
- With flexible machine learning, there may be **overfitting bias** due to training the prediction model and evaluating the treatment effect on the same data.

How to accomodate these biases?

To **eliminate bias** from the estimated mean outcomes, we must:

How to accomodate these biases?

To **eliminate bias** from the estimated mean outcomes, we must:

Step 1: **Model fitting/Training**
Train a prediction model for outcome Y given treatment allocation A and baseline covariates X.

Step 2: **Predicting**
Use this model to predict outcome **under treatment and control** for **all** patients.

How to accomodate these biases?

To **eliminate bias** from the estimated mean outcomes, we must:

Step 1: **Model fitting/Training**
Train a prediction model for outcome Y given treatment allocation A and baseline covariates X.

Step 2: **Predicting**
Use this model to predict outcome **under treatment and control** for **all** patients.

Step 3: **Targeting**
Shift predictions by a constant so that the average prediction in the treated equals the average outcome in the treated (similar for control).

How to accomodate these biases?

To **eliminate bias** from the estimated mean outcomes, we must:

Step 1: **Model fitting/Training**
Train a prediction model for outcome Y given treatment allocation A and baseline covariates X.

Step 2: **Predicting**
Use this model to predict outcome **under treatment and control** for **all** patients.

Step 3: **Targeting**
Shift predictions by a constant so that the average prediction in the treated equals the average outcome in the treated (similar for control).

Step 4: **Averaging**
Take the average of threse predictions, and calculate treatment effect of interest.

How to accomodate these biases?

- **Overfitting bias** can be eliminated by **training the model on one subsample and evaluating the treatment effect on another subsample**.

  (this can be done **without precision loss**)

How to accomodate these biases?

- **Overfitting bias** can be eliminated by **training the model on one subsample and evaluating the treatment effect on another subsample**.
  (this can be done **without precision loss**)

- A by-product of targeting is that **standard errors are simple to calculate**
  because they are made immune to the uncertainty in the machine learning predictions.

Reflections: targeted learning

- This strategy is called **targeted learning or Targeted Maximum Likelihood Estimation** (TMLE; Van Der Laan and Rubin (2006)).

Reflections: targeted learning

- This strategy is called **targeted learning or Targeted Maximum Likelihood Estimation** (TMLE; Van Der Laan and Rubin (2006)).
- It updates initial predictions and targets them towards the **estimand of interest**.

Reflections: targeted learning

- This strategy is called **targeted learning or Targeted Maximum Likelihood Estimation** (TMLE; Van Der Laan and Rubin (2006)).
- It updates initial predictions and targets them towards the **estimand of interest**.
- It brings **data analysis back to its essence**:
  translating a scientific question into an **estimands**, doing **sanity checks**, ...
  with **automated** model building strategies in background.

Reflections: targeted learning

- This strategy is called **targeted learning or Targeted Maximum Likelihood Estimation** (TMLE; Van Der Laan and Rubin (2006)).
- It updates initial predictions and targets them towards the **estimand of interest**.
- It brings **data analysis back to its essence**:
  translating a scientific question into an **estimands**, doing **sanity checks**, ...
  with **automated** model building strategies in background.
- It has the potential to deliver an **additional power increase** over standardization,
- but arguably the main benefit is that it **enables pre-specification of the analysis**.

Thank you for your attention!

Interested? ▸ Pre-print and ▸ Tutorials

E-mail: kelly.vanlancker@ugent.be

website: KellyVanLancker.com

The opinions in this presentation are of the authors and do not necessarily represent those of anyone else.

References I

Avagyan, V. and S. Vansteelandt (2021). High-dimensional inference for the average treatment effect under model misspecification using penalized bias-reduced double-robust estimation. *Biostatistics & Epidemiology*, 1–18.

Benkeser, D., I. Díaz, A. Luedtke, J. Segal, D. Scharfstein, and M. Rosenblum (2020). Improving precision and power in randomized trials for covid-19 treatments using covariate adjustment, for binary, ordinal, and time-to-event outcomes. *Biometrics*.

Daniel, R., J. Zhang, and D. Farewell (2021). Making apples from oranges: Comparing noncollapsible effect estimators and their standard errors after adjustment for different covariate sets. *Biometrical Journal 63*(3), 528–557.

EMA (2015). Guideline on adjustment for baseline covariates in clinical trials. Last checked: 2022-05-30.

References II

FDA (2023). Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products. Guidance for Industry. https://www.fda.gov/media/148910/download. Last checked: 2023-09-05.

Jiang, F., L. Tian, H. Fu, T. Hasegawa, and L. J. Wei (2018). Robust alternatives to ANCOVA for estimating the treatment effect via a randomized comparative study. *Journal of the American Statistical Association 0*, 1–37.

Koch, G. G., C. M. Tangen, J.-W. Jung, and I. A. Amara (1998). Issues for covariance analysis of dichotomous and ordered categorical data from randomized clinical trials and non-parametric strategies for addressing them. *Stat. Med. 17*(15-16), 1863–1892.

References III

Moore, K. and M. J. van der Laan (2009a). Covariate adjustment in randomized trials with binary outcomes: Targeted maximum likelihood estimation. *Stat. Med. 28*(1), 39–64.

Moore, K. L. and M. J. van der Laan (2009b). Increasing power in randomized trials with right censored outcomes through covariate adjustment. *Journal of Biopharmaceutical Statistics 19*(6), 1099–1131. PMID: 20183467.

Rosenblum, M. and M. J. Van Der Laan (2009). Using regression models to analyze randomized trials: Asymptotically valid hypothesis tests despite incorrectly specified models. *Biometrics 65*(3), 937–945.

References IV

Rubin, D. and M. van der Laan (2008). Covariate adjustment for the intention-to-treat parameter with empirical efficiency maximization. *U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 229*, https://biostats.bepress.com/ucbbiostat/paper229.

Schuler, A., D. Walsh, D. Hall, J. Walsh, C. Fisher, A. D. N. Initiative, et al. (2022). Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score. *The International Journal of Biostatistics 18*(2), 329–356.

Tsiatis, A. A., M. Davidian, M. Zhang, and X. Lu (2008). Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. *Statistics in medicine 27*(23), 4658–4677.

References V

Van Der Laan, M. J. and D. Rubin (2006). Targeted maximum likelihood learning. *The international journal of biostatistics 2*(1).

Yang, L. and A. Tsiatis (2001). Efficiency study of estimators for a treatment effect in a pretest-posttest trial. *The American Statistician 55*(4), 314–321.

Ye, T., M. Bannick, Y. Yi, and J. Shao (2023). Robust variance estimation for covariate-adjusted unconditional treatment effect in randomized clinical trials with binary outcomes. *Statistical Theory and Related Fields*, 1–5.

Ye, T., J. Shao, Y. Yi, and Q. Zhao (2022). Toward better practice of covariate adjustment in analyzing randomized clinical trials. *Journal of the American Statistical Association*, 1–13.

References VI

Zhang, M. (2015, Jan). Robust methods to improve efficiency and reduce bias in estimating survival curves in randomized clinical trials. *Lifetime Data Analysis 21*(1), 119–137.