Introduction to Causal Inference and Causal Data Science

Introduction to Target Trials and Target Trial Emulation

Bas Penning de Vries

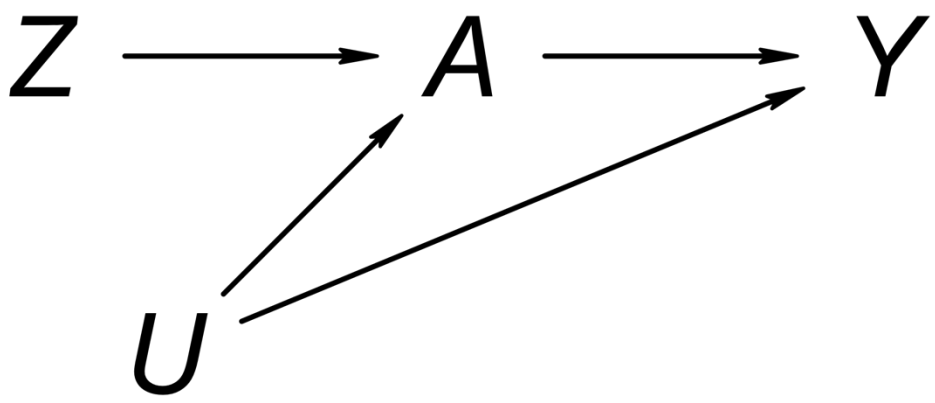
Question 1

Which of the following options best describes what a potential outcome is according to the counterfactual (or potential) outcomes framework?

A. A possible value of the outcome variable
B. The outcome of an individual that would be observed had treatment been set (by intervention) to a certain value
C. The best outcome an individual can achieve
D. A possible outcome of a study

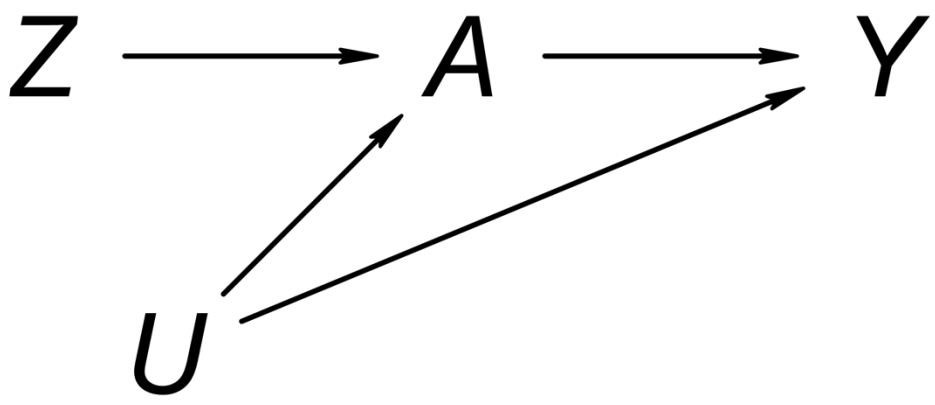
Question 2

True or false? The **backdoor criterion** is fulfilled by a set of variables if it closes *at least one* backdoor path from treatment to outcome and none of the variables is a descendant of the first variable on the path.



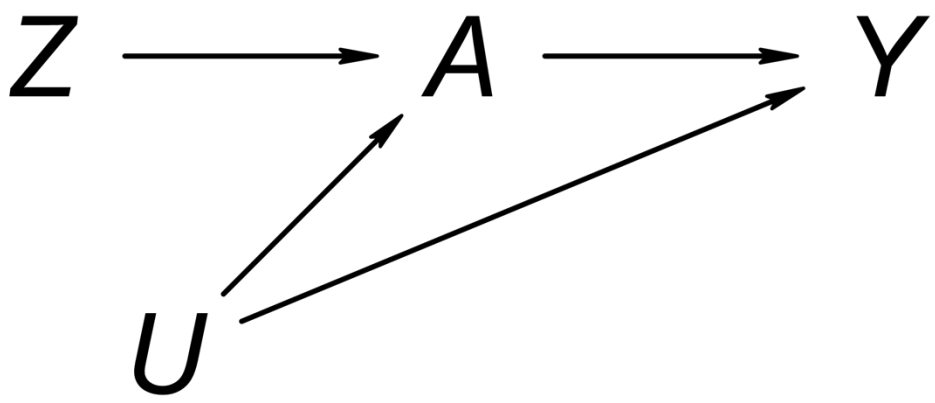
Question 3

True or false? The **backdoor criterion** is satisified for the treatment/exposure *Z* and outcome *Y*.



Question 4

True or false? If the **backdoor criterion** is satisfied for *Z* and *Y*, then the exposure groups (defined by *Z*) are **exchangeable** with respect to the outcome *Y*



Question 4

True or false? The **backdoor criterion** is satisified for the treatment/exposure *A* and outcome *Y*.

Question 5

True or false? Recent methodological developments allow epidemiologists to falsify the presence of confounding using a statistical test that does not rely on causal assumptions.

Learning objectives

By the end of today/week, you'll be able to

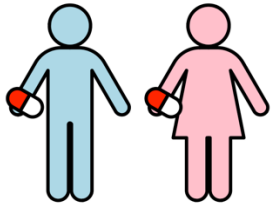
- Describe what is meant by a target trial and **target trial emulation**
- Identify **key components** of target trial emulation, including the determination of the start of follow-up (**time zero**)
- Recognise a **taxonomy of estimands** relevant to target trial emulation and distinguish between common targets such as intention-to-treat and per-protocol effects *(Friday!)*
- Describe the **relevance** of target trial emulation in causal inference from observational data
- Recognise **common deviations** from a target trial in observational studies
- Explain the basics of commonly used **methods** to address these deviations

"**Causal inference from observational data** can be viewed as an attempt to **emulate a hypothetical randomised trial**"

Hernán and Robins, 2020, *Causal Inference: What if*

Why trials?

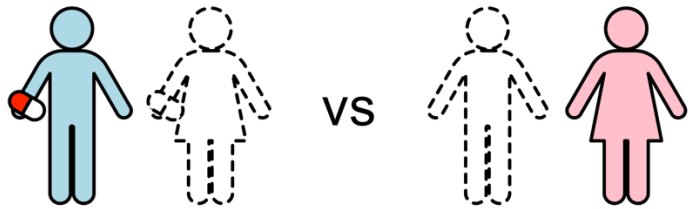
Causal inference is about speculating what would happen if …



A **causal effect** is a contrast between the answers to what-if questions

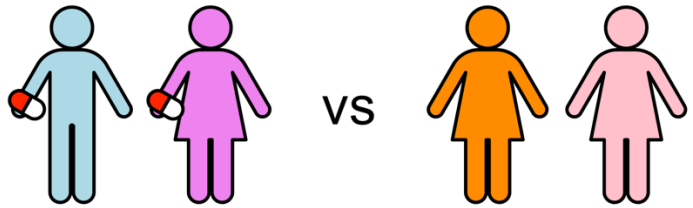
Fundamental obstacle

Impossible to observe the consequences of ≥ 2 mutually exclusive actions (interventions, treatments, etc.)

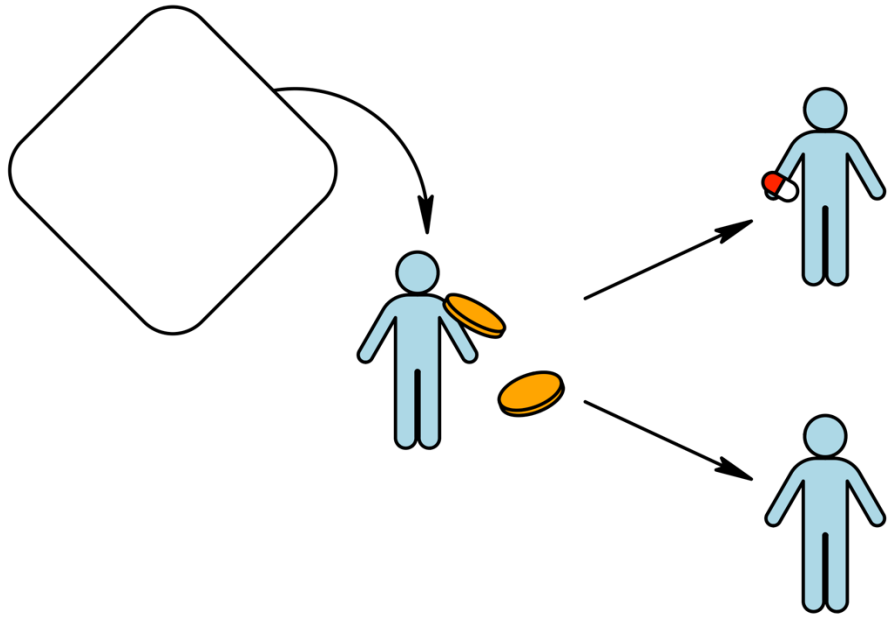


Solution?

Instead of comparing the same individual between different counterfactual ("what-if") situations, …

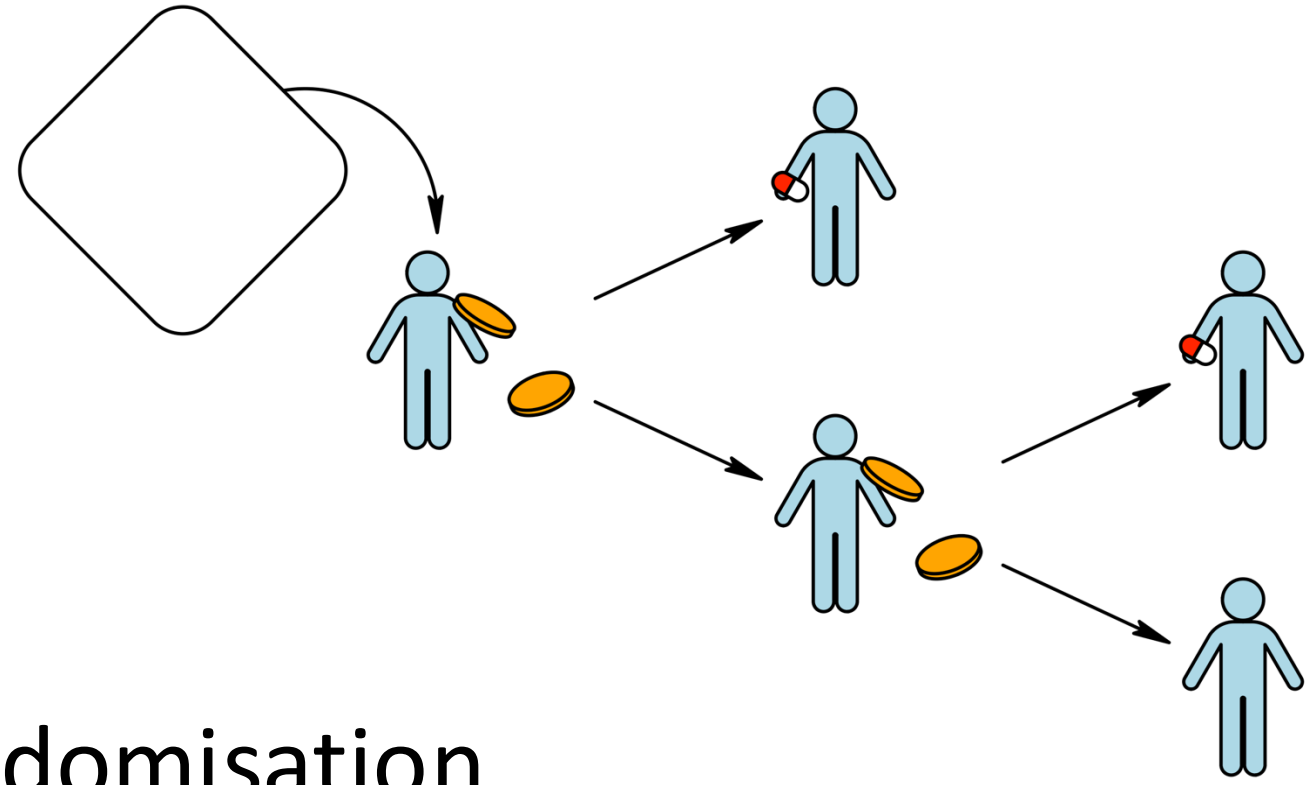


… compare *different individuals* who are *actually treated differently*



Randomisation

Instead of subjecting exact copies of the same individual to different levels of treatment, with randomisation you get *differently treated individuals* whose characteristics – other than treatment and its consequences – are *identical in distribution*



Randomisation

- Powerful (conceptual) tool
- Can accommodate all sorts of interventions
  (single or multiple time-point interventions, static or dynamic)

Why not do trials?

- Expensive
- Unethical
- Impractical
- Untimely
- …

Target trial

A *hypothetical trial* that – if implemented – would readily allow us to answer our what-if question

- To help *communicate causal estimand* (because identification is “straightforward”)
- To *facilitate appraisal* of actual research designs (and *avoid methodological problems* with your study)

Target trial emulation

Explicit attempt to address deviations from a target trial, given the (observational) study data at hand

Step 1. *Specify* target trial
Step 2. *Emulate* it!

Step 1: specify target trial

What do you need to know to implement (and replicate) it?

	Target trial
Eligibility criteria	...
Treatment strategies	...
Outcome	...
Time zero and follow-up	...
Causal contrasts	...
Data analysis	...

Example: identification from target trials should be straightforward

$$\text{ATE} = E[Y^{A=1}] - E[Y^{A=0}]$$ (the causal estimand)

$$= E[Y^{A=1} \mid A = 1] - E[Y^{A=0} \mid A = 0]$$

(randomisation ⇒ *exchangeability*, i.e., A indep. of $Y^{A=a}$ for $a = 0,1$; conditionals are defined only under *positivity*, which too is controlled by design)

$$= E[Y \mid A = 1] - E[Y \mid A = 0]$$

(*consistency*, i.e., $Y^{A=a} = Y$ if $a = A$)

Target trial emulation vs “silly” questions

Formulating a target trial helps to communicate the causal estimand and helps to avoid asking vague or “silly” questions (about ill-defined or irrelevant interventions)

- Eligibility defined by post-baseline events
- Causal effect of (a reduction/increase in) BMI?
- “Does water kill?” (Hernán, Ann Epidemiol., 2016;26(10):674–680)
- Unclear treatment strategies (e.g., stopping rules, dosage, etc.)

Treatment-variation (ir)relevance and well-definedness

- There may be many variations on an intervention and their impact on the outcome of interest need not be the same
- Interventions are sufficiently **well-defined** if there is no ambiguity about the variation or all possible variations equally affect the outcome variables of interest (i.e., there is **treatment-variation irrelevance**)
- Prerequisite of consistency

Having to write a trial protocol forces you to be explicit and precise!

Step 2: emulate target trial

Compare and address departures from target trial (analytically)

	Target trial	Emulation study
Eligibility criteria	...	...
Treatment strategies	...	...
Outcome	...	...
Time zero and follow-up	...	...
Causal contrasts	...	...
Data analysis	...	...

Example: do statins prevent cancer?

nature medicine

ANALYSIS

https://doi.org/10.1038/s41591-019-0597-x

Avoidable flaws in observational analyses: an application to statins and cancer

Barbra A. Dickerman[1*], Xabier García-Albéniz[1,2], Roger W. Logan[1], Spiros Denaxas[3,4,5] and Miguel A. Hernán[1,6,7]

Table 2—*Effect of Statin Therapy and Its Duration on the Odds for Lung Cancer*

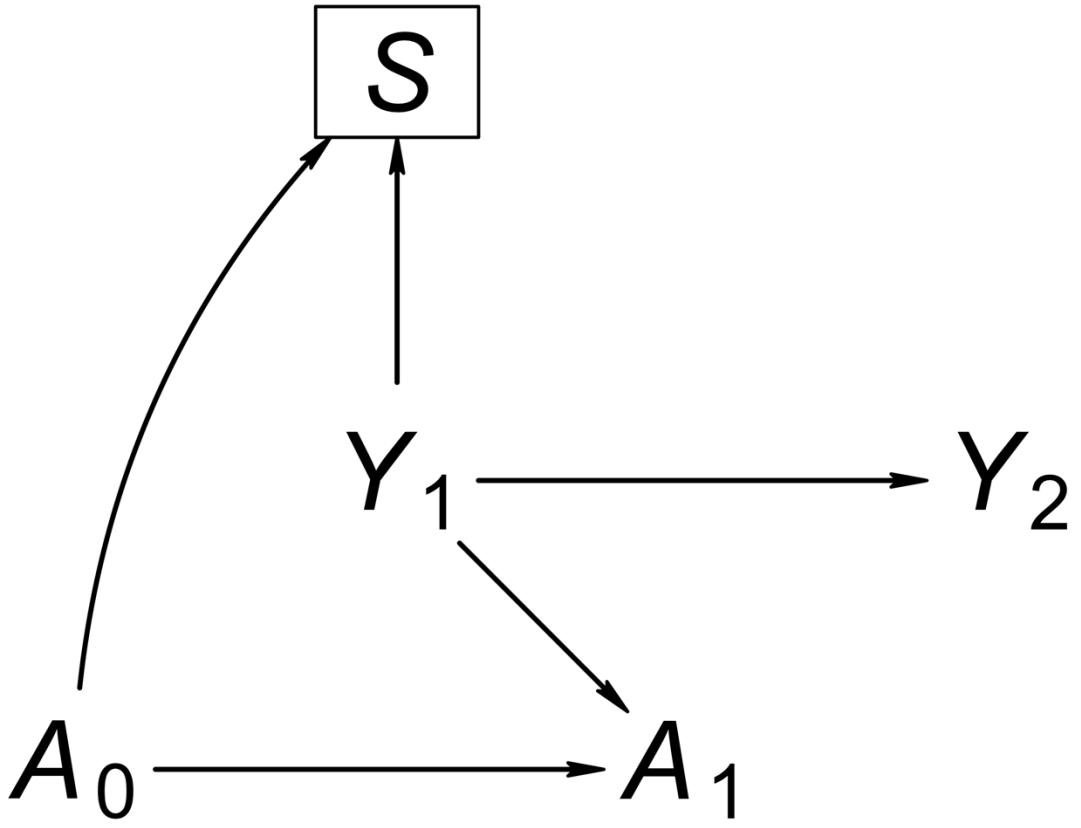
Variables	Cases, No.	Controls, No.	Crude OR	95% CI	Adjusted OR*	95% CI*	p Value for Adjusted OR
Overall	7,280	476,453					
Not exposed to statins	5,286	314,785					
Exposed to statins before lung cancer diagnosis (statin use > 0 yr)	1,994	161,668	0.73	0.70–0.77	0.55	0.52–0.59	< 0.01
Duration of statin use, yr							
0–0.5	446	10,259	2.59	2.34–2.86	2.32	2.05–2.63	< 0.01
0.5–1.0	214	15,564	0.82	0.71–0.94	0.75	0.63–0.89	< 0.01
1.0–2.0	416	30,590	0.81	0.73–0.90	0.70	0.61–0.79	< 0.01
2.0–4.0	649	55,516	0.70	0.64–0.76	0.49	0.44–0.55	< 0.01
> 4.0	269	49,739	0.32	0.28–0.36	0.23	0.20–0.26	< 0.01

*Adjusted for effects of age, race, sex, BMI, smoking, alcohol use, and diabetes.

Khurana et al. *Chest* 2007;131(5):1282-8

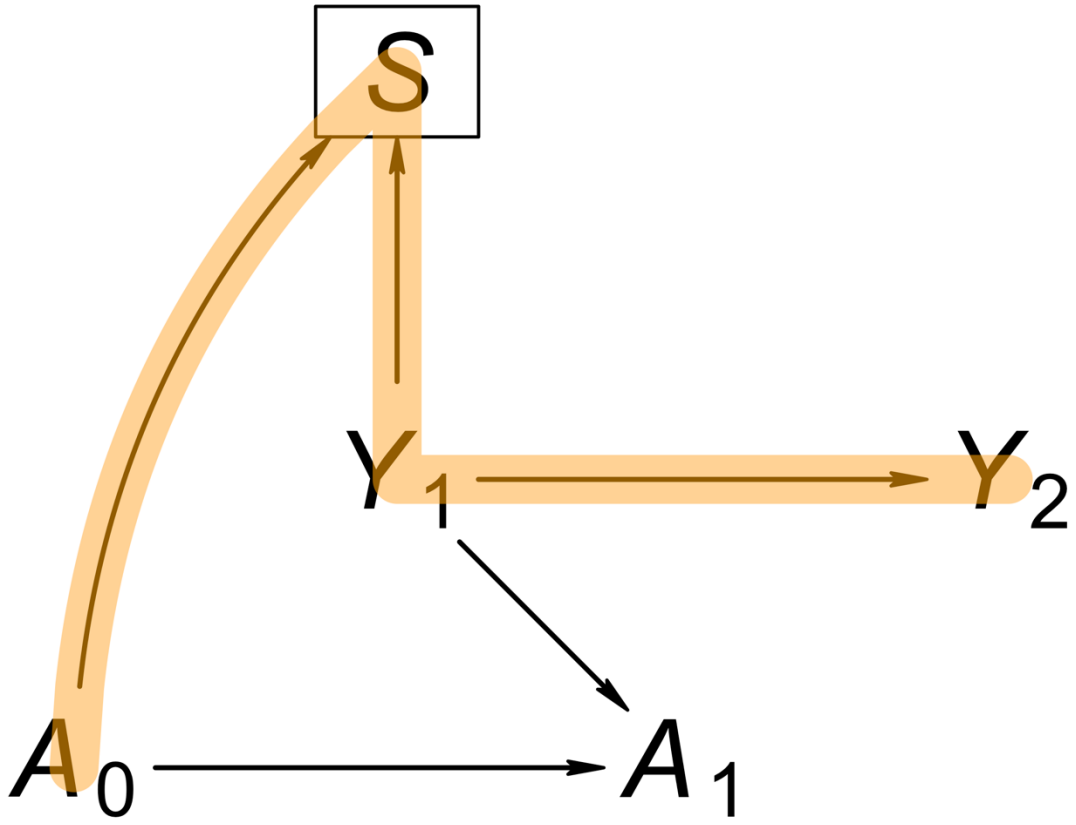
Immortal time bias

- **Immortal time** – a period of follow-up during which death or the study outcome cannot occur *by design*
- Arises from using postbaseline information to define (1) inclusion/eligibility/selection (selection) or (2) the exposure/contrast – *against trial principles!*
- May result in bias depending on *how it is handled*!
- Key to depicting this in a *DAG* is to include time-specific instances of variables



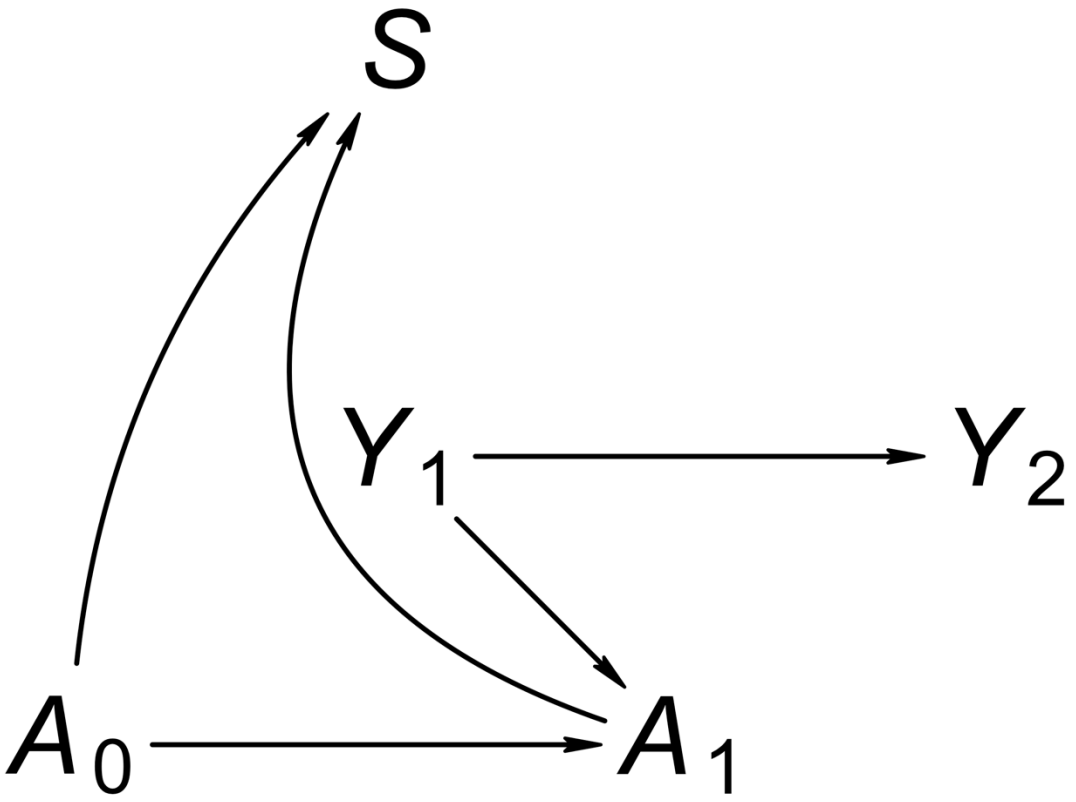
Immortal time bias by selection

Are groups defined by A_0 exchangeable relative to outcome Y_2 conditional on S=1? *Hint: use the backdoor criterion!*



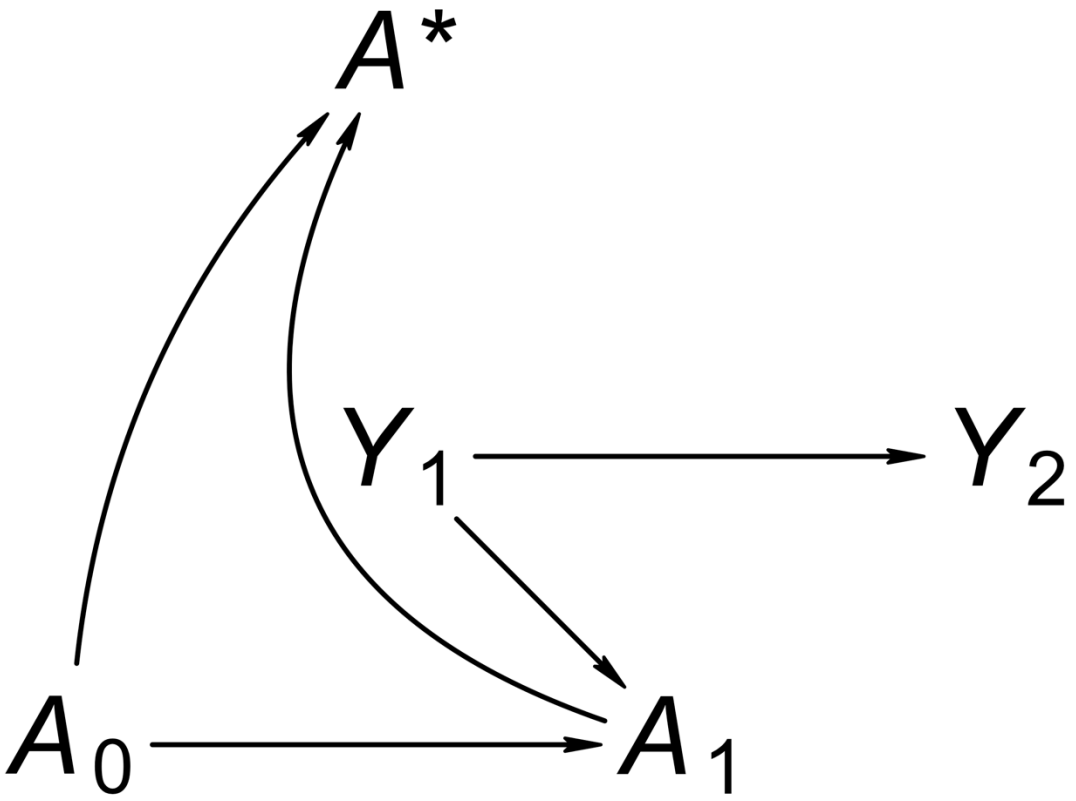
Immortal time bias by selection

- S is a descendant of A_0 – violation of backdoor criterion!
- A_0 and Y_2 are marginally independent (d-separated) but not necessarily conditional on S=1!



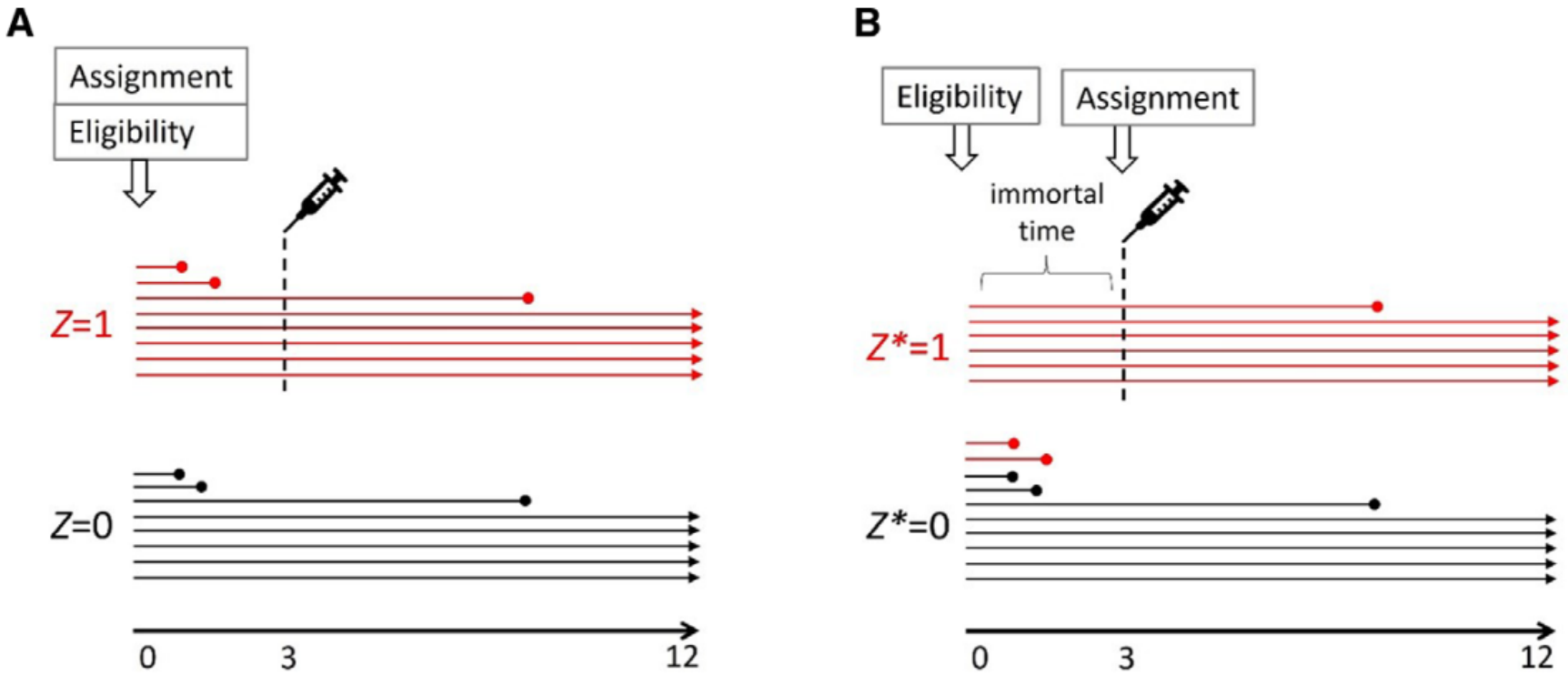
Immortal time bias by selection (2)

- S is a descendant of A_0 – violation of backdoor criterion!
- A_0 and Y_2 are marginally independent (d-separated) but not necessarily conditional on S=1!



Immortal time bias by misclassification

- Eg: compare surgery with wait-time vs no surgery
- Immortal time bias can arise when we include everyone but make the wrong contrast (surgery actually received vs not)



Hernán, et al., *Epidemiology* 2025;36:107-114

Including **prevalent users** to study effects **incident use**

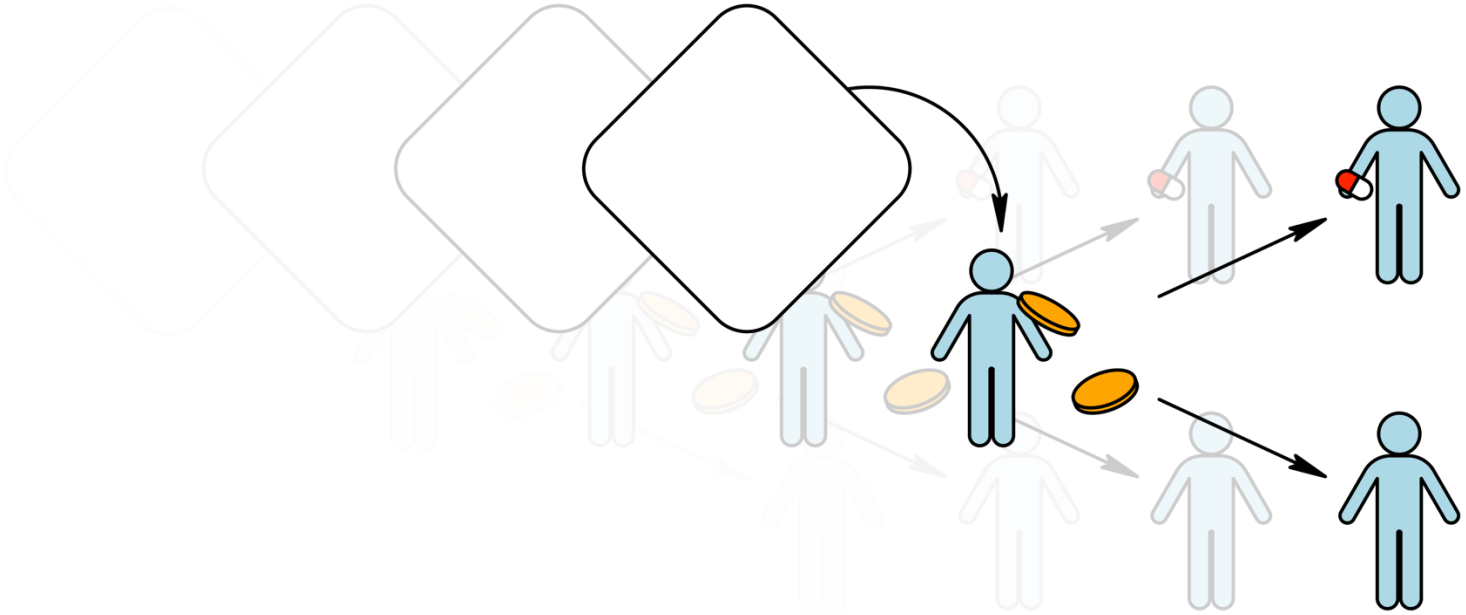
- Would you consider *initiating* a treatment regime now (at baseline) for a patient who is already on treatment (*prevalent user*)?
- Prevalent users are not part of the target population!
- Inclusion might result in (selection) bias!

"**Misalignment** of **eligibility**, **treatment assignment**, and the **start of follow-up** can result in time-related bias such as immortal time and selection of prevalent users"

Matthews et al., *JCE* 2023;164:112-115

Too few incident users at any given time?

- In studies on the effect of statin use and cancer incidence, there are *few* incident users at specified time baseline t_0 (or in short period starting at t_0)
- Consider trials that are identical except for their baseline time
- To gain efficiency, could emulate *multiple* such trials and analyse simultaneously (possibly according to flexible modelling assumptions to reflect heterogeneity across trials)
- NB: because individuals can be eligible for randomisation in multiple trials, need to respect *clustering* in estimating standard errors and constructing confidence intervals!



Sequential trial emulation

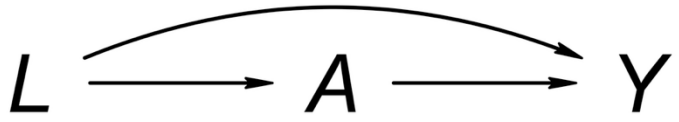
Statin-cancer example revisited

- Dickerman et al. (Nat Med, 2019;25(10):1601-1606):
  - When applying trial principles to analyse observational data (emulating a trial), they found effect estimates close to null
  - When reanalysing the observational data using the same approach as in earlier analyses, they found effect estimates similar to those found in earlier observational studies
- Discrepancies between trials and observational studies are often attributable largely to sources of bias other than residual confounding!

Addressing departures from randomisation

Any method may be used for confounding control

- Restriction
- Regression adjustment
- (Propensity score) matching
- G-computation
- Inverse probability weighting (IPW)
- …



Choice should be influenced in part by estimand

- Propensity score matching and g-methods (IPW and g-computation) target quantities typically estimated in trials

Propensity score methods

A collection of methods based on the propensity score (PS):

- PS **stratification**
- **Regression** on the PS
- PS **matching**
- Inverse probability **weighting** (IPW)

Propensity score, ps(*L*): **conditional probability of** (propensity for) **treatment** (or exposure) *A* given a set *L* of variables:

$$\text{ps}(L) = \Pr(A = 1 \mid L)$$

Donald Rubin

The PS as a summary and balancing score

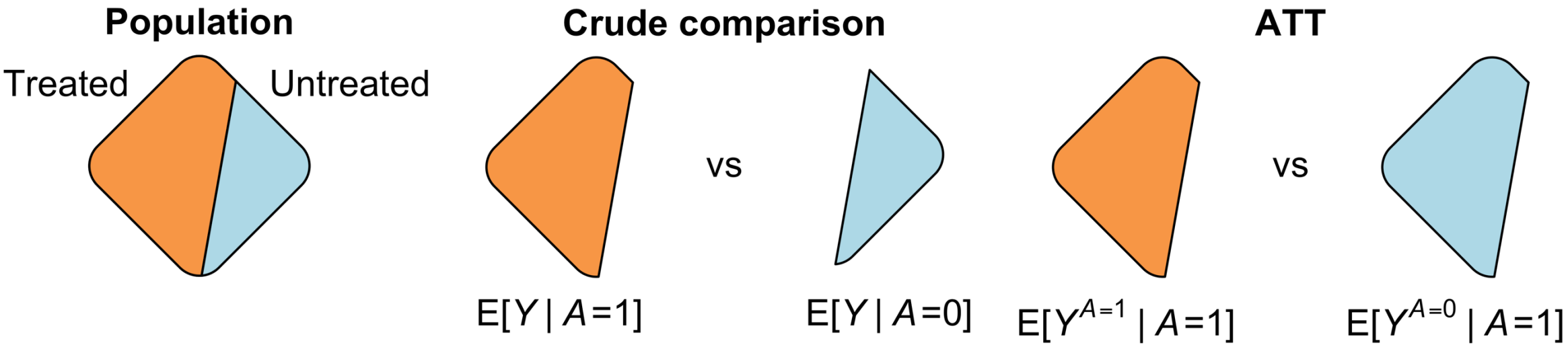
Rosenbaum and Rubin (1983) demonstrated **balancing** property:

- Conditional on the ps(L), the distribution of L is the same among the treated ($A = 1$) as it is among the untreated ($A = 0$)
- More importantly, if $\Pr(A = a \mid L) > 0$ (**positivity given** L),

$Y^{A=a}$ independent of A (**exchangeability**) **given** L

$\Downarrow$

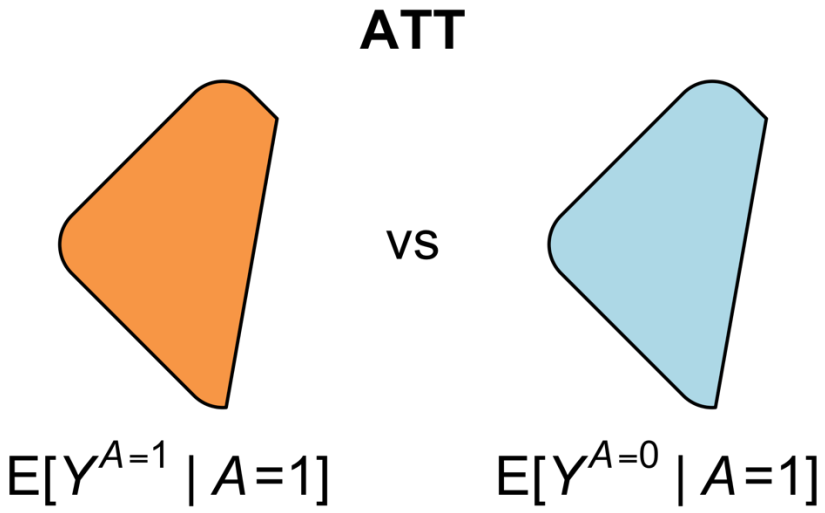
$Y^{A=a}$ independent of A (**exchangeability**) **given** ps(L)



PS matching

- Typical estimand is *average treatment effect* among *treated* (ATT), but exact estimand depends on implementation/variation

PS matching



- E[$Y^{A=1}$ | A=1] = E[Y | A=1] (consistency),
  but what about the other half of the contrast?
- If we could do **exact matching**,
  i.e., find Y_{match} such that

$$Y_{\text{match}} | A=1, \text{ps}(L) \sim Y | A=0, \text{ps}(L),$$

it turns out that, under conditional exchangeability, positivity and consistency,

$$\text{E}[Y^{A=0} \mid A=1] = \text{E}[Y_{\text{match}} \mid A = 1]$$

PS matching

Problem with finite samples: exact matching often (nearly always) impossible for all treated individuals

Solution: use approximate matching

- Many algorithms to choose from (e.g., greedy 1:1 nearest neightbour matching with distance defined as absolute difference between logit ps(*L*); with calliper)

But

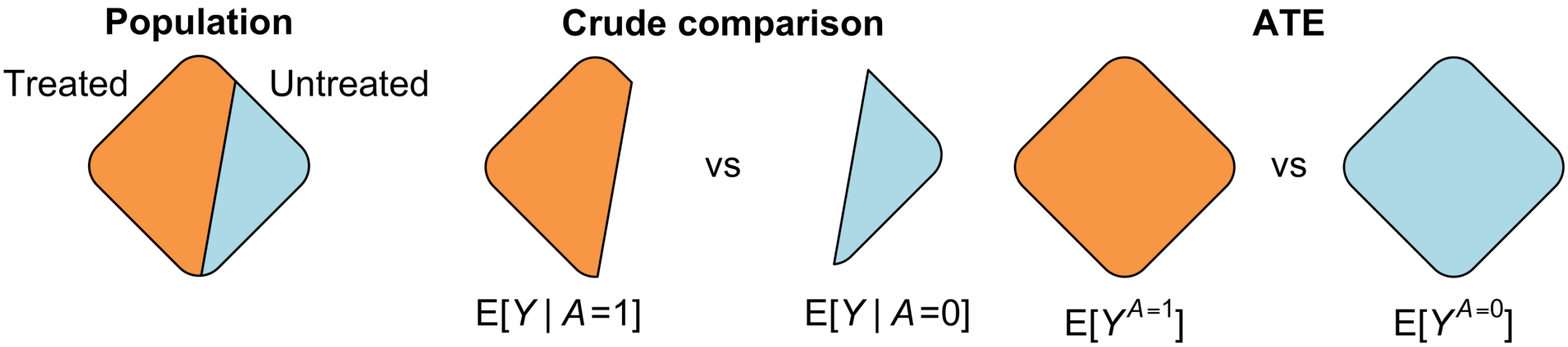
- Allowing matching of *distant* subjects may result in bias (**residual confounding**)
- Excluding subjects for whom no close-enough match can be found may result in **poor precision** (or **selection bias**)

PS matching in practice

- In practice, the true propensity score ps(L) is **unknown**
- Two-step procedure: (1) **propensity score estimation** and (2) treatment-outcome **effect estimation** based on the estimated instead of the true propensity score
- First step involves **modelling the distribution of treatment** A **given** L
  - For example, if L is a length-p vector, use MLE of the parameters of a logistic regression model

$$\text{logit Pr}(A = 1 \mid L) = \alpha_0 + \alpha_1 L_1 + \alpha_2 L_2 + \ldots + \alpha_p L_p$$

  - There is potential for **model misspecification** (and in turn bias)



IPW for binary treatments

- Typical estimand is *average treatment effect* (ATE)
- Key idea: reweight the treated and untreated subpopulations so that they look (in some respects) like the entire population

Goal of IPW for binary treatments

Create a **pseudopopulation** by weighting the original population such that

distribution of $(L, Y^{A=a})$ in the (original) population
=
distribution of $(L, Y^{A=a})$ in the pseudopopulation
among both the treated and among the untreated

This means that, in the pseudopopulation:

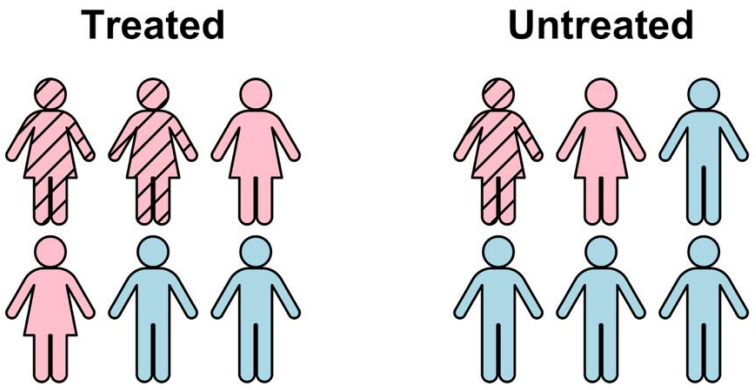
- L and $Y^{A=a}$ independent of A (exchangeability, as in RCT)
- Causal effect = crude association, under consistency

IPW for binary treatments: how?

The pseudopopulation with this property can be made under exchangeability and positivity given *L*, by weighting

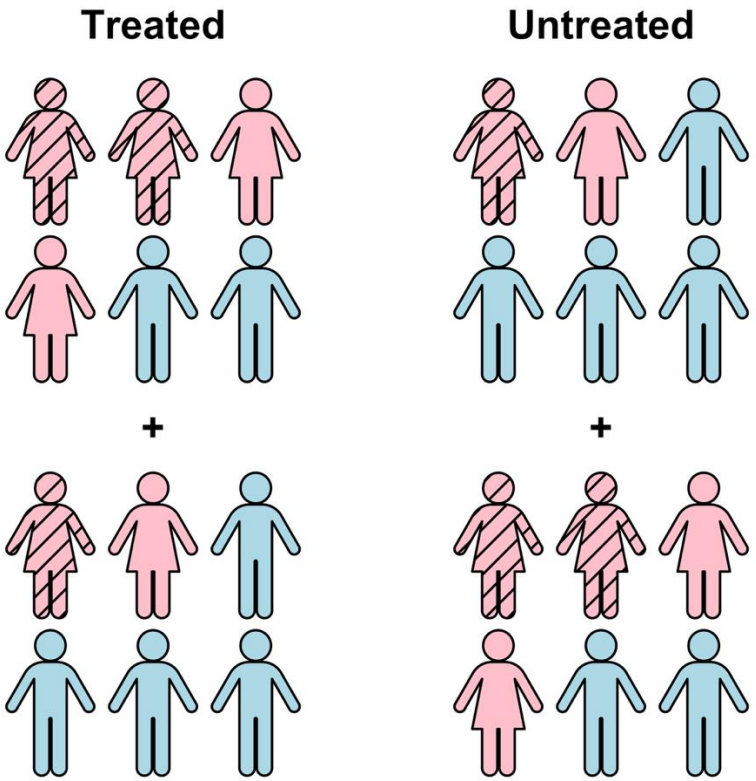
- treated subjects with $\frac{1}{\text{ps}(L)}$
- untreated subjects with $\frac{1}{1 - \text{ps}(L)}$

where ps(*L*) is the propensity score Pr(*A* = 1 | *L*)



Sex	Treated	Propensity for treatment	Weight
Female	No	?	?
Female	Yes	?	?
Male	No	?	?
Male	Yes	?	?

IPW: example



Sex	Treated	Propensity for treatment	Weight
Female	No	4/6	3.0
Female	Yes	4/6	1.5
Male	No	2/6	1.5
Male	Yes	2/6	3.0

IPW: example

IPW: extensions

- For inference about average effect among the treated (ATT)
- To accommodate more than two treatment levels
- Point interventions with more than two levels or interventions on time-varying variables
- But with many possible treatment regimes to consider: there may be a need to make modelling assumptions regarding the distribution of the outcome and the treatment(s) in the pseudopopulation (marginal structural modelling)
- ‘Stabilisation’ of weights recommended (to reduce their variability)
- To address selective censoring or account for missing data

G-computation for binary treatments

- Alternative to IPW that comes under same (non-parametric) identifiability conditions, for the same estimand
- Relies on outcome modelling rather than treatment modelling
- Essentially “standardisation” of conditional quantities

Identification of marginal causal mean

$E[Y^{A=a}] = E\{E[Y^{A=a} \mid L]\}$ (Law of iterated expectations)

$= E\{E[Y^{A=a} \mid A=a, L]\}$ (Conditional exchangeability + positivity)

$= E\{E[Y \mid A=a, L]\}$ (Consistency)

The gist of g-estimation (for binary treatments)

- Postulate model for contrast between treatments within levels of covariates; eg, $E[Y^{A=a} \mid A=a,L] - E[Y^{A=0} \mid A=a,L] = \beta a$
- Express $E[Y^{A=0} \mid A=a,L]$ in terms of factuals under consistency; $E[Y^{A=0} \mid A=a,L] = E[Y - \beta A \mid A=a,L]$
- *Search* for model parameters that are *compatible with conditional exchangeability*; $E[Y^{A=0} \mid A,L] = E[Y^{A=0} \mid L]$, so search for $\hat{\beta}$ that renders A independent of $E[Y - \hat{\beta}A \mid A,L]$ given L

The gist of g-estimation (for binary treatments)

- Postulate model for contrast between treatments within levels of covariates; eg, $E[Y^{A=a} \mid L] - E[Y^{A=0} \mid L] = \beta a$
- Apply conditional exchangeability and consistency to express $E[Y^{A=0} \mid L]$ in terms of factuals;
  $E[Y^{A=0} \mid L] = E[Y - \beta A \mid A=a, L]$ for all a
- *Search* for model parameters that are *compatible* with the result; search for $\hat{\beta}$ that renders $E[Y - \hat{\beta} A \mid A,L]$ independent of A given L

International Journal of Epidemiology, 2017, 756–762
doi: 10.1093/ije/dyw323
Advance Access Publication Date: 30 December 2016
Education Corner

Education Corner

An introduction to g methods

Ashley I Naimi[1]*, Stephen R Cole[2] and Edward H Kennedy[3]

[1]Department of Epidemiology, University of Pittsburgh, [2]Department of Epidemiology, University of North Carolina at Chapel Hill and [3]Department of Statistics, Carnegie Mellon University

*Corresponding author. Department of Epidemiology University of Pittsburgh 130 DeSoto Street 503 Parran Hall Pittsburgh, PA 15261 ashley.naimi@pitt.edu

Accepted 17 October 2016

Abstract

Robins' generalized methods (g methods) provide consistent estimates of contrasts (e.g. differences, ratios) of potential outcomes under a less restrictive set of identification conditions than do standard regression methods (e.g. linear, logistic, Cox regression). Uptake of g methods by epidemiologists has been hampered by limitations in understanding both conceptual and technical details. We present a simple worked example that illustrates basic concepts, while minimizing technical complications.

Extentions of g-methods to time-varying treatment settings

- IPW, g-computation and g-estimation suitable to estimate causal effects of time-varying treatment effects
- … more on this later!

Statin-cancer example revisited (2)

Previous studies implicitly compared long-term statin users versus non-users – don't necessarily answer questions like …

- What would be my 10-year cancer risk if – possibly contrary to fact – I would start statin treatment now? And what if I wouldn't?
- What would be my 10-year cancer risk if – possibly contrary to fact – I would start statin treatment now and **adhered** to it? And what if I wouldn't start now *or* in the future?

Summary

Target trial emulation = *explicit* attempt to address deviations from a target trial, given the (observational) study data at hand

Step 1. *Specify* target trial
Step 2. *Emulate* it!

"The target trial framework provides an **organizing principle** for the design of observational studies that leads to **clinically interpretable results** and analytic approaches that can **prevent common biases**. Explicitly documenting the target trial that can be emulated in available observational data provides **a base for in-depth discussion** between experts to decide what is and is not acceptable in relation to study design. It also provides a link between observational studies and randomized trials, so the design quality of all studies that ask questions about the effectiveness and safety of medical treatments can be judged symmetrically."

Matthews et al., *JCE* 2023;164:112-115