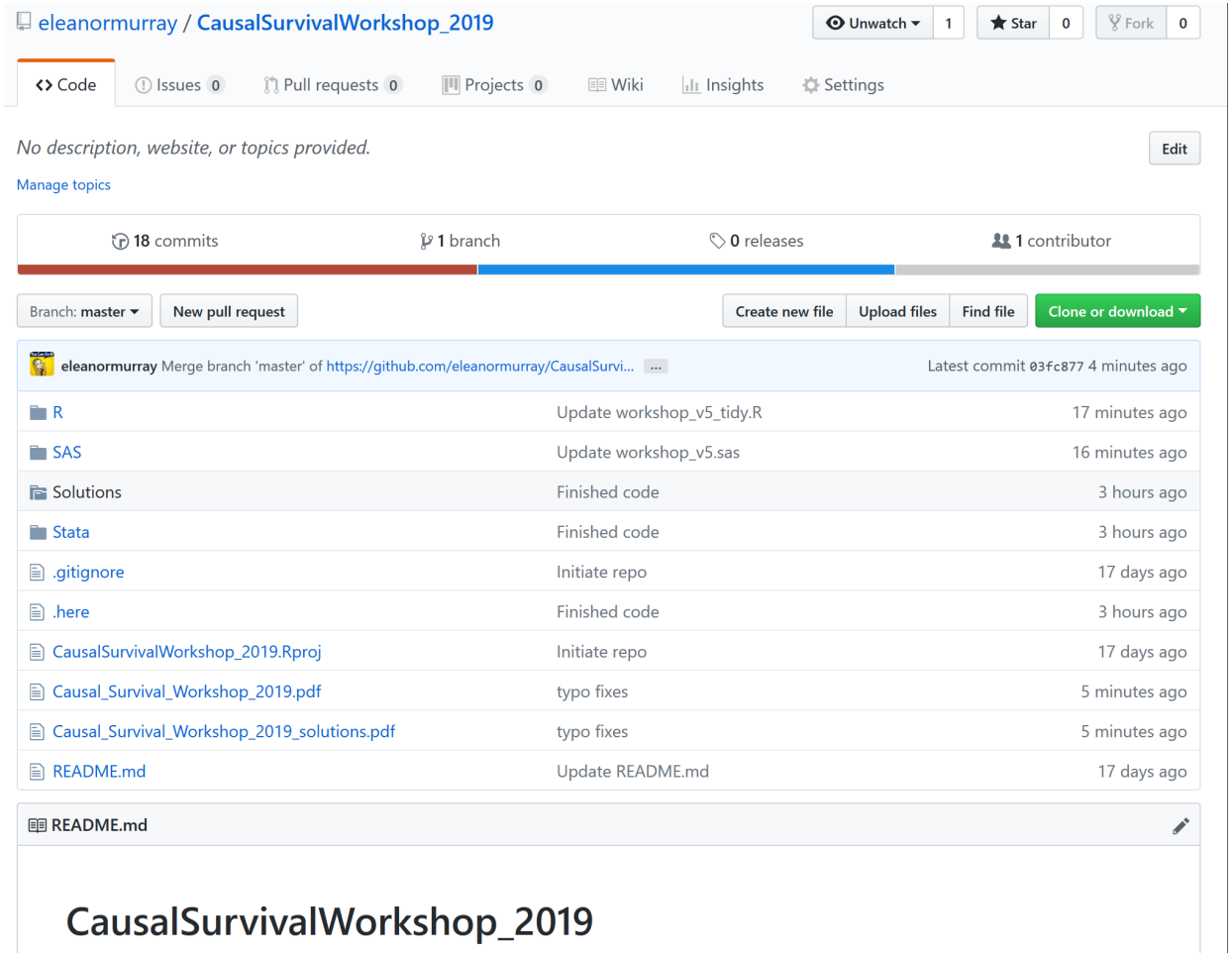
Download the workshop materials from:

https://github.com/eleanormurray/CausalSurvivalWorkshop_2019



Methods for Causal Inference from Randomized Trials with Loss to Follow-up or Non-adherence

Written by

Eleanor Murray, Ellen Caniglia, & Lucia Petito

ASA Biopharmaceutical Industry Regulatory Workshop

September 23, 2019

Instructors

Lucia Petito

Division of Biostatistics, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine

Ellen (Ellie) Caniglia

Department of Population Health, NYU Langone School of Medicine

Workshop objectives

At the end of this workshop, you will be able to:

- Understand the relative advantages and disadvantages of various effects frequently estimated in randomized trials and observational data, and their associated challenges
- Draw causal diagrams for intention-to-treat and per-protocol effects
- Estimate intention-to-treat hazard ratios and standardized survival curves
- Estimate inverse probability of treatment weights to adjust for loss to follow-up and differential non-adherence
- Estimate per-protocol hazard ratios and standardized survival curves

Schedule

1. Overview	30 min
2. Directed acyclic graphs for survival analysis	30 min
Break	15 min
3. Exercise 2: Estimating intention-to-treat effects	1 hour
Interlude: Adjusting for loss to follow-up	15 min
Break	15 min
4. Exercise 3: Estimating per-protocol effects	1 hour

Workshop outline

1. Overview

Why are we here and what are we doing?

This workshop is about causal survival analysis for longitudinal or follow-up data

We'll learn how to estimate causal effects of point exposures and static sustained exposures on survival

The case study: Coronary Drug Project (CDP)

NHLBI funded trial of lipid-lowering medications in men with a history of heart attacks

- Enrollment began in 1966, and follow-up ended in 1975
- 5 active treatments versus placebo

Coronary drug project research group *JAMA* 1975

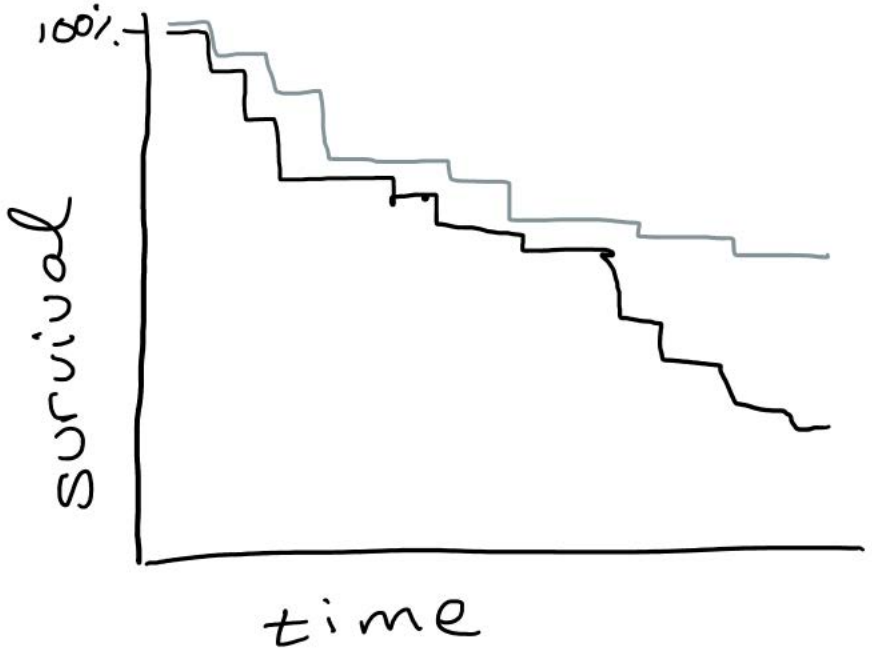
The case study: Coronary Drug Project (CDP)

Protocol	Description
Eligibility criteria	Men with a history of a myocardial infarction in previous 3 months 30–64 years old
Treatment arms	5 lipid-influencing drugs vs. placebo
Follow-up	Randomization until 5 years, LTFU, or death
Outcome	5-year mortality risk
Causal contrasts of interest	Intention-to-treat effect Effect of good adherence to trial protocol versus poor adherence, stratified by randomization arm

Coronary drug project research group *JAMA* 1975

Survival analysis basics

Survival analysis allows us to compare the **time-to-event** rather than just the number of events



The first challenge: administrative censoring

Let's define T as the time to death

- $T=1$ for subjects who die in month 1
- $T=2$ for subjects who die in month 2, etc.
- T is unknown for subjects who are alive at end of follow-up

This is called administrative censoring: time of death T is unknown for subjects who have not died by end of follow-up

Administrative censoring is different than loss to follow-up

Some people will drop out of our study. For these individuals,

- T is unknown after some point during follow-up

This is called loss to follow-up. We often need to deal with this in our studies, but for the workshop we have data with no loss to follow-up.

Some survival analysis definitions

Survival probability: $\Pr[T>k]$

Risk (cumulative incidence): $\Pr[T\leq k]$

- All deaths between baseline and time k
- Compared to the number of individuals at baseline

Discrete hazard: $\Pr[T=k+1|T>k]$

- Recent deaths occurring during interval k to $k+1$
- Compared to the number individuals alive (and under follow-up) at k

Defining exposures strategies

Point exposures are exposures that happen at a single point in time, typically at baseline. Randomization is (usually) a point exposure.

Sustained exposures are exposures that happen repeatedly over time. Many randomized trials assign participants to sustained exposures.

Defining exposures strategies

Static sustained exposures are sustained exposures that don't change over time. An example is "always eat vegetables".

Dynamic sustained exposures are sustained exposures which change over time, especially if the change is dependent on time-varying individual characteristics. An example is "take treatment unless a contraindication develops."

Stochastic sustained exposures are sustained exposures which probabilistically change over time – beyond the scope of this workshop

Defining exposures strategies

One last definition that may be useful:

A grace period is a pre-specified time frame in which you are interested in exposure happening. This is important for sustained exposures, especially in observational data.

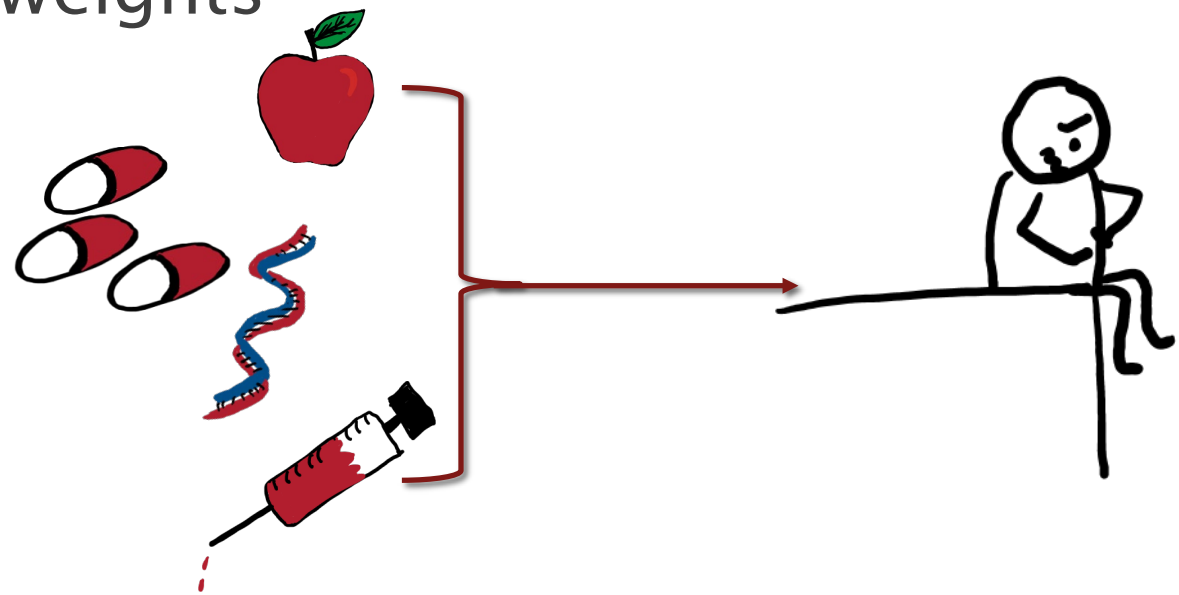
Causal inference for survival analysis is difficult because…

- Time-varying confounding–exposure feedback
- Time zero & immortal time bias
- Exposures can be hard to define well

These problems are less likely with a randomized trial, but common in observational studies

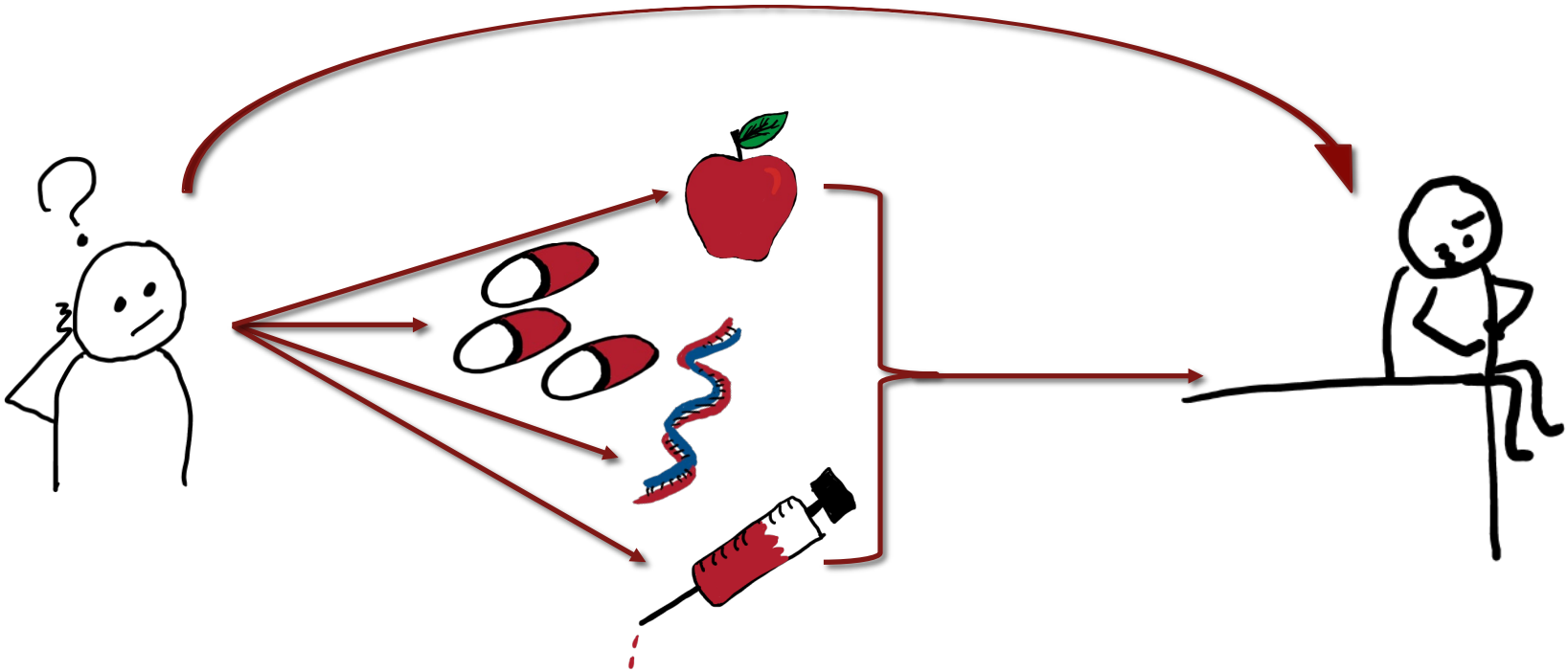
Why are well-defined exposures important?

- A well-defined exposure is one for which we can specify an 'intervention' that creates that exposure
- When there are multiple possible interventions, our answer is a weighted average of all 'interventions' but we don't know the weights



Why are well-defined exposures important?

- Worse, if the 'intervention' is ill-defined, the confounding is probably also ill-defined!



Solution: the target trial framework

First, specify in detail a randomized trial you would like to conduct

Second, design your observational study so that it closely mimics this target trial

But, to design a good target trial, we need to understand randomized trials!

A little about our data

The datasets on the workshop github are simulated versions of the Coronary Drug Project trial

The data are in long format, which means we have 1 observation per person per visit

Long-format data

ID	time	Z	L_0	L_t	C_t	Y_t
1	0	1	0	0	0	0
1	1	1	0	1	0	0
1	2	1	0	1	0	1
2	0	0	1	1	0	0
2	1	0	1	0	0	0
2	2	0	1	0	1	.
3	0	1	1	1	0	0
3	1	1	1	0	0	0
3	2	1	1	0	0	0

We have 15 time-varying covariates, and 1 baseline-only variable

Table 2: Data Description

Variable	Notation	Definition	Values
simID	id	ID variable	Range: 0 to 4042
rand	Z	Randomization group	0: placebo; 1: clofibrate
visit	t	Visit number: visits occur quarterly	Range: 0 to 14
death	Y_t	Death recorded at visit t	0: alive; 1: died $\mathrm{T} \in [t, t+1)$
adhpre0bin	A_{-1}	Adherence to placebo during pre-randomization run-in period	0: adherence $\geq$ 80%; 1: adherence $<$ 80%
adhr_b	A_0	Adherence to placebo at baseline (recorded at a special visit 2 weeks after randomization)	0: adherence $\geq$ 80%; 1: adherence $<$ 80%
adhr	A_t	Adherence to assigned treatment at visit t	0: adherence $\geq$ 80%; 1: adherence $<$ 80%
mi_bin	L	Myocardial infarction at baseline	0: 1 or 2; 1: $\geq$2
AntiHyp	L	Antihypotensive medication use at visit t	0: no; 1: yes
AnyQQS	L	ECG findings: Q/QS pattern at visit t	0: no; 1: yes
AnySTDep	L	ECG findings: ST depression at visit t	0: no; 1: yes
AP	L	Angina pectoris at visit t	0: no; 1: yes
CardioM	L	Cardiomegaly at visit t	0: no; 1: yes
CHF	L	Coronary heart failure at visit t	0: no; 1: yes
DIUR	L	Diuretic use at visit t	0: no; 1: yes
FVEB	L	ECG finding: Frequent ventricular beats at visit t	0: no; 1: yes
HiHeart	L	High heart rate at visit t	0: $<$70 bpm; 1: $\geq$ 70 bpm
HiSerChol	L	High serum cholesterol at visit t	0: $<$250; 1: $\geq$ 250
HiSerTrigly	L	High serum triglyesterol at visit t	0: $<$5.0; 1: $\geq$ 5.0
IC	L	Intermittent claudication at visit t	0: no; 1: yes
NIHA	L	New York Heart Association class at visit t	0: no limitations; 1: any limitation
OralHyp	L	Oral hypoglycemic agents use at visit t	0: no; 1: yes
VCD	L	ECG finding: Ventricular conduction defect at visit t	0: no; 1: yes

2. Directed acyclic graphs for survival analysis

First, we need to choose an estimand

Directed acyclic graphs (DAGs) are a tool for summarizing the information we know about a research question we want to answer

Each DAG should be targeted to a specific question, so we first need to decide what we want to estimate

What types of causal estimands?

In randomized trials we can estimate:

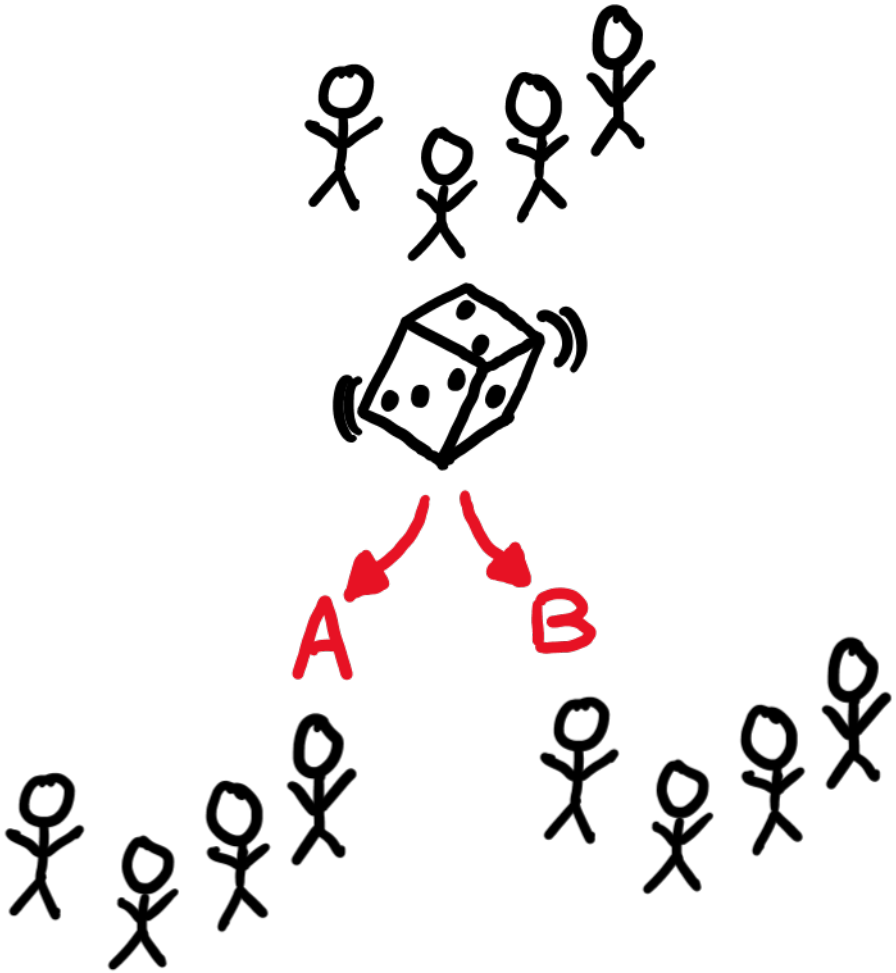
- Intention-to-treat effect
- Per-protocol effect(s)

In an observational study we can generally only estimate:

- Per-protocol effect(s)

Per-protocol analyses have a bad reputation!

But per-protocol effects are what we actually want to know!



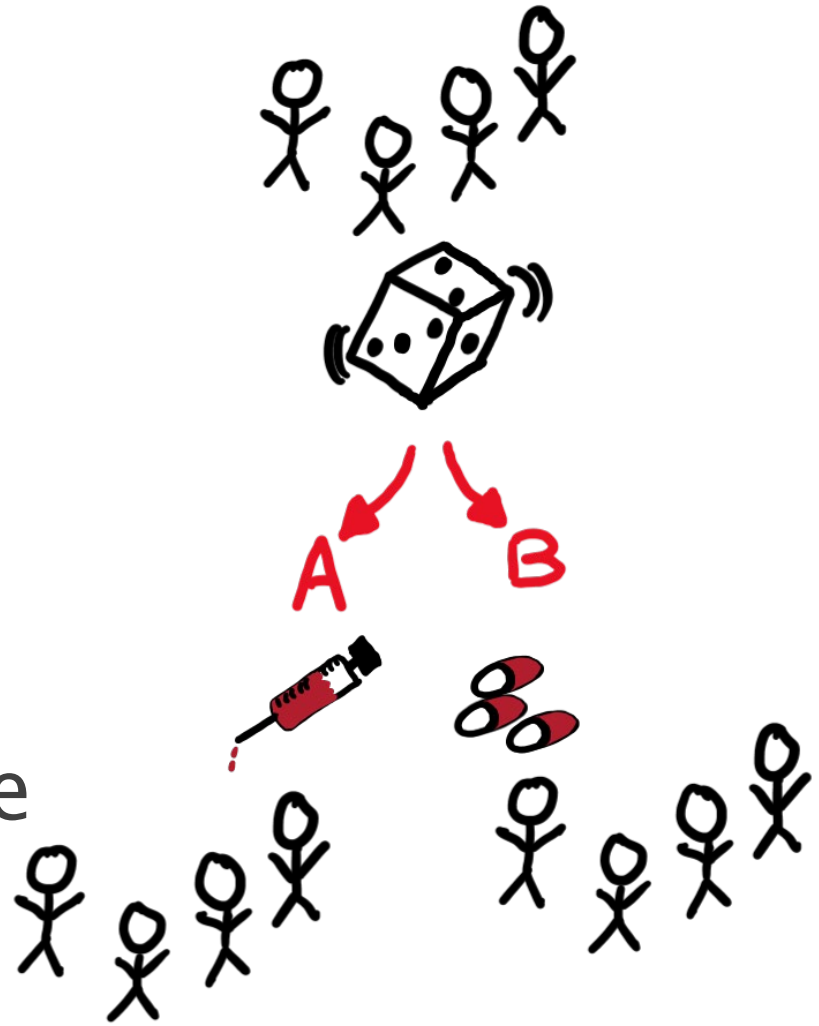
The intention-to-treat effect is the effect of **randomization**

If everyone initiates treatment, it can also be interpreted as the effect of initiating assigned treatment

But per-protocol effects are what we actually want to know!

The per-protocol effect is the effect of **receiving assigned treatment**

This doesn't have to be continuous, it can be according to some treatment *protocol*



What assumptions do we make to assume a causal interpretation?

Causal inference relies on three main assumptions:

- Exchangeability

$$Y_i^a \coprod A_i$$

- Positivity

$$\Pr[A_i = a] > 0 \; \forall \; a \in \mathbf{A}$$

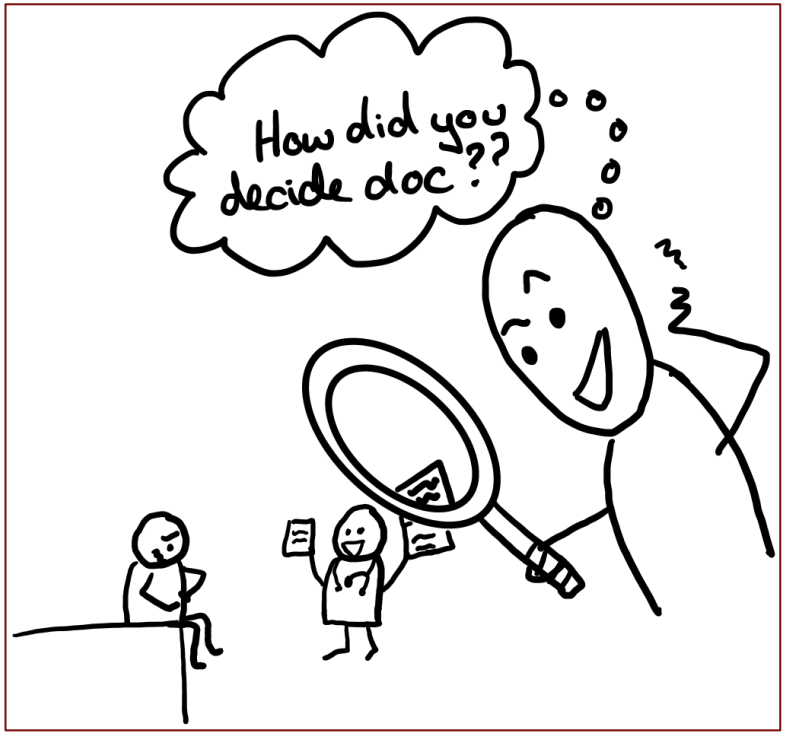
- Consistency

$$Y_i^a = Y_i \mid A_i = a$$

What is exchangeability?

- No unmeasured confounding: all common causes of the treatment and outcome are known and measured in the data

- No selection bias: we haven't conditioned or restricted on a variable that is a common effect of exposure and the outcome (or outcome cause)

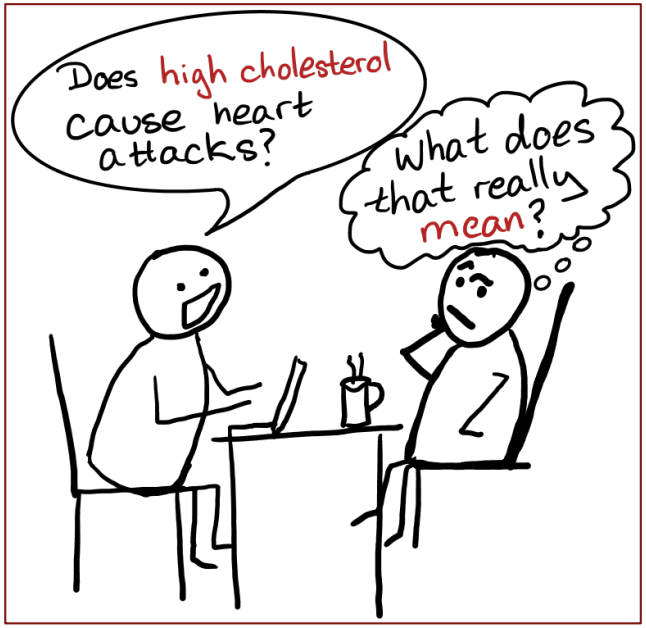


What is positivity?

- Positivity: there is a non-zero probability of all levels of treatment for all types of individuals in our population

What is consistency?

- Consistency: our treatment levels are clearly specified, aka:
  - Well-defined interventions
  - Well-defined causal questions



Ok, but why do we bother with intention-to-treat effects?

Randomization ensures no confounding at baseline for treatment assignment

Randomization also ensures positivity for treatment assignment

Randomization is a well-defined intervention

So, intention-to-treat analyses often give unbiased estimates of intention-to-treat effects

But not so fast!

Randomization ensures no confounding at baseline for treatment assignment

- Treatment happens after randomization
- Loss to follow-up happens after randomization

Post-randomization events are not guaranteed to be unconfounded!

Also the effect of randomization is not very interesting

- Often a lower bound on the effect of treatment compared to placebo
- Lower bound is insufficient for adverse events or safety
- When comparing active treatments, ITT can vary towards or away from the null
- No real world, clinical, equivalent of randomization
- Depends on the distribution of adherence in the trial & this can affect external validity

Per-protocol effects we could estimate:

- Effect of initiating treatment
- Effect of adhering to treatment protocol
- Effect of receiving point intervention, among the compliers (note, not adherers!)

Effect of treatment is an interesting effect

- Relevant for real world, clinical, decision making
- Allows better risk assessment for adverse events or safety
- Interpretable for both placebo and active / usual care comparators
- Doesn't depend on adherence patterns!

Per-protocol effect is the effect we really want!

Plus, per-protocol effects are patient-centered causal effects

- Per-protocol effects are most relevant for people who can determine their own exposure, and have a high level of concern
- Discuss how/ why exposure occurred for maximum patient relevance

Murray EJ, et al. *Journal of Clinical Epidemiology*. 2018; 103: 10-21.

But, since we have a trial why not both?!

What about loss to follow-up?

- Our simulated data has no loss to follow-up
- If it did, we may need to adjust for differential loss to follow-up
- Remember, randomization only ensures exchangeability at baseline
- If loss to follow-up is differential, even the intention-to-treat effect can be biased

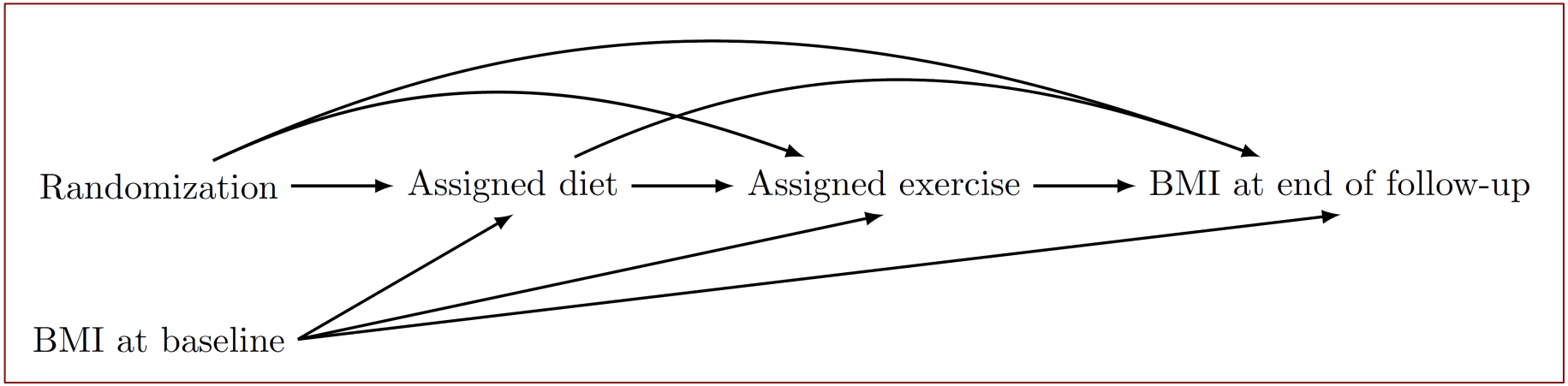
What effects are we really estimating?

- The intention-to-treat effect is the effect of randomization had no one been lost
- The per-protocol effect is the effect of following the protocol and not being lost
- These are really joint interventions

Exercise 1: Drawing directed acyclic graphs to inform our analyses

An aside on DAGs

A directed acyclic graph (DAG) is a tool for visualizing the relationships between factors in the real world & in your study



An aside on DAGs

DAGs have 2 main components:

- Nodes: represent variables of interest for your study, including known or unknown confounders
- Arrows: represent possible causal relationships between variables in the direction of the arrow

An aside on DAGs

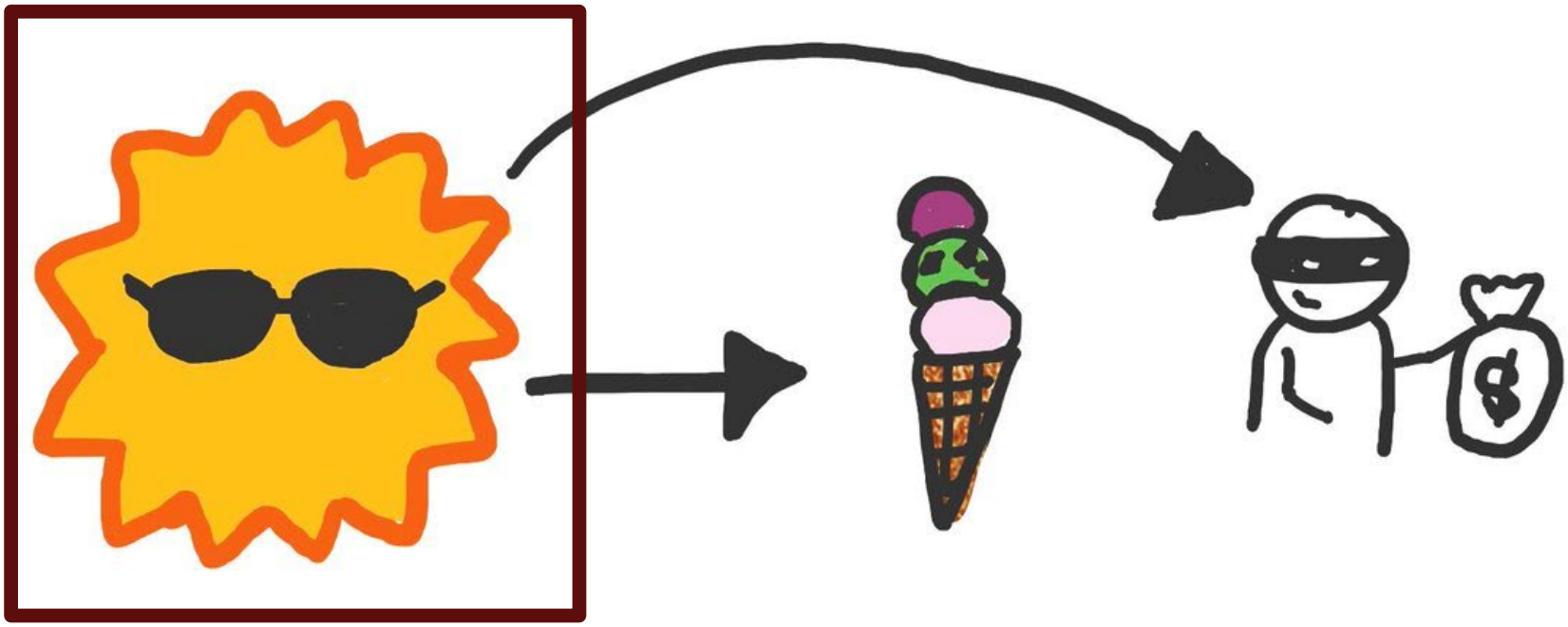
DAGs encode an underlying factorization of the observed data distribution

The absence of an arrow between 2 nodes implies their independence – this is a ***strong*** assumption

DAGs an be interpreted causally if you assume an underlying causal data generating process

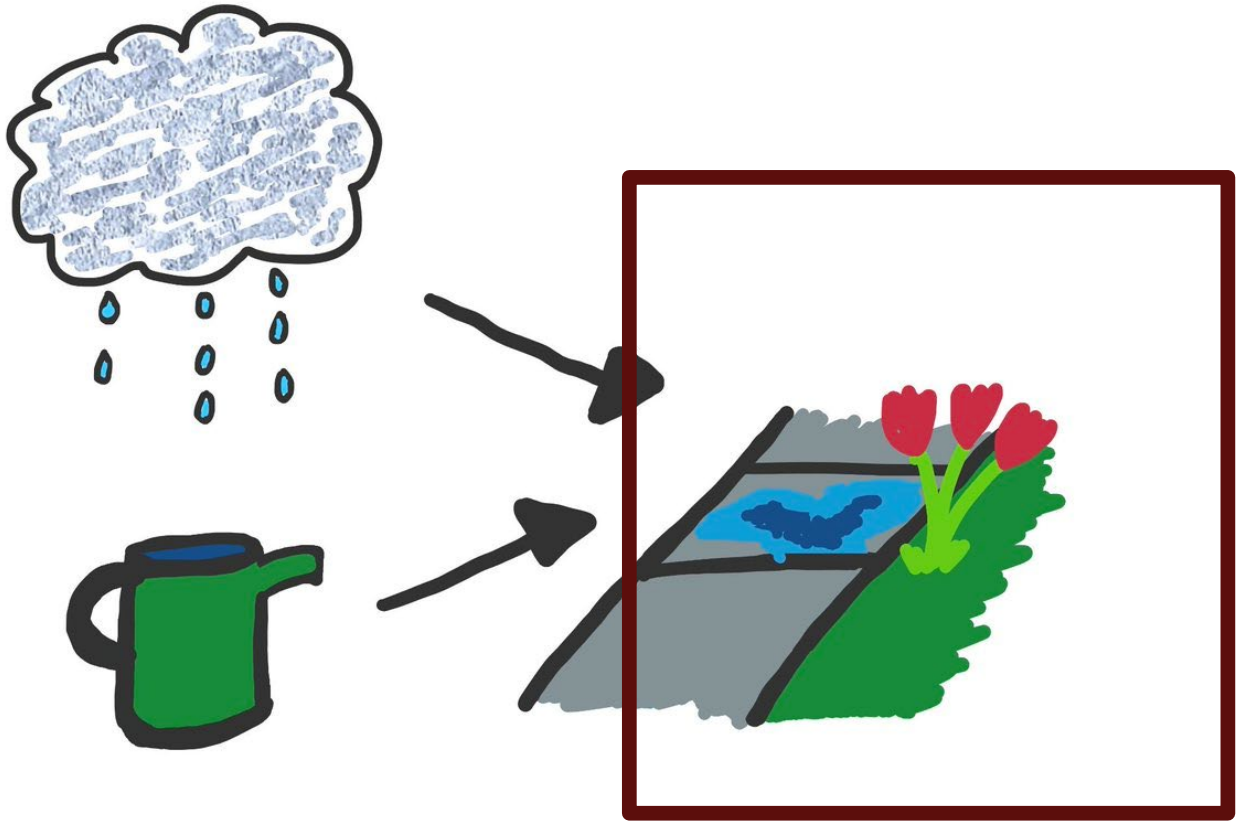
An aside on DAGs

DAGs let us read potential biases easily



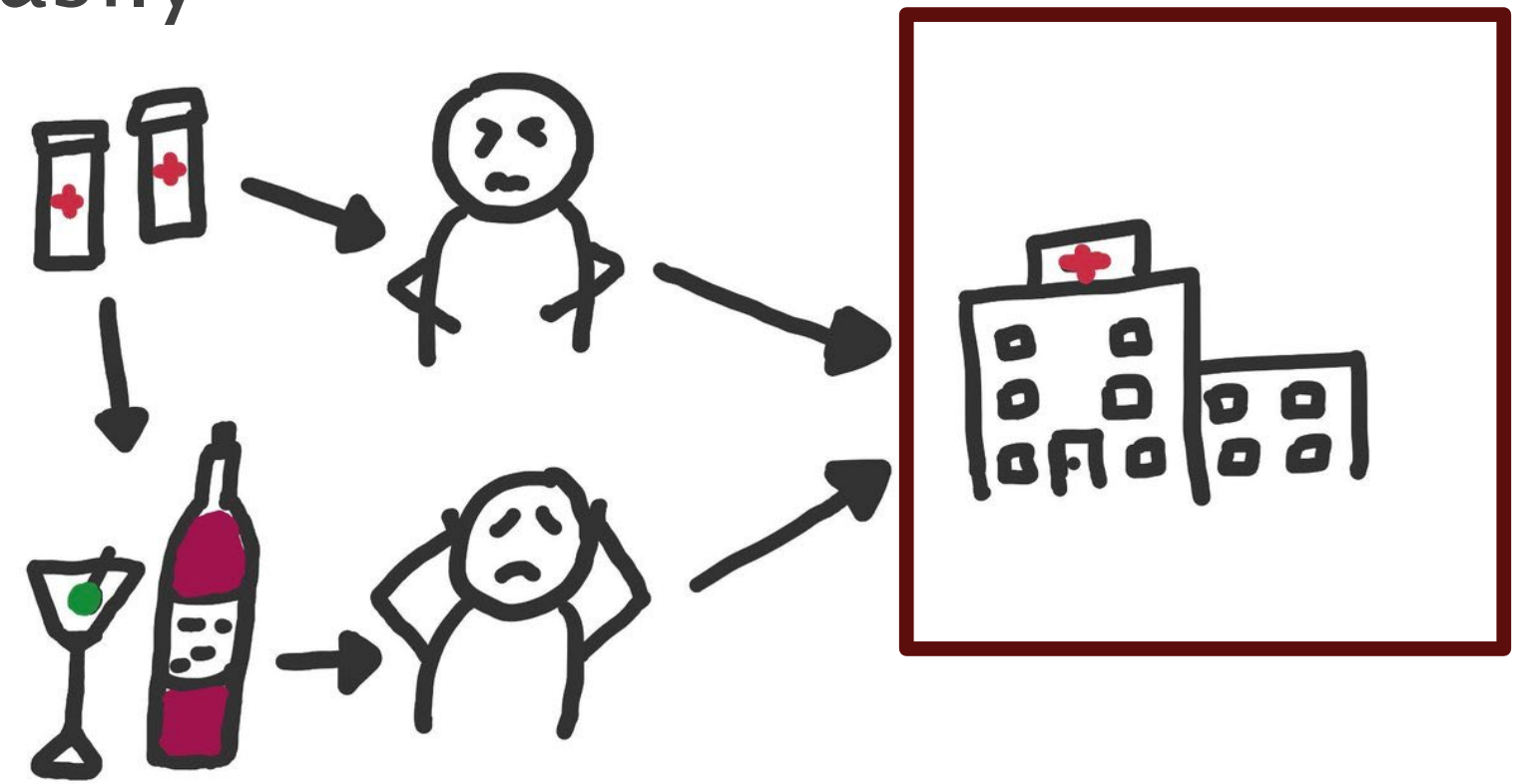
An aside on DAGs

DAGs let us read potential biases easilv



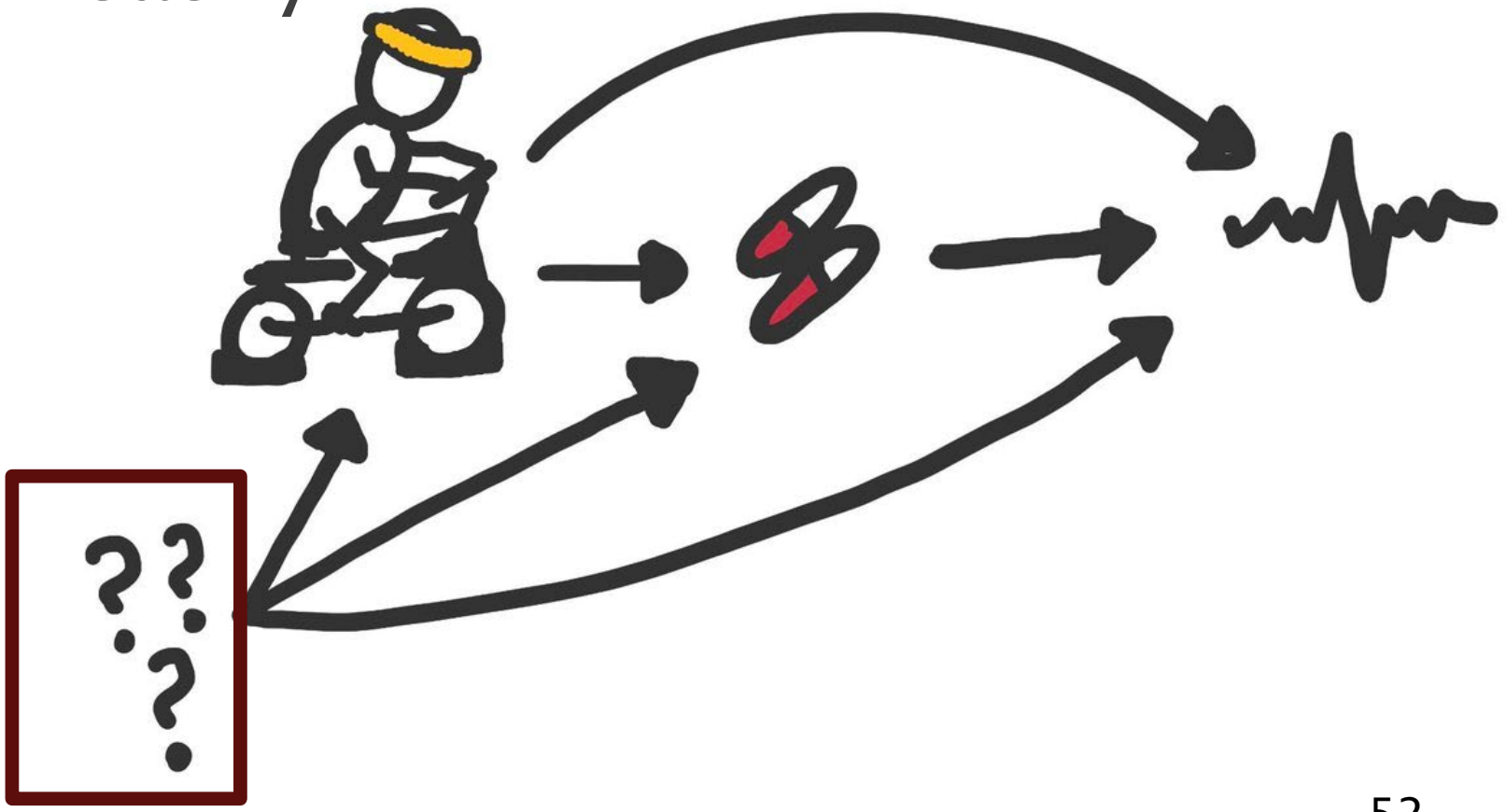
An aside on DAGs

DAGs let us read potential biases easily



An aside on DAGs

DAGs let us read potential biases easily



An aside on DAGs

Where to go for more resources:

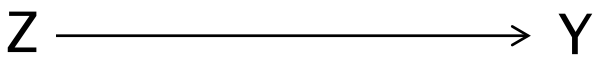
- Greenland, Pearl, & Robins. 1999. *Epidemiology*, 10(1):37–48
- Hernan & Robins. *Causal Inference*. Available online at: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
- Murray. *Graphical models for causal inference using LaTeX*. Available online at: https://github.com/eleanormurray/causalgraphs_latex

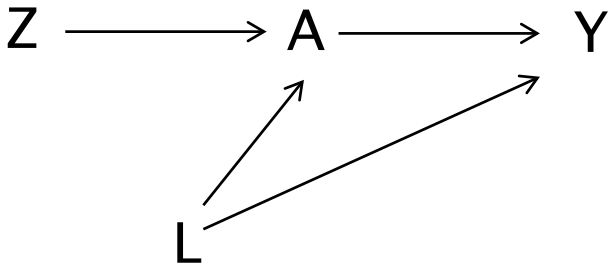
So, let's draw an intention-to-treat DAG

Go to handout, page 7:

2.2 Exercise 1: Drawing the DAG

2 ways to draw an intention-to-treat DAGs (assuming no loss to follow-up)



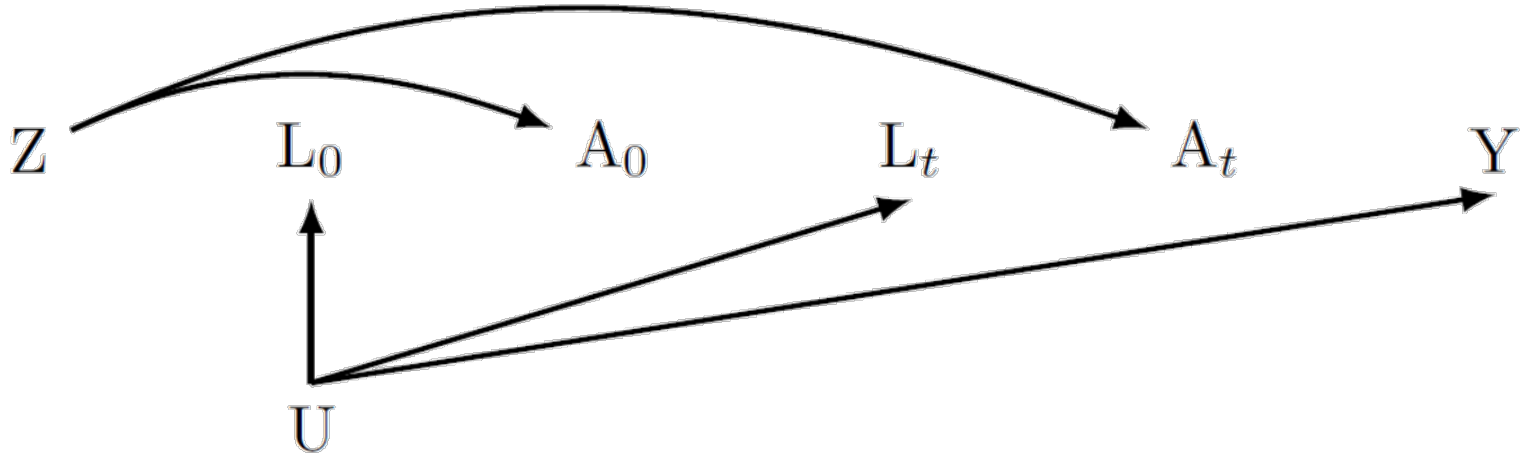


Next, let's draw a per-protocol effect DAG

Go to handout, page 8:

2.2.2 Exercise 1: DAG for the per-protocol effect

Different assumptions lead to different DAGs, and different analyses

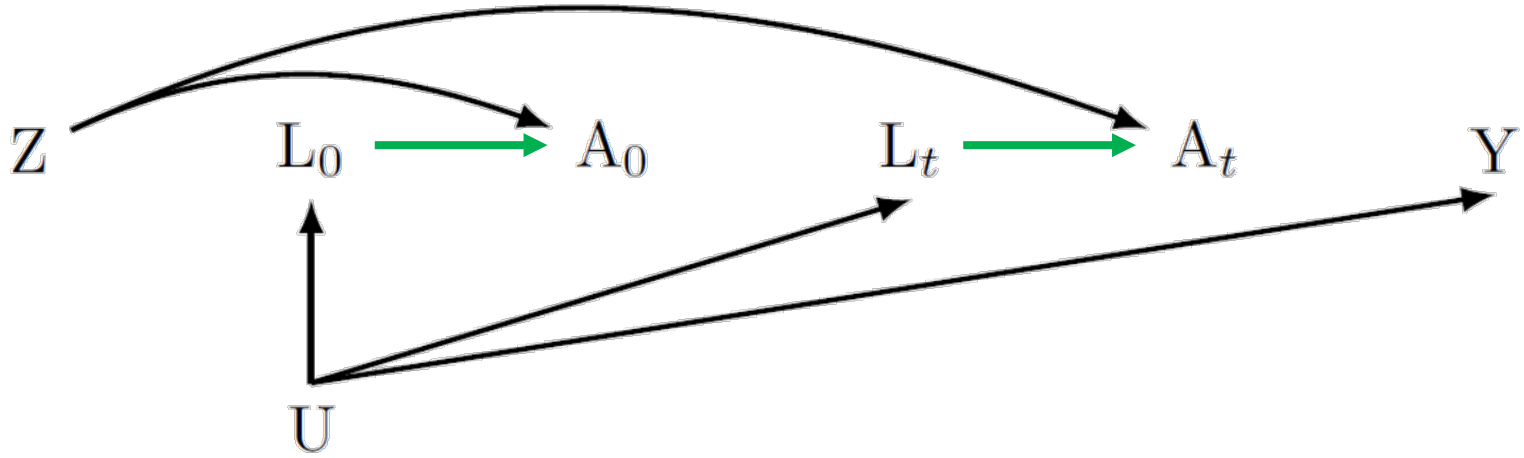


Random non-adherence

- No confounding adjustment needed

Hernan & Robins. 2017, NEJM 377:14

Different assumptions lead to different DAGs, and different analyses

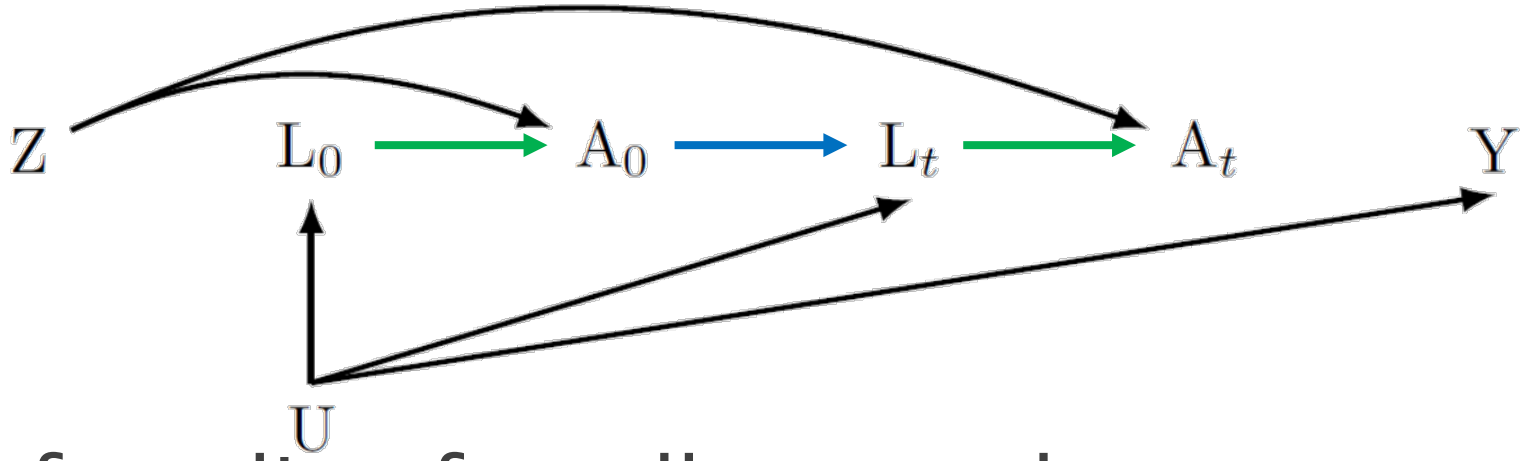


Confounding for adherence by measured covariates

- Adjustment required using any method

Hernan & Robins. 2017, NEJM 377:14

Different assumptions lead to different DAGs, and different analyses

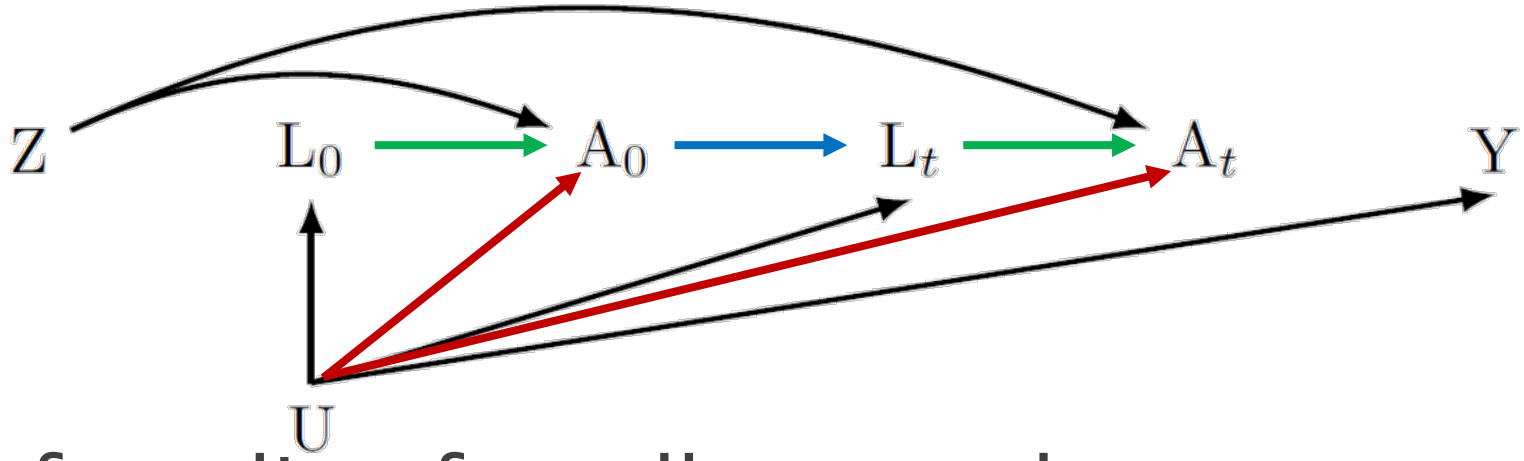


Confounding for adherence by measured covariates and prior adherence

- G-methods required

Hernan & Robins. 2017, NEJM 377:14

Different assumptions lead to different DAGs, and different analyses



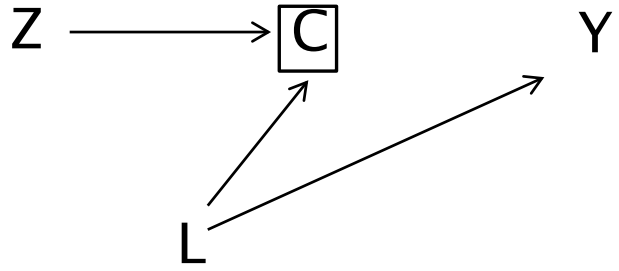
Confounding for adherence by measured covariates, prior adherence, and unmeasured covariates

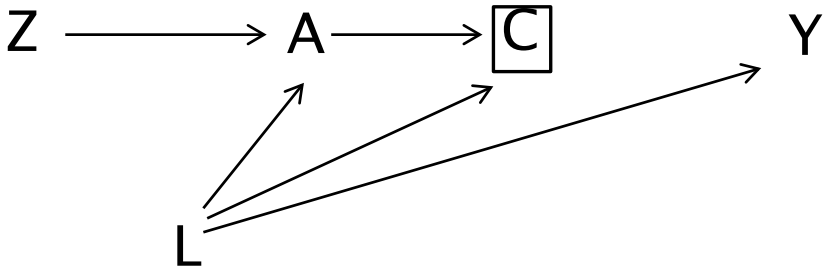
- Strong assumptions + structural nested models

Hernan & Robins. 2017, NEJM 377:14

Finally, let's draw an intention-to-treat DAG with loss to follow-up

2 ways to draw an intention-to-treat DAGs (with loss to follow-up)





15 minute break

3. Exercise 2: Estimating intention-to-treat effects

Make sure you have the data downloaded

In the handout, page 9:
Section 3.1 Data exploration

In your preferred coding language, go to:
Exercise 2, Code Section 1 Data exploration

Answer Questions 1-5

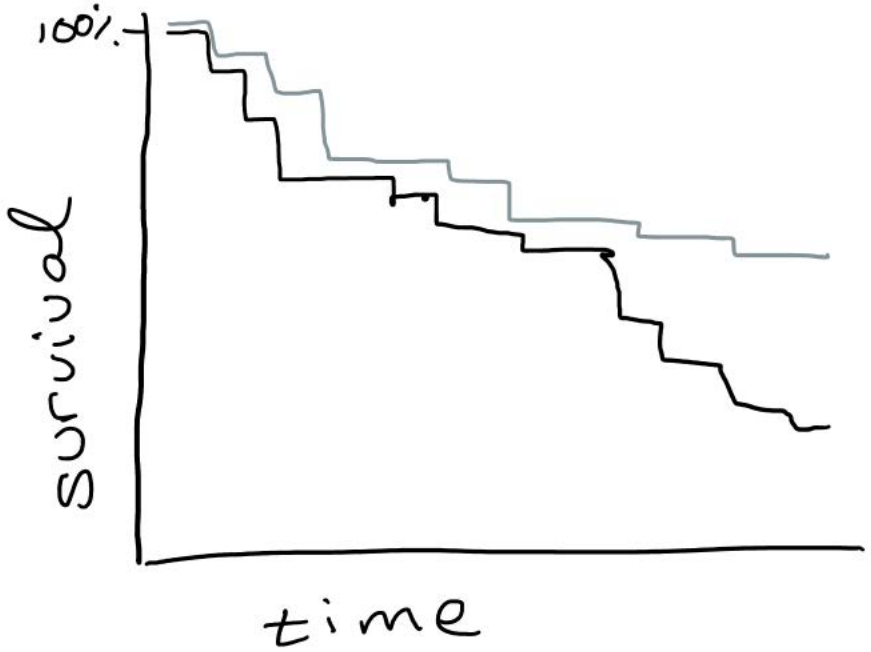
Non-parametric survival estimates

Since we have no loss to follow-up, we can estimate the causal effect of randomization on survival without making any additional assumptions

We can do this by comparing the Kaplan-Meier survival curves

Reminder: Survival curves

Kaplan-Meier curves look like staircases



Reminder: some definitions

Let's define T as the time to death

- $T=1$ for subjects who die in month 1
- $T=2$ for subjects who die in month 2, etc.
- T is unknown for subjects who are alive at end of follow-up

Quick reminder: survival analysis definitions

Survival probability: $\Pr[T>k]$

Risk (cumulative incidence): $\Pr[T\leq k]$

- All deaths between baseline and time k
- Compared to the number of individuals at baseline

Discrete hazard: $\Pr[T=k+1|T>k]$

- Recent deaths occurring during interval k to $k+1$
- Compared to the number individuals alive (and under follow-up) at k

Kaplan–Meier survival

Kaplan–Meier estimates survival at *k* as the product of conditional probabilities of having survived each interval between 0 and *k*

The conditional probability at each time is one minus the discrete hazard

For example, for $k=2$

$$\Pr[D_2=0] = \Pr[D_1=0]\Pr[D_2=0|D_1=0]$$

In the handout, page 11:

3.2 Kaplan-Meier survival curves

In your preferred coding language, go to:

Exercise 2, Code Section 2 Kaplan-Meier

Answer Questions 1-3

Semi-parametric ITT estimate

We can easily estimate the intention-to-treat effect on the hazard ratio scale if we're willing to assume:

- The hazard ratio is constant over follow-up

Now we can use Cox proportional hazards regression

Parametric ITT estimate

If we're willing to make one more assumption, we can estimate the hazard ratio and the survival, plus the risk difference and risk ratio

- The baseline hazard has a known specified functional form
- (here we assume quadratic)

Now we can use pooled logistic regression*

*Other types of regression models can be used to estimate the hazard in discrete time; for brevity, in this workshop we only discuss pooled logistic regression

In the handout, page 12:
3.3.1 Unadjusted intention-to-treat effects

In your preferred coding language, go to:
Exercise 2, Code Section 3a Unadjusted Hazard Ratios

Fill in the first 3 rows of the table on pg. 13
Answer questions 1-3 on pg. 13

Baseline covariate adjustment

Both semi-parametric and parametric survival methods allow adjustment for baseline covariates to estimate conditional hazard ratios

In the handout, page 14:

3.3.2 Baseline-adjusted intention-to-treat effects

In your preferred coding language, go to:

Exercise 2, Code Section 3b Adjusted Hazard Ratios

Fill in the last 3 rows of the table on pg. 13

Answer questions 1-5 on pages 14-15

Maintaining interpretability

Conditional hazard ratios are hard to interpret

We would prefer to know the average hazard ratio, or even better the average survival and risks

We can estimate these from our pooled logistic regression model using standardization

id	rand	sex	age	race	death
1	0	0	65	1	0
2	1	0	58	1	1
3	0	1	67	1	1
4	1	0	54	0	0
5	0	1	48	0	0

id	rand	sex	age	race	death	Interv
1	0	0	65	1	0	0
2	1	0	58	1	1	0
3	0	1	67	1	1	0
4	1	0	54	0	0	0
5	0	1	48	0	0	0
1	0	0	65	1	0	1
2	1	0	58	1	1	1
3	0	1	67	1	1	1
4	1	0	54	0	0	1
5	0	1	48	0	0	1

```
expand 2, gen(interv)
```

```
expand 2 if interv == 0, gen(interv2
```

id	rand	sex	age	race	death	Interv	Interv 2
1	0	0	65	1	0	0	0
2	1	0	58	1	1	0	0
3	0	1	67	1	1	0	0
4	0	0	54	0	0	0	0
5	0	1	48	0	0	0	0
1	0	0	65	1	0	1	0
2	1	0	58	1	1	1	0
3	0	1	67	1	1	1	0
4	1	0	54	0	0	1	0
5	0	1	48	0	0	1	0
1	0	0	65	1	0	.	1
2	1	0	58	1	1	.	1
3	0	1	67	1	1	.	1
4	1	0	54	0	0	.	1
5	0	1	48	0	0	.	1

id	rand	sex	age	race	death	Interv	Interv 2
1	0	0	65	1	0	0	0
2	1	0	58	1	1	0	0
3	0	1	67	1	1	0	0
4	0	0	54	0	0	0	0
5	0	1	48	0	0	0	0
1	0	0	65	1	0	1	0
2	1	0	58	1	1	1	0
3	0	1	67	1	1	1	0
4	1	0	54	0	0	1	0
5	0	1	48	0	0	1	0
1	0	0	65	1	0	-1	1
2	1	0	58	1	1	-1	1
3	0	1	67	1	1	-1	1
4	1	0	54	0	0	-1	1
5	0	1	48	0	0	-1	1

```
replace interv = -1 if interv2 == 1
```

id	rand	sex	age	race	death	Interv
1	0	0	65	1	0	0
2	1	0	58	1	1	0
3	0	1	67	1	1	0
4	0	0	54	0	0	0
5	0	1	48	0	0	0
1	0	0	65	1	0	1
2	1	0	58	1	1	1
3	0	1	67	1	1	1
4	0	0	54	0	0	1
5	0	1	48	0	0	1
1	0	0	65	1	0	-1
2	1	0	58	1	1	-1
3	0	1	67	1	1	-1
4	0	0	54	0	0	-1
5	0	1	48	0	0	-1

```
Drop interv2
```

id	rand	sex	age	race	death	Interv
1	0	0	65	1	.	0
2	1	0	58	1	.	0
3	0	1	67	1	.	0
4	1	0	54	0	.	0
5	0	1	48	0	.	0
1	0	0	65	1	.	1
2	1	0	58	1	.	1
3	0	1	67	1	.	1
4	1	0	54	0	.	1
5	0	1	48	0	.	1
1	0	0	65	1	0	-1
2	1	0	58	1	1	-1
3	0	1	67	1	1	-1
4	1	0	54	0	0	-1
5	0	1	48	0	0	-1

```
Replace death = . If interv != -1
```

id	rand	sex	age	race	death	Interv
1	0	0	65	1	.	0
2	0	0	58	1	.	0
3	0	1	67	1	.	0
4	0	0	54	0	.	0
5	0	1	48	0	.	0
1	0	0	65	1	.	1
2	1	0	58	1	.	1
3	0	1	67	1	.	1
4	1	0	54	0	.	1
5	0	1	48	0	.	1
1	0	0	65	1	0	-1
2	1	0	58	1	1	-1
3	0	1	67	1	1	-1
4	1	0	54	0	0	-1
5	0	1	48	0	0	-1

```
Replace death = . If interv != -1
Replace rand = 0 if interv == 0
```

id	rand	sex	age	race	death	Interv
1	0	0	65	1	.	0
2	0	0	58	1	.	0
3	0	1	67	1	.	0
4	0	0	54	0	.	0
5	0	1	48	0	.	0
1	1	0	65	1	.	1
2	1	0	58	1	.	1
3	1	1	67	1	.	1
4	1	0	54	0	.	1
5	1	1	48	0	.	1
1	0	0	65	1	0	-1
2	1	0	58	1	1	-1
3	0	1	67	1	1	-1
4	1	0	54	0	0	-1
5	0	1	48	0	0	-1

```
Replace death = . If interv != -1
Replace rand = 0 if interv == 0
Replace rand = 1 if interv == 1
```

id	rand	sex	age	race	death	Interv
1	0	0	65	1	.	0
2	0	0	58	1	.	0
3	0	1	67	1	.	0
4	0	0	54	0	.	0
5	0	1	48	0	.	0
1	1	0	65	1	.	1
2	1	0	58	1	.	1
3	1	1	67	1	.	1
4	1	0	54	0	.	1
5	1	1	48	0	.	1
1	0	0	65	1	0	-1
2	1	0	58	1	1	-1
3	0	1	67	1	1	-1
4	1	0	54	0	0	-1
5	0	1	48	0	0	-1

```
Replace death = . If interv != -1
Replace rand = 0 if interv == 0
Replace rand = 1 if interv == 1
```

Model will be fitted in this data, since it doesn't have missing values

id	rand	sex	age	race	death	Interv	pr_d
1	0	0	65	1	.	0	#
2	0	0	58	1	.	0	#
3	0	1	67	1	.	0	#
4	0	0	54	0	.	0	#
5	0	1	48	0	.	0	#
1	1	0	65	1	.	1	#
2	1	0	58	1	.	1	#
3	1	1	67	1	.	1	#
4	1	0	54	0	.	1	#
5	1	1	48	0	.	1	#
1	0	0	65	1	0	-1	#
2	1	0	58	1	1	-1	#
3	0	1	67	1	1	-1	#
4	1	0	54	0	0	-1	#
5	0	1	48	0	0	-1	#

```
Predict pr_d, pr
```

id	rand	sex	age	race	death	Interv	pr_d
1	0	0	65	1	.	0	#
2	0	0	58	1	.	0	#
3	0	1	67	1	.	0	#
4	0	0	54	0	.	0	#
5	0	1	48	0	.	0	#
1	1	0	65	1	.	1	#
2	1	0	58	1	.	1	#
3	1	1	67	1	.	1	#
4	1	0	54	0	.	1	#
5	1	1	48	0	.	1	#
1	0	0	65	1	0	−1	#
2	1	0	58	1	1	−1	#
3	0	1	67	1	1	−1	#
4	1	0	54	0	0	−1	#
5	0	1	48	0	0	−1	#

$E[Y^{a=0}]$

$E[Y^{a=1}]$

Kaplan–Meier survival

Kaplan–Meier estimates survival at *k* as the product of conditional probabilities of having survived each interval between 0 and *k*

The conditional probability at each time is one minus the discrete hazard

For example, for $k=2$

$$\Pr[D_2=0] = \Pr[D_1=0]\Pr[D_2=0|D_1=0]$$

In the handout, page 15:

3.4 Standardizing over baseline covariates to estimate marginal causal effects

In your preferred coding language, go to:
Exercise 2, Code Section 4 Marginal Effects

Complete table on page 18
Answer questions 1–4 on page 18–19

Interlude: Adjusting for loss to follow-up

Reminder

- Randomization only ensures exchangeability at baseline
- If loss to follow-up is differential, even the intention-to-treat effect can be biased
- Inverse probability of censoring weights can be used to adjust for differential loss to follow-up in intention-to-treat and per-protocol analyses

Inverse probability of censoring weights

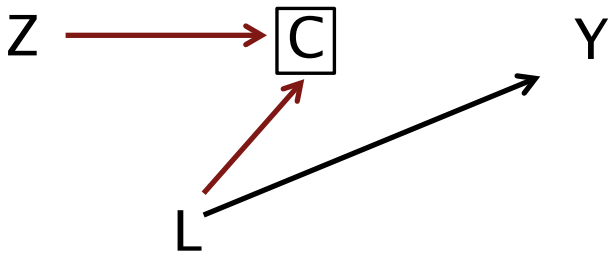
- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[C_j=0|Z,\bar{L}_j,\bar{C}_{j-1}=0]}$

- $SW_t = \prod_{j=0}^{t} \frac{Pr[C_j=0|Z]}{\Pr[C_j=0|Z,\bar{L}_j,\bar{C}_{j-1}=0]}$

- At each time, each person receives a weight inversely proportional to the probability of remaining uncensored, conditional on randomization and time-varying covariates

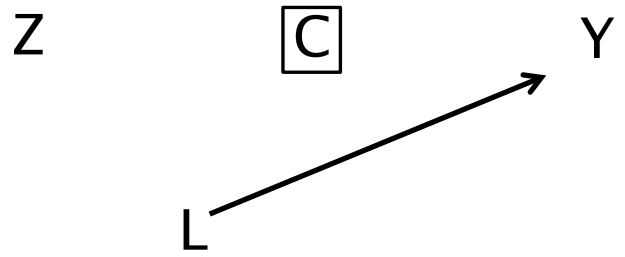
Adjusting for loss to follow-up

- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[C_j=0|Z,\bar{L}_j,\bar{C}_{j-1}=0]}$



Adjusting for loss to follow-up

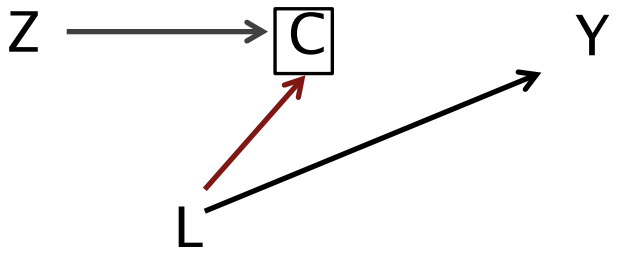
- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[C_j=0|Z,\bar{L}_j,\bar{C}_{j-1}=0]}$



Non-stabilized weights create a pseudo-population with no selection!

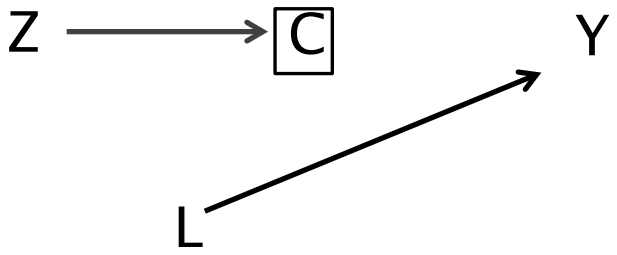
Adjusting for loss to follow-up

- $SW_t = \prod_{j=0}^{t} \frac{Pr[C_j=0|Z,\bar{C}_{j-1}=0]}{\Pr[C_j=0|Z,\bar{L}_j,\bar{C}_{j-1}=0]}$



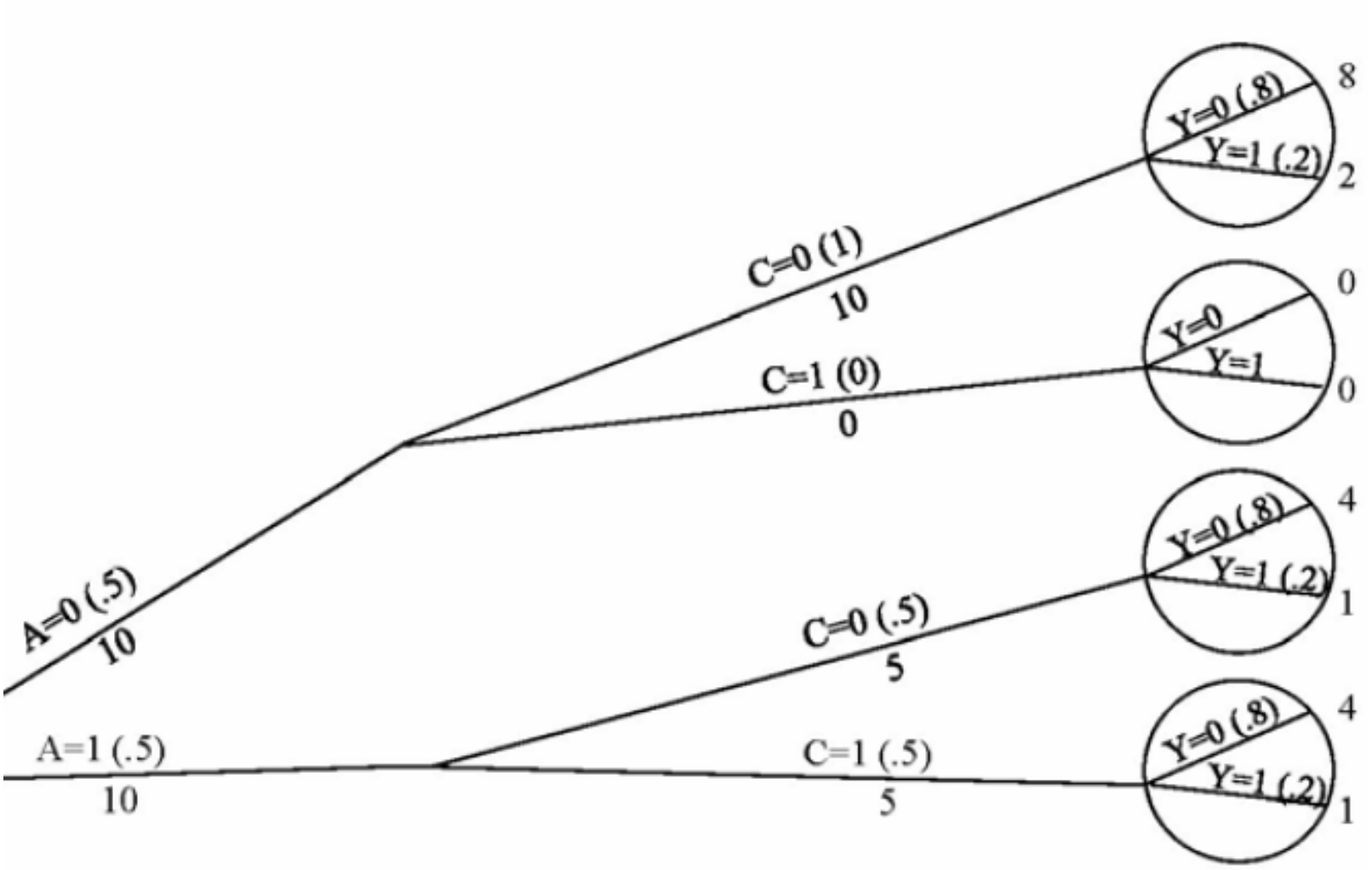
Adjusting for loss to follow-up

- $SW_t = \prod_{j=0}^{t} \frac{Pr[C_j=0|Z,\bar{C}_{j-1}=0]}{\Pr[C_j=0|Z,\bar{L}_j,\bar{C}_{j-1}=0]}$

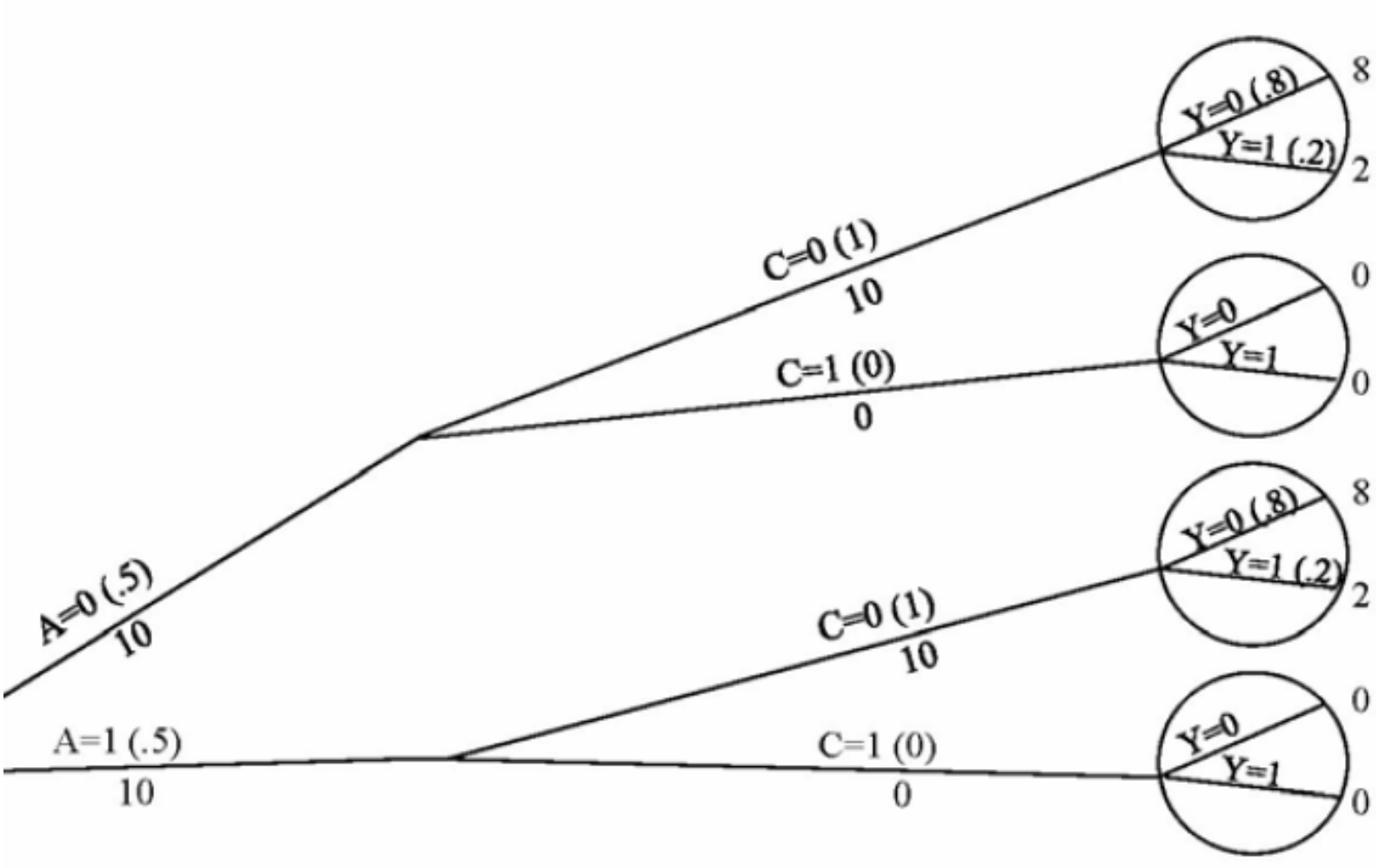


Stabilized weights create a pseudo-population with selection but no selection bias!

Adjusting for loss to follow-up



Adjusting for loss to follow-up

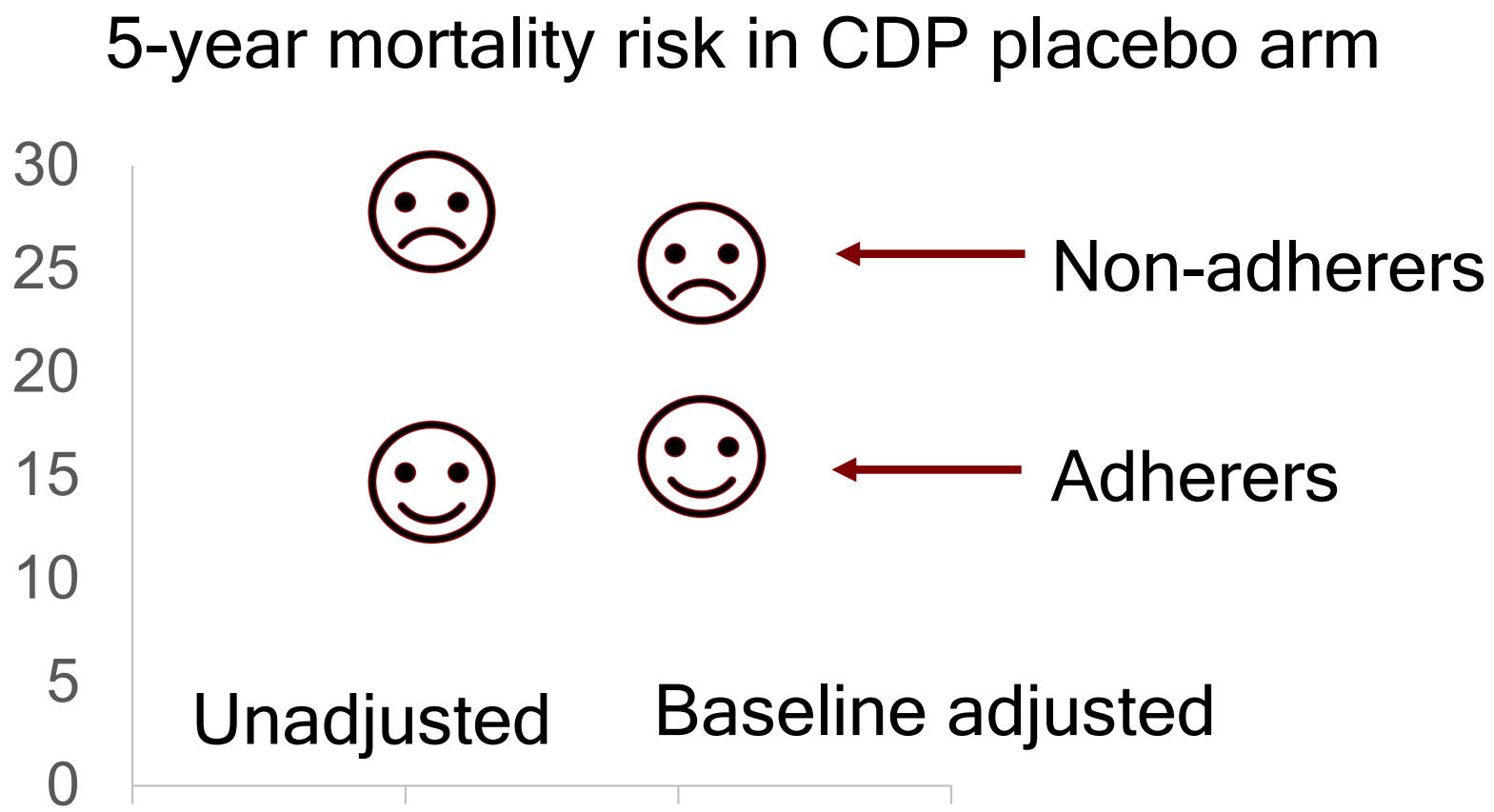


15 minute break

4. Exercise 3: Per-protocol effects

Reminder: Per-protocol analyses have a bad reputation!

But isn't adherence *intractably* confounded?



Coronary Drug Project. 1980, NEJM; 303: 1038-41.

What is a per-protocol analysis?

Common approaches

- censor when non-adherent, don't adjust for confounding
- add adherence to regression model, adjust only for baseline confounders

Common ≠ correct!

Per-protocol analyses tell us:

how did trial outcomes differ between those who did adhere to, or recieved, assignment A and those who did adhere to, or received, assignment B?

Per-protocol analyses in the literature

Approach	Description

Per-protocol analyses in the literature

Approach	Description
1. “Modified ITT”	▪ censor never initiators

Per-protocol analyses in the literature

Approach	Description
1. “Modified ITT”	▪ censor never initiators
2. “Per-protocol population”	▪ censor if never initiate, cross-over, or discontinuation

Per-protocol analyses in the literature

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor if never initiate, cross-over, or discontinuation
3. "As-treated"	▪ allow cross-over ▪ censor non-initiators or discontinuers

Methods 1 to 3: Censor without adjustment

Per-protocol analyses in the literature

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor at initiation, cross-over or discontinuation
3. "As-treated"	▪ allow cross-over ▪ censor non-initiators or discontinuers
4. Adherence adjustment	▪ include adherence in model for outcome model

Methods 1 to 3: Censor without adjustment

Method 4: Adjustment for baseline confounding only

Potential per-protocol analyses

Approach	Description
1. "Modified ITT"	▪ censor never initiators
2. "Per-protocol population"	▪ censor never initiators, cross-over or discontinuation
3. "As-treated"	▪ allow cross-over ▪ censor non-initiators or discontinuers
4. Adherence adjustment	▪ include adherence in model for outcome model
5. Instrumental variables, aka "contamination-adjusted ITT"	▪ compare outcome by trial arm, and correct using adherence by trial arm

Methods 1 to 3: Censor without adjustment

Method 4: Adjustment for baseline confounding only

Effects are different from analyses

Per-protocol effect tells us

> "how would trial outcomes differ if everyone adhered to assignment A versus if everyone adhered to assignment B"

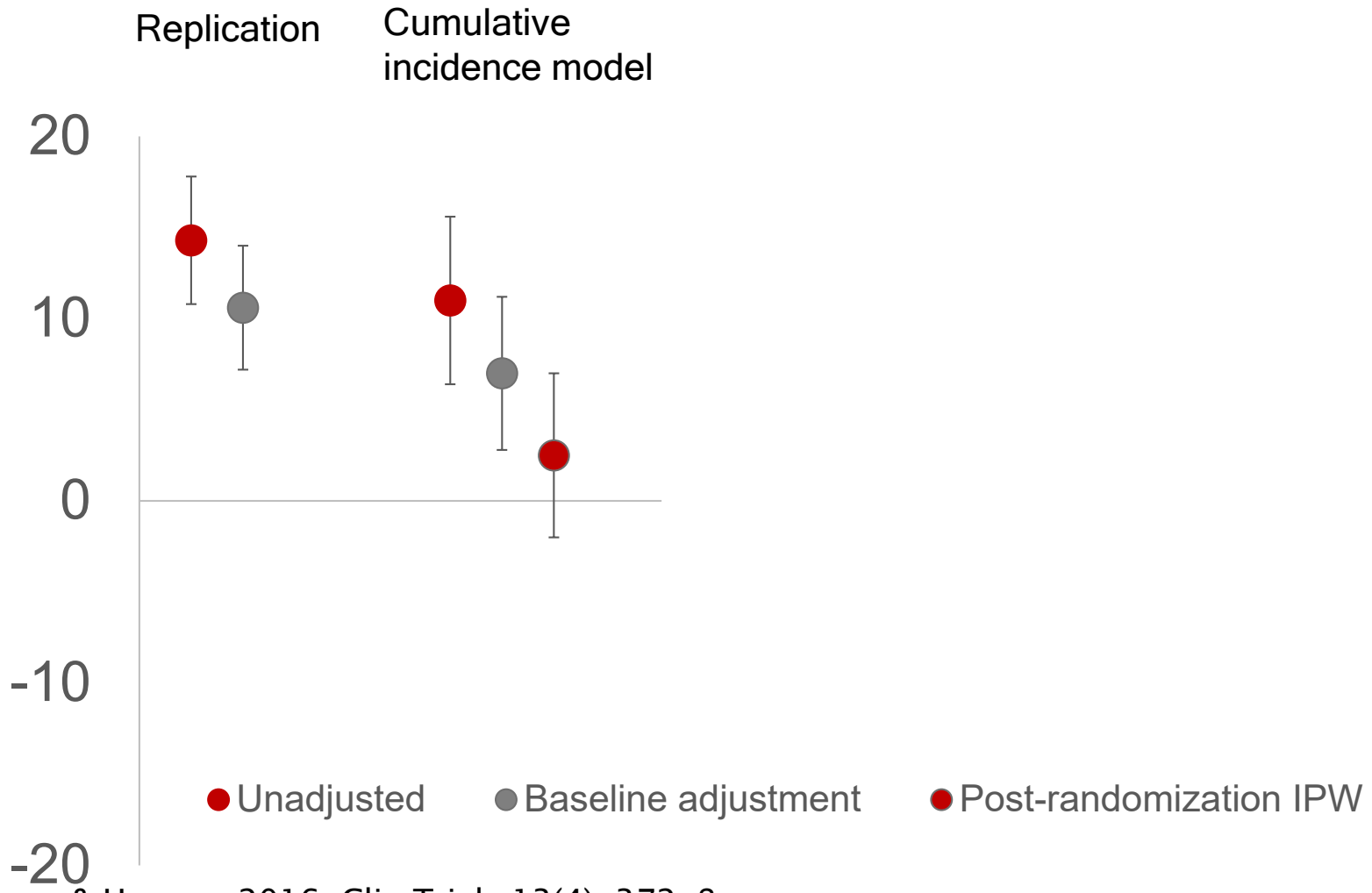
Better per-protocol analyses

Approach	Description
Per-protocol effect estimation	▪ censor if deviate from protocol or include adherence in outcome model ▪ adjust for censoring or time-varying confounding
5. Instrumental variables, ~~aka "contamination-adjusted ITT"~~	▪ compare outcome by trial arm, and correct using adherence by trial arm

Revisiting the Coronary Drug Project

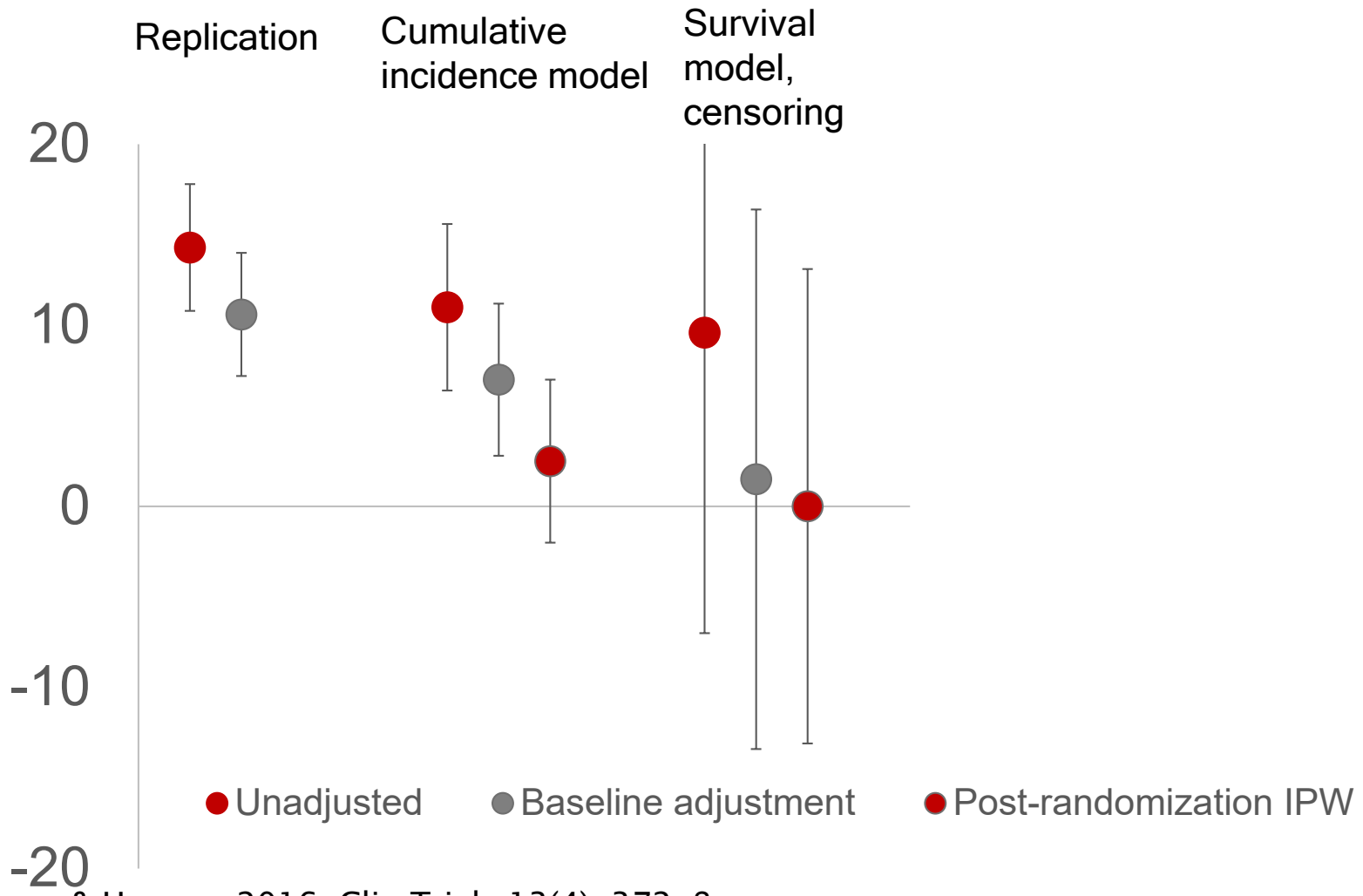
Murray & Hernan. 2016, Clin Trials 13(4): 372-8.
Murray & Hernan. 2018, Trials 19: 158.
Murray. 2019. Pragmatic Randomized Trials

Revisiting the Coronary Drug Project



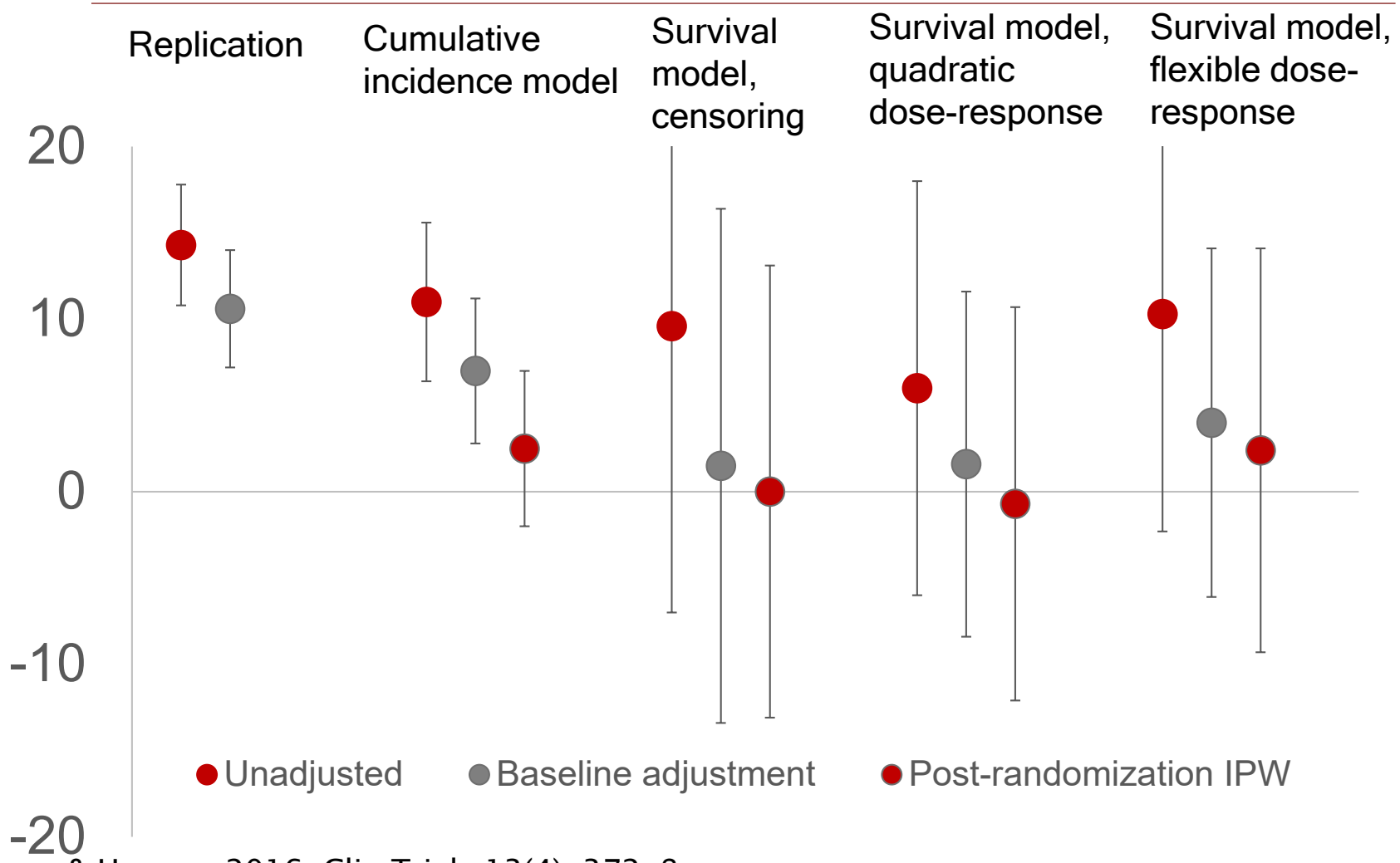
Murray & Hernan. 2016, Clin Trials 13(4): 372-8.
Murray & Hernan. 2018, Trials 19: 158.
Murray. 2019. Pragmatic Randomized Trials

Revisiting the Coronary Drug Project



Murray & Hernan. 2016, Clin Trials 13(4): 372-8.
Murray & Hernan. 2018, Trials 19: 158.
Murray. 2019. Pragmatic Randomized Trials

Revisiting the Coronary Drug Project



Murray & Hernan. 2016, Clin Trials 13(4): 372-8.
Murray & Hernan. 2018, Trials 19: 158.
Murray. 2019. Pragmatic Randomized Trials

In the handout, page 21:

4.2 Data Cleaning

In your preferred coding language, go to:

Exercise 3, Code Section 5 Data Cleaning for Exercise 3

Answer questions 1-2 on page 21

Inverse probability weighting

Step 1: identify adherent person-time

Step 2: build inverse probability weights for adherence and fit separately in each trial arm

Step 3: censor when non-adherent

Step 4: fit an IP-weighted model for the outcome given trial arm, baseline covariates

Step 5: generate estimated survival & risks standardized over baseline covariates

Inverse probability of adherence weights

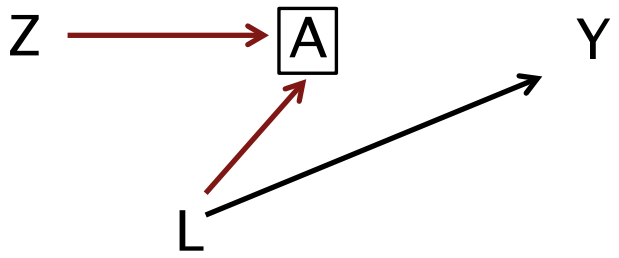
- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[A_j = a_j | Z, \bar{L}_j, \bar{A}_{j-1}]}$

- $SW_t = \prod_{j=0}^{t} \frac{Pr[A_j = a_j | Z, \bar{A}_{j-1}]}{\Pr[A_j = a_j | Z, \bar{L}_j, \bar{A}_{j-1}]}$

- At each time, each person receives a weight inversely proportional to the probability of the adherence pattern they have, conditional on randomization, time-varying covariates, and adherence history

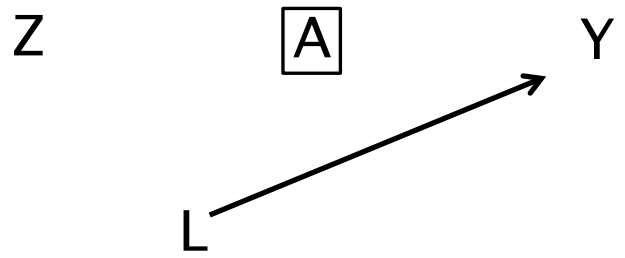
Adjusting for non-adherence

- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[A_j = a_j | Z, \bar{L}_j]}$



Adjusting for non-adherence

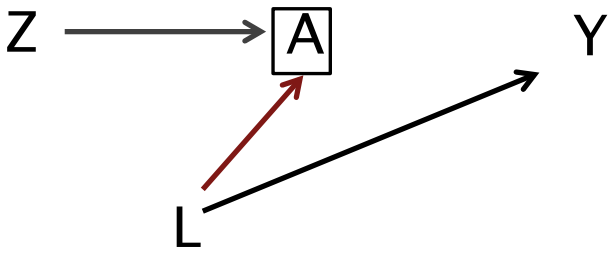
- $W_t = \prod_{j=0}^{t} \frac{1}{\Pr[A_j = a_j | Z, \bar{L}_j]}$



Non-stabilized weights create a pseudo-population with no adherence problems!

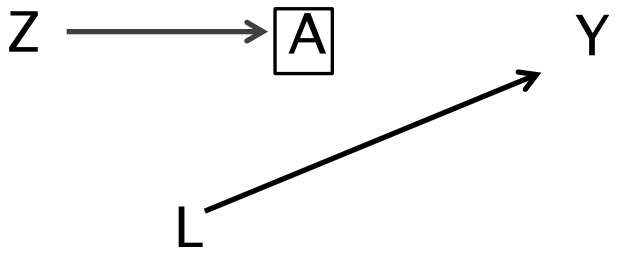
Adjusting for non-adherence

- $SW_t = \prod_{j=0}^{t} \frac{Pr[A_j = a_j | Z]}{\Pr[A_j = a_j | Z, \bar{L}_j]}$



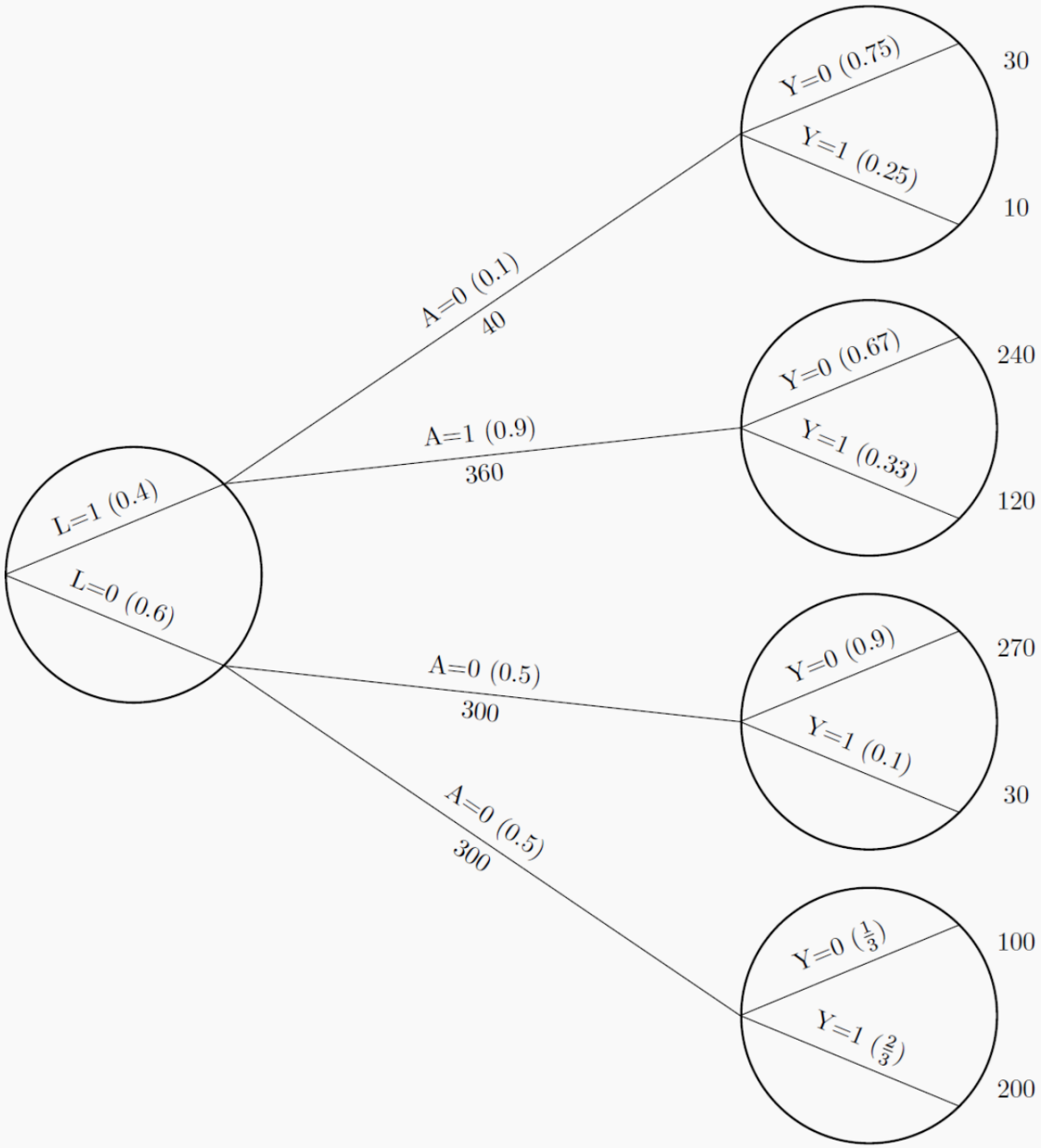
Adjusting for non-adherence

- $SW_t = \prod_{j=0}^{t} \frac{Pr[A_j = a_j | Z]}{\Pr[A_j = a_j | Z, \bar{L}_j]}$

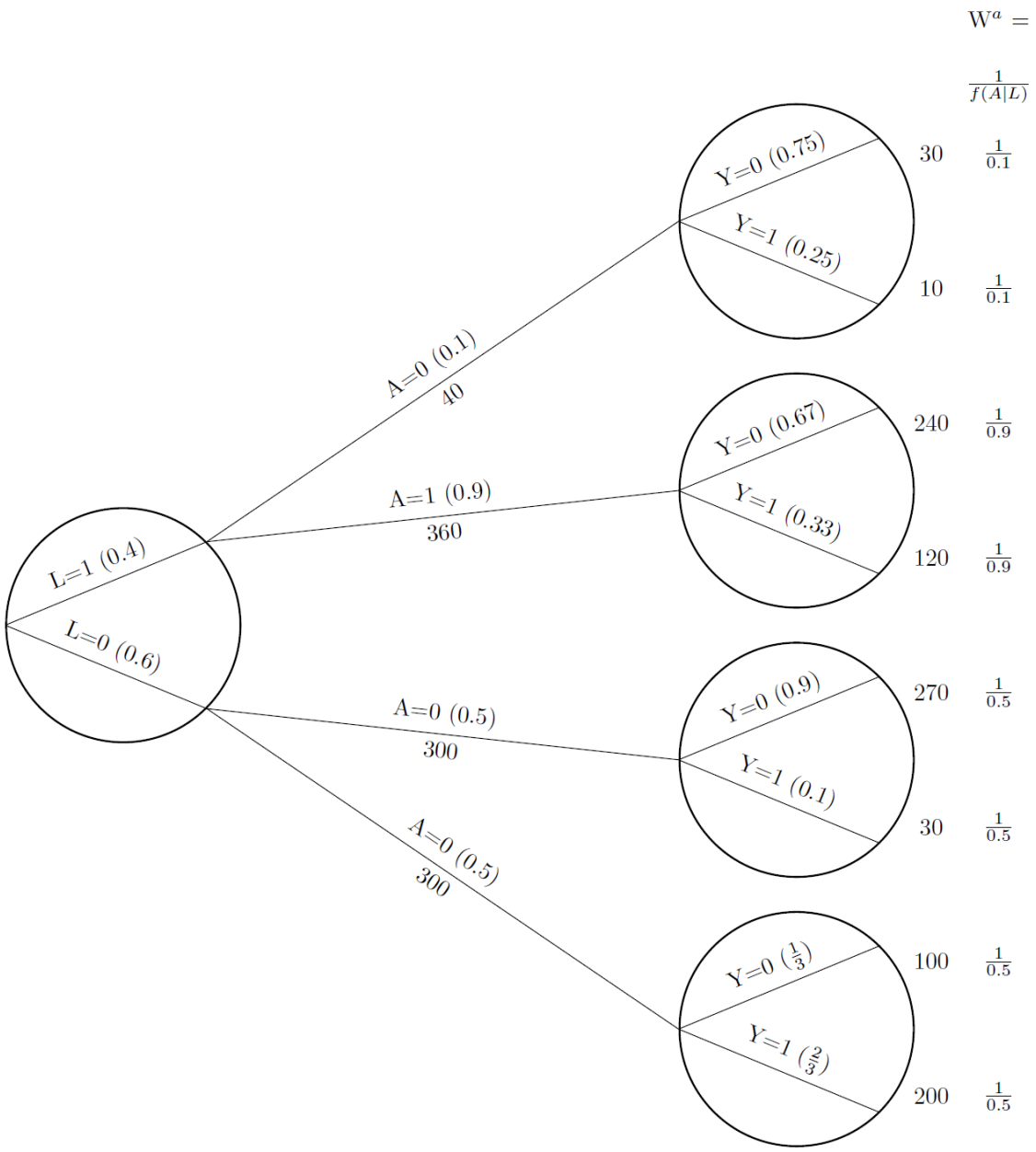


Stabilized weights create a pseudo-population with selection but no selection bias!

Probability tree:



Probability tree:



Estimating the per-protocol effect

Step 1: identify adherent person-time

Step 2: build inverse probability weights for adherence and fit separately in each trial arm

Step 3: censor when non-adherent

Step 4: fit an IP-weighted model for the outcome given trial arm, baseline covariates

Step 5: generate estimated survival & risks standardized over baseline covariates

In the handout, page 22:

4.3.1 Estimate inverse probability of adherence weights

In your preferred coding language, go to:
Exercise 3, Code Section 6 Weight Creation

Complete table on page 23

Answer question 1 on page 23

In the handout, page 24:
4.3.2 Estimate the conditional hazard ratio

In your preferred coding language, go to:
Exercise 3, Code Section 7 Weighted Conditional Hazard Ratios

Complete table on page 24
Answer questions 1-3 on page 24

In the handout, page 26:

4.3.3 Estimating the average survival curves

In your preferred coding language, go to:

Exercise 3, Code Section 8 Weighted Survival Curves

Complete table on page 26

Answer questions 1-4 on page 26-27

What about inference?

Estimators presented here are singly robust

- Doubly robust extensions exist

Robust standard errors ("sandwich" estimator)

- Provide overly conservative inference
- Assume weights are fixed, not estimated

Bootstrapping

- Allows for inference robust to the entire estimating procedure (weights included)

Wrap-up

What we learned

For point exposures:

- Non-parametric survival curves
- Unadjusted hazard ratios from semi-parametric and parametric models
- Conditional hazard ratios from semi-parametric and parametric models
- Average hazard ratios, survival curves, risk differences, and cumulative incidence ratios from parametric survival models

What we learned

For static sustained exposures:

- Non-parametric survival curves
- Unadjusted hazard ratios from semi-parametric and parametric models
- Conditional hazard ratios from semi-parametric and parametric models
- Average hazard ratios, survival curves, risk differences, and cumulative incidence ratios from parametric survival models

What about observational studies?

Observational studies have confounding

If we have a point exposure, we need to worry about baseline confounding

If we have a sustained exposure, we need to worry about baseline and time-varying confounding

We've learned how to handle both of these already!

Some more complicated scenarios you may encounter:

- Loss to follow-up and non-adherence – address this by multiplying inverse probability weights
- Dynamic sustained strategies – these typically require unstabilized weights
- Grace periods – whenever someone is following multiple strategies, clone them, and when you know which strategy they're following censor them & adjust with IPW
- Competing events – think carefully about the causal effect of interest

Where to get more information

Some references:

- Proposed pragmatic trial guidelines: https://www.hsph.harvard.edu/causal/pragmatictrials/
- Patient-centered causal effects: Murray et al. 2018. J Clin Epi 103:10-21.
- Choosing a causal effect: Hernan & Scharfstein. 2018, Ann Intern Med; 168(7):515-6.
- Per-protocol effect estimation: Hernan & Robins. 2017, NEJM 377:14; Lodi et al, 2016. AIDS; 30(17):2659-63.
- Placebo arm adherence analyses: Murray & Hernan. 2018, Trials 19:158; Murray & Hernan. 2016, Clin Trials 13(4): 372-8.
- G-methods: Causal Inference, Hernan & Robins. Available online at: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

Contact us:

@lucia_petito
lucia.petito@northwestern.edu

@EllieCaniglia
ellen.caniglia@nyulangone.org

THANK YOU