Recitation 4

Rachel Lee

September 29, 2021

What is causal inference?

Inferring the effects of any treatment/policy/intervention/etc.

Examples:

- Effect of treatment on a disease
- Effect of climate change policy on emissions
- Effect of social media on mental health
- Many more (effect of X on Y)

Motivating example: Simpson's paradox

Correlation does not imply causation

Then, what does imply causation?

Causation in observational studies

Simpson's paradox: COVID-27

New disease: COVID-27

Treatment T: A (0) and B (1)

Condition C: mild (0) or severe (1)

Outcome Y: alive (0) or dead (1)

Simpson's paradox: mortality rate table

Condition

Treatment	Mild	Severe
A	15% (210/1400)	30% (30/100)
B	**10%** (5/50)	**20%** (100/500)
	$\mathbb{E}[Y\|T, C = 0]$	$\mathbb{E}[Y\|T, C = 1]$

Simpson's paradox: mortality rate table

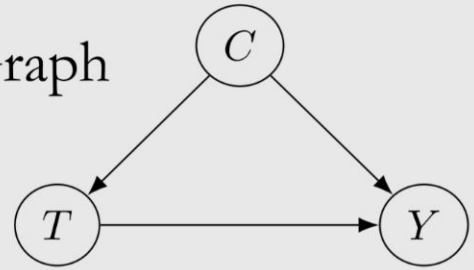
Condition

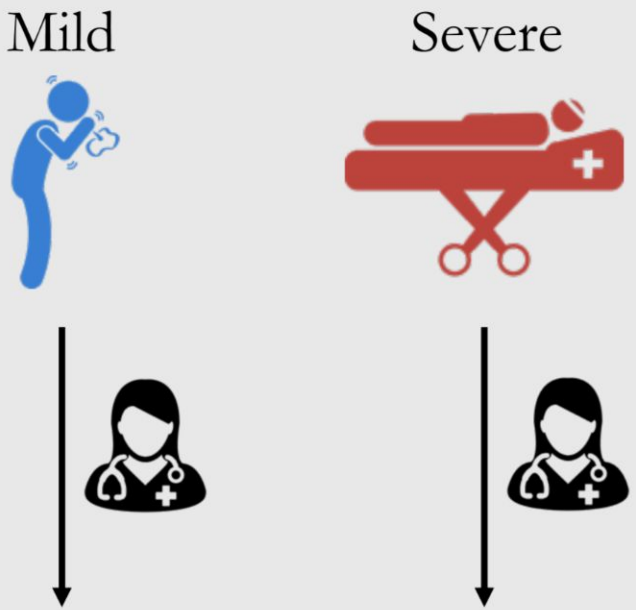
Treatment	Mild	Severe	Total
A	15% (210/1400)	30% (30/100)	**16%** (240/1500)
B	**10%** (5/50)	**20%** (100/500)	19% (105/550)
	$\mathbb{E}[Y\|T, C=0]$	$\mathbb{E}[Y\|T, C=1]$	$\mathbb{E}[Y\|T]$

Simpson's paradox: scenario 1 (treatment B)

Treatment	Condition: Mild	Condition: Severe	Condition: Total
A	15% (210/1400)	30% (30/100)	**16%** (240/1500)
B	**10%** (5/50)	**20%** (100/500)	19% (105/550)

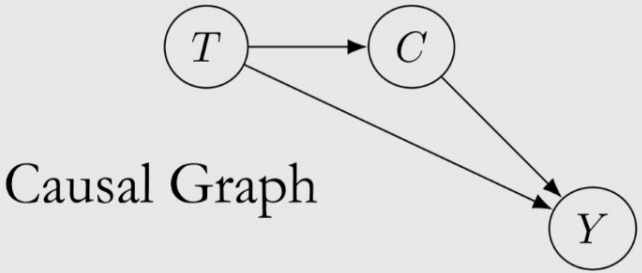
Causal Graph



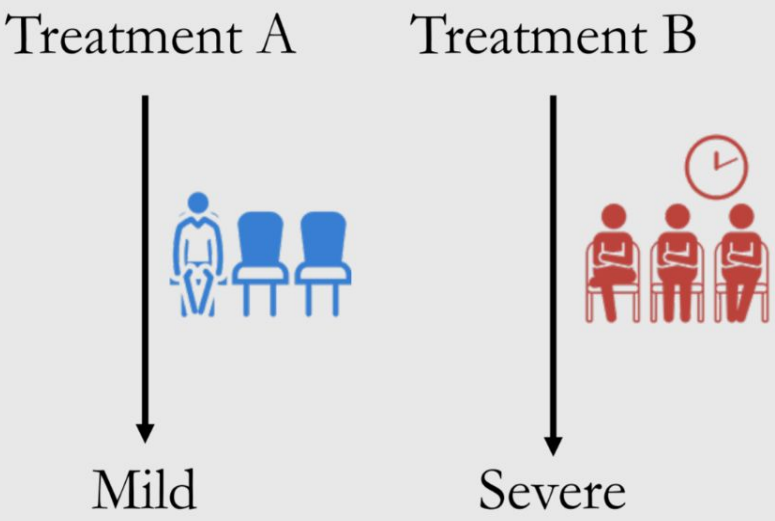


Simpson's paradox: scenario 2 (treatment A)

Treatment	Mild	Severe	Total
A	15% (210/1400)	30% (30/100)	**16%** (240/1500)
B	**10%** (5/50)	**20%** (100/500)	19% (105/550)

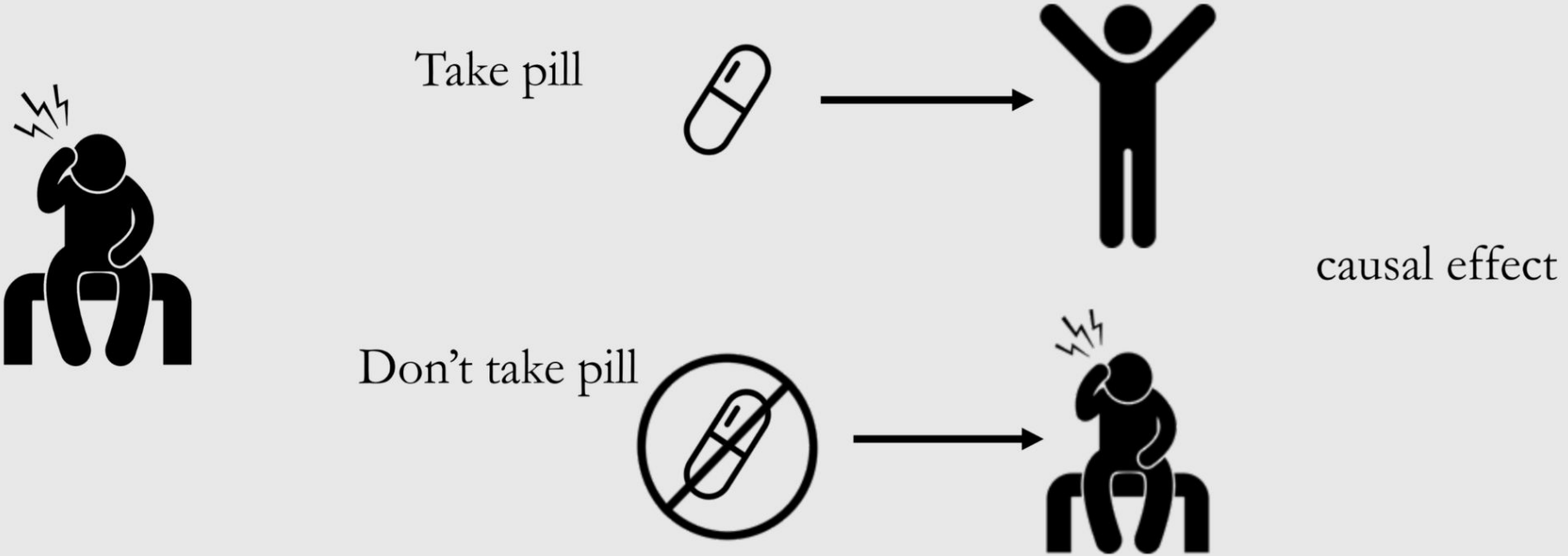


Causal Graph



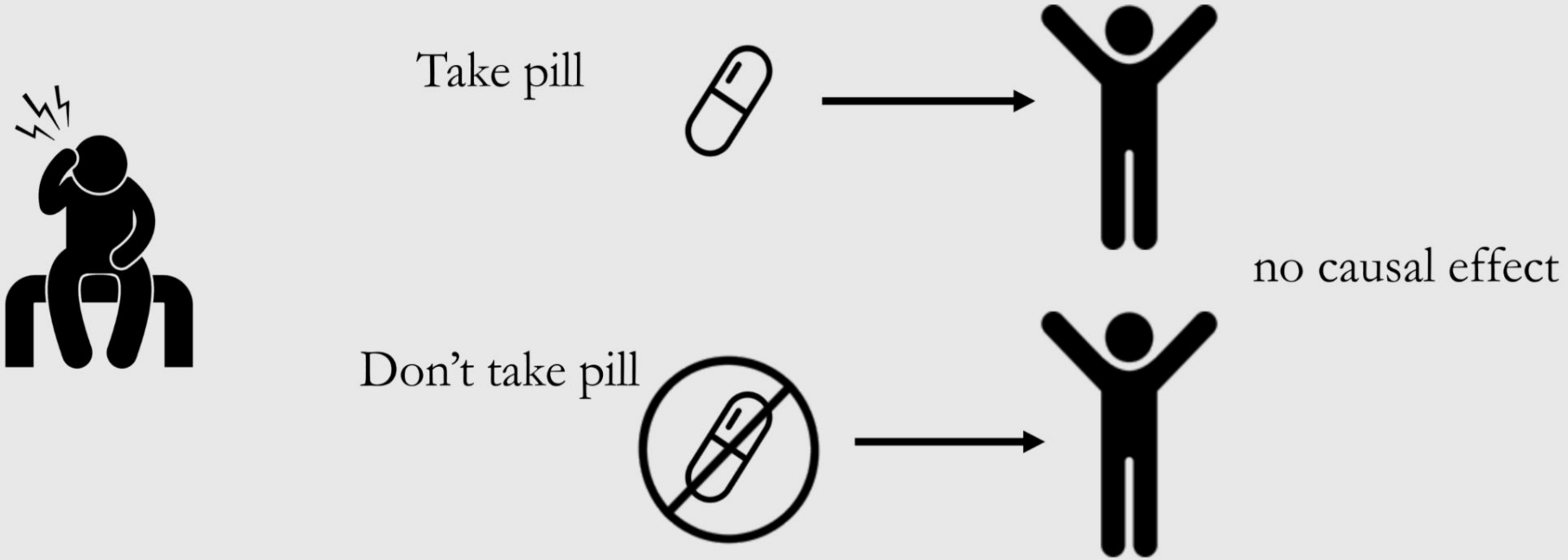
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome



Potential outcomes: intuition

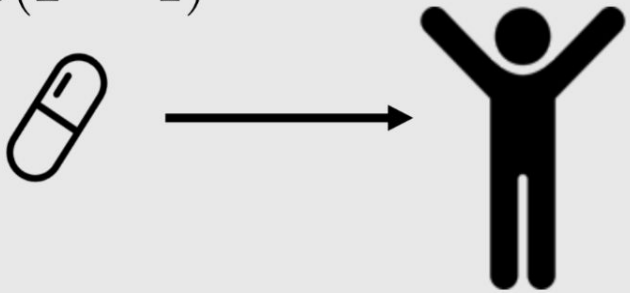
Inferring the effect of treatment/policy on some outcome



Potential outcomes: notation

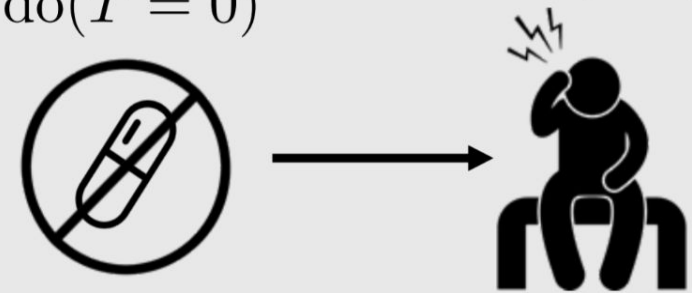
$\text{do}(T = 1)$

$Y_i|_{\text{do}(T=1)} \triangleq Y_i(1)$



$\text{do}(T = 0)$

$Y_i|_{\text{do}(T=0)} \triangleq Y_i(0)$



T	: observed treatment
Y	: observed outcome
i	: used in subscript to denote a specific unit/individual
$Y_i(1)$	: potential outcome under treatment
$Y_i(0)$	: potential outcome under no treatment

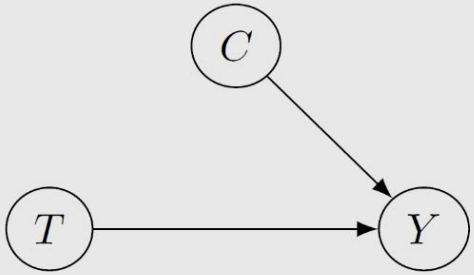
Causal effect

$$Y_i(1) - Y_i(0)$$

Observational studies

Can't always randomize treatment

- **Ethical reasons** (e.g. unethical to randomize people to smoke for measuring effect on lung cancer)
- **Infeasibility** (e.g. can't randomize countries into communist/capitalist systems to measure effect on GDP)
- **Impossibility** (e.g. can't change a living person's DNA at birth for measuring effect on breast cancer)



Formulating your hypothesis

Formulating your hypothesis