Causal Inference for Health Data

(STATS C160/C260 — Winter 2026)

Lecture 19: Causal Artificial Intelligence for Health Data

Drago Plečko

UCLA College | Physical Sciences
Statistics & Data Science

Medicine before the Data Revolution

Disease severity: low

	Survived	Not Survived
Treated	*20*	*10*
Untreated	*35*	*35*

Disease severity: high

	Survived	Not Survived
Treated	*25*	*25*
Untreated	*20*	*30*

- Data difficult to collect, analyze,
- Basic questions considered, such as treatment effects,
- Data types constrained, such as count data.

Medicine today (post-Data Revolution)

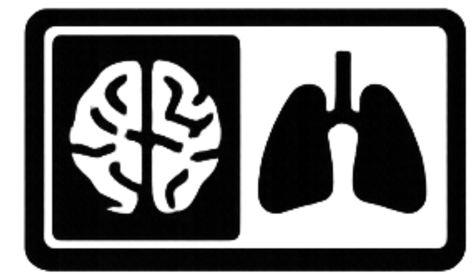
Disease severity: low

	Survived	Not Survived
Treated	*20*	*10*
Untreated	*35*	*35*

Disease severity: high

	Survived	Not Survived
Treated	*25*	*25*
Untreated	*20*	*30*

individual EHR data



radiology imaging

patient text notes

waveform data

omics

...

Medicine today (post-Data Revolution)

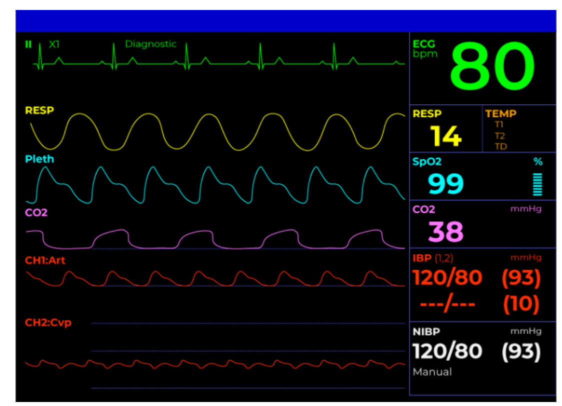
Our discussion so far

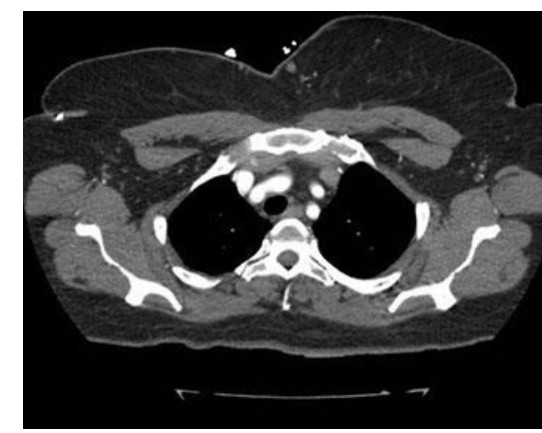
Clinical Reality Today

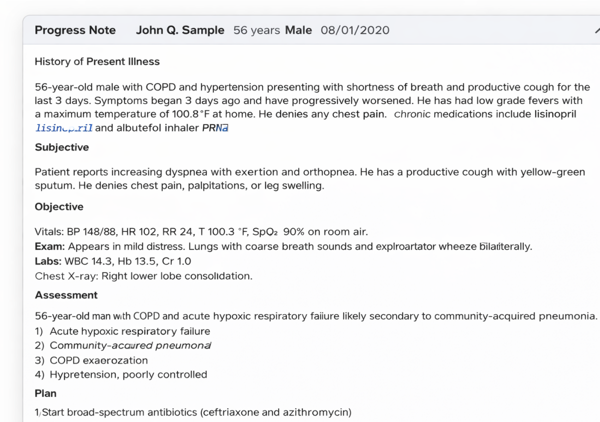
Z (confounders)	X (treatment)	Y (outcome)
(0.3, 0.7)	*0*	*1*
…	…	…
(0.3, 0.7)	*1*	*1*

Question: what is the treatment effect?

For reaching the decision, physicians process all kinds of information:






Progress Note **John Q. Sample** 56 years **Male** 08/01/2020

History of Present Illness

56-year-old male with COPD and hypertension presenting with shortness of breath and productive cough for the last 3 days. Symptoms began 3 days ago and have progressively worsened. He has had low grade fevers with a maximum temperature of 100.8°F at home. He denies any chest pain. *chronic* medications include lisinopril lisinopril and albutefol inhaler *PRN*

Subjective

Patient reports increasing dyspnea with exertion and orthopnea. He has a productive cough with yellow-green sputum. He denies chest pain, palpitations, or leg swelling.

Objective

Vitals: BP 148/88, HR 102, RR 24, T 100.3 °F, SpO_2 90% on room air.
Exam: Appears in mild distress. Lungs with coarse breath sounds and explroartator wheeze bilaiterally.
Labs: WBC 14.3, Hb 13.5, Cr 1.0
Chest X-ray: Right lower lobe consolidation.

Assessment

56-year-old man with COPD and acute hypoxic respiratory faiiure likely secondary to community-acquired pneumonia.
1) Acute hypoxic respiratory failure
2) Community-*acqured pneumona*
3) COPD exaerozation
4) Hypretension, poorly controlled

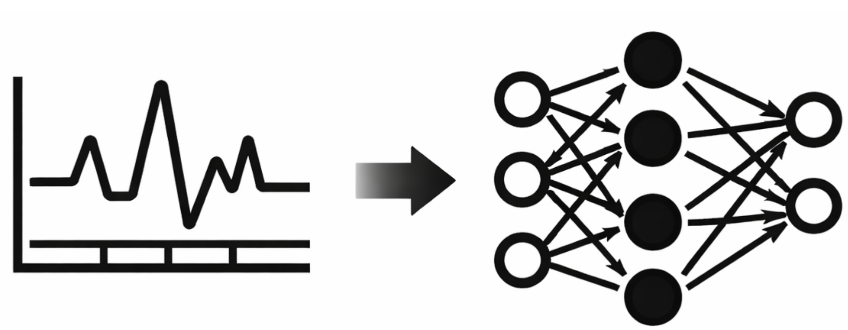
Plan

1 Start broad-spectrum antibiotics (ceftriaxone and azithromycin)
2 Initiate supplemental oxygen via nasal cannula

Treatment Decision

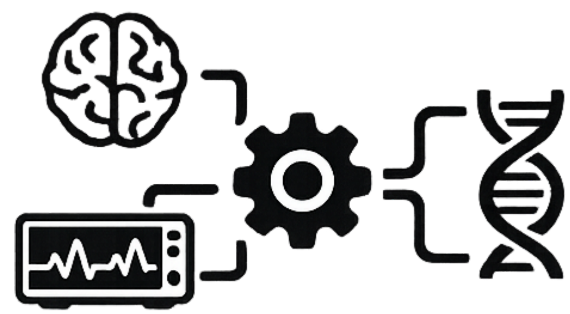
AI Tools & Their Use-Cases

Predictive Analytics



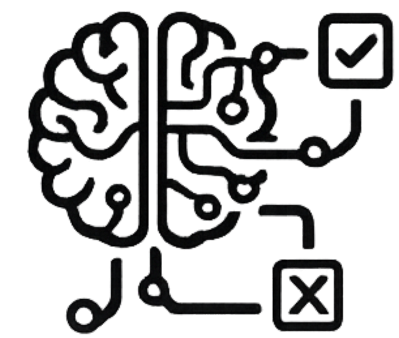
- We have access to observational data, which may include complex data such as text, images, etc.
- The query of interest lives in Layer 1 of the PCH,
- Causal Inference not needed.

Multi-Modal Causal Inference



- Complex data involved, but our query of interest is either Layer 2 or Layer 3 of the PCH,
- The key challenge is on how to adapt our identification & estimation techniques for complex data (e.g., what does adjusting for an image mean?)

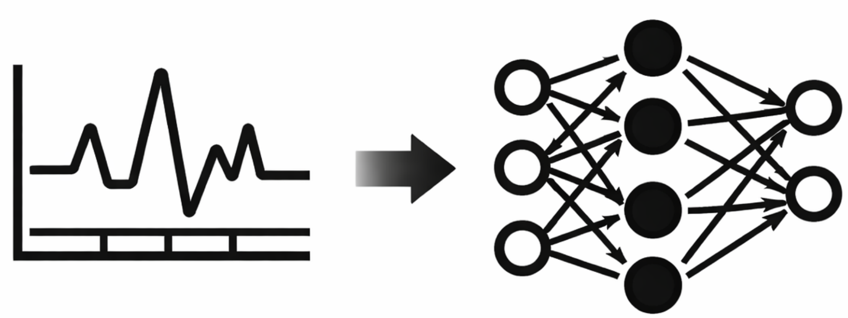
AI Decision-Making



- In Decision-Making, our goal is to build a system which is able to make decisions entirely on its own,
- Such decisions may be based on treatment effect estimates (Layer 2/3 queries), but may also consider other downstream effects.

AI Tools & Their Use-Cases

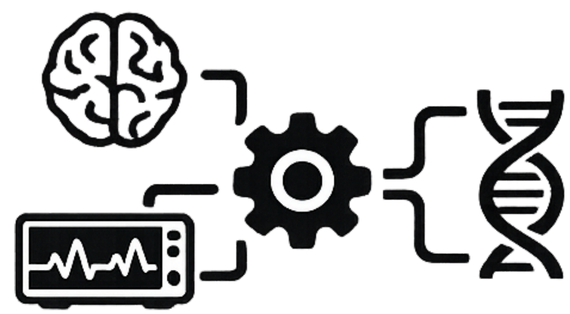
Predictive Analytics



- We have access t[illegible] observational [illegible]hich may include [illegible]ex data such as t[illegible]ages, etc.
- The q[illegible] interest lives in L[illegible] of the PCH,
- C[illegible] inference not nee[illegible]ed.

Non-Causal

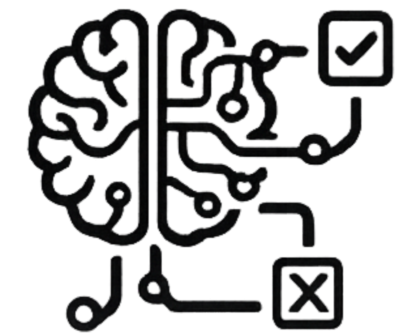
Multi-Modal Causal Inference



- Complex data inv[illegible]lved, but our query o[illegible]est is either Lay[illegible] Layer 3 of the P[illegible]
- The ke[illegible]enge is on how[illegible]pt our id[illegible]ation & esti[illegible]ation techniques for complex data (e.g., what does adjusting for an image mean?)

Causal

AI Decision-Making



- In Decision-Maki[illegible], our goal is to build [illegible]tem which is abl[illegible]ke decisions [illegible] on its own,
- Suc[illegible]ions may be b[illegible] treatment effe[illegible] estimates (Layer 2/3 queries), but may also consider other downstream effects.

Causal

Predictive Analytics Examples

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > articles > article

Article | Open access | Published: 22 May 2024

A whole-slide foundation model for digital pathology from real-world data

Hanwen Xu, Naoto Usuyama, Jaspreet Bagga, Sheng Zhang, Rajesh Rao, Tristan Naumann, Cliff Wong, Zelalem Gero, Javier González, Yu Gu, Yanbo Xu, Mu Wei, Wenhui Wang, Shuming Ma, Furu Wei, Jianwei Yang, Chunyuan Li, Jianfeng Gao, Jaylen Rosemon, Tucker Bower, Soohee Lee, Roshanthi Weerasinghe, Bill J. Wright, Ari Robicsek, … Hoifung Poon + Show authors

Nature **630**, 181–188 (2024) | Cite this article

Predictive Analytics Examples

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > articles > article

Article

A wh... gy from...

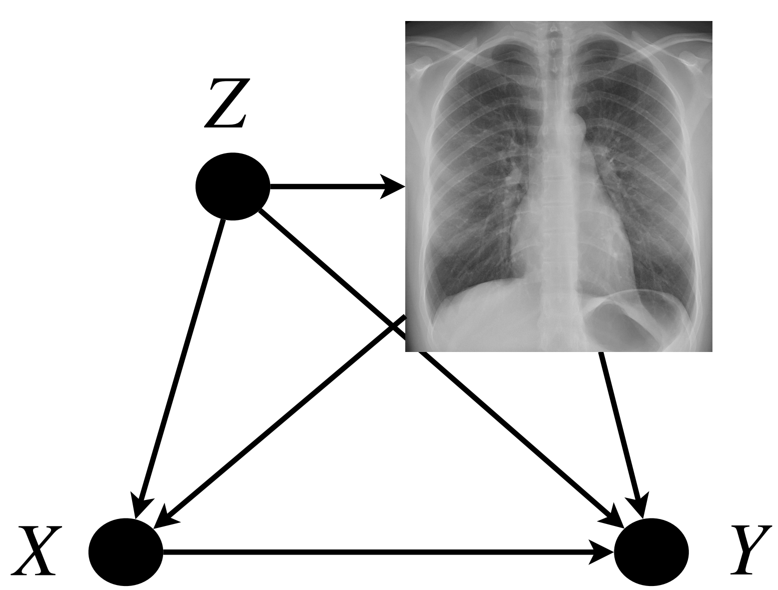
However, as this setting is not causal, we do not discuss in detail in this course

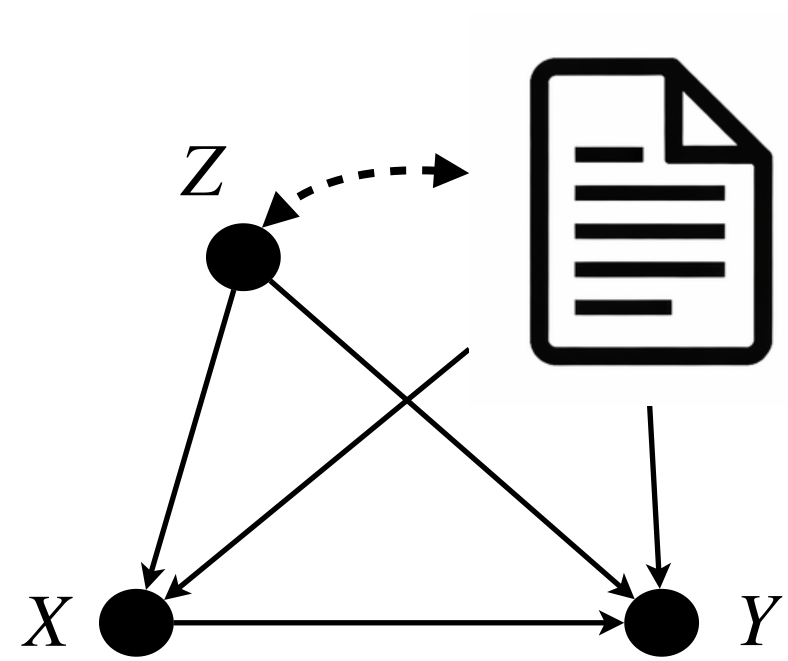
Hanwen Xu, Naoto Usuyama, Jaspreet Bagga, Sheng Zhang, Rajesh Rao, Tristan Naumann, Cliff Wong, Zelalem Gero, Javier González, Yu Gu, Yanbo Xu, Mu Wei, Wenhui Wang, Shuming Ma, Furu Wei, Jianwei Yang, Chunyuan Li, Jianfeng Gao, Jaylen Rosemon, Tucker Bower, Soohee Lee, Roshanthi Weerasinghe, Bill J. Wright, Ari Robicsek, ... Hoifung Poon + Show authors

Nature **630**, 181–188 (2024) | Cite this article

Multi-Modal Causal Inference Settings

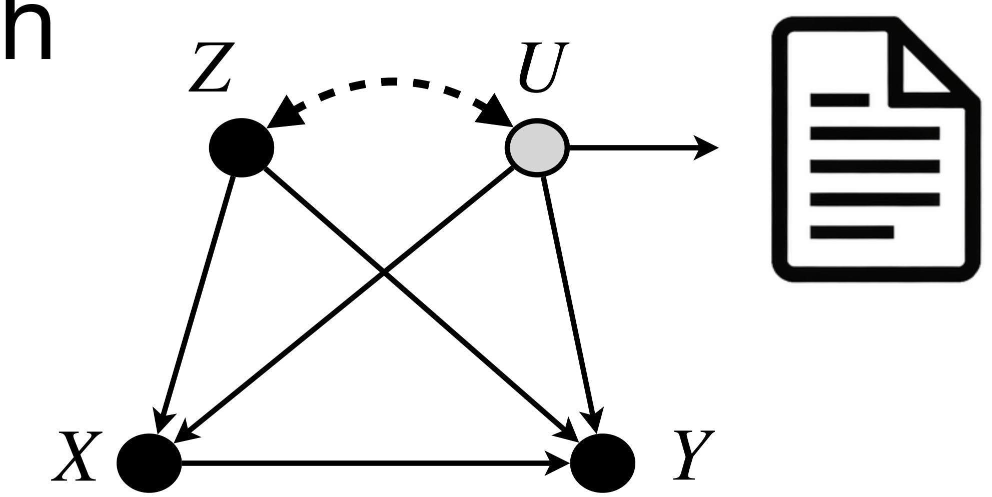
- Under multi-modal causal inference, we include all settings where data is of multiple modalities,
- This commonly includes tabular data in addition with images or text data, e.g.,





Multi-Modal CI: Task 1 — Known Confounder Recovery

- Suppose we have a setting with observed confounders Z, and textual data T, treatment X, outcome Y



- Further, suppose we know that the text contains known confounders $U_1, \ldots, U_k$ which we wish to extract from the task.

Input:

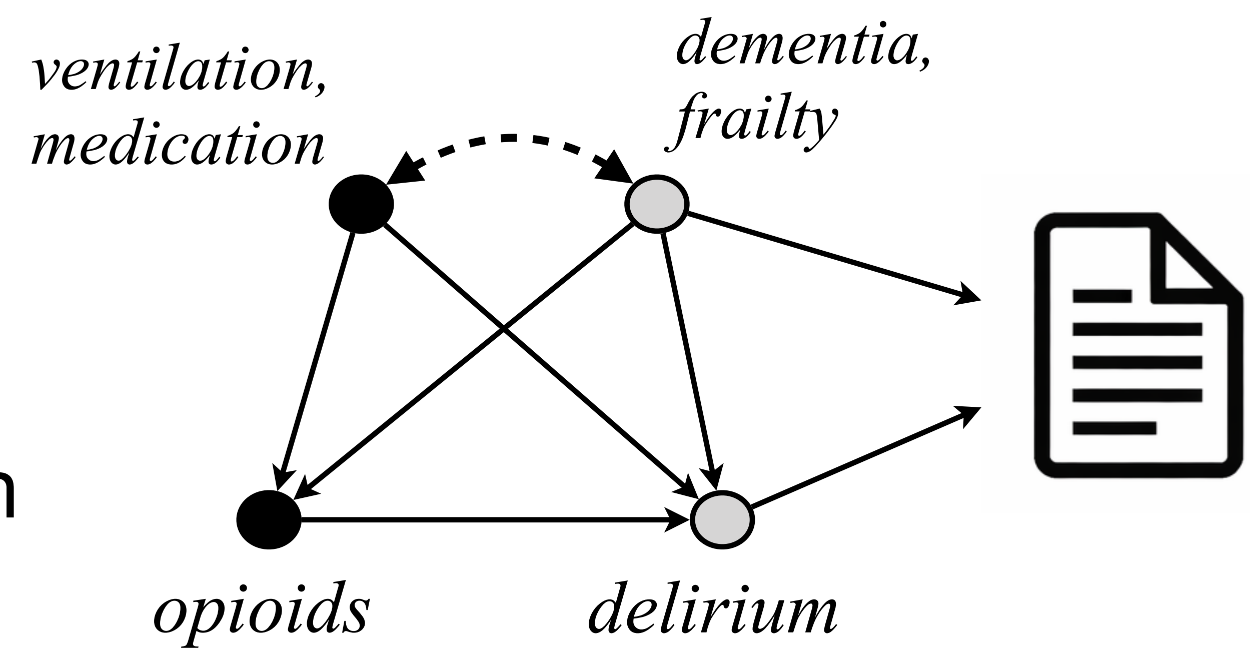
- Data (Z, X, Y)
- Text T
- Known, but unavailable confounders U

Output:

$$P(y \mid do(x))$$

Multi-Modal CI: Task 1 — Delirium in Intensive Care

- We are interested in analyzing the effect of opiods on delirium in intensive care units,
- Delirium is a serious change in mental abilities, resulting in confused thinking and a lack of awareness of someone's surroundings,

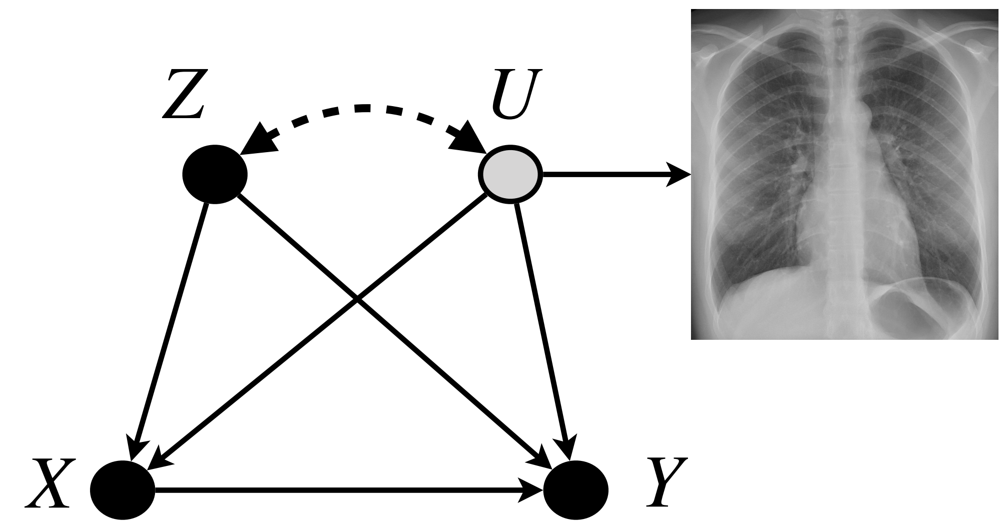


Goal: annotate delirium, dementia, frailty from text

- However, delirium is rarely explicitly coded in EHR data, but may be available in text notes.
- Confounders include ventilation, other medication (propofol, benzodiazepines), but also dementia/frailty (coded in text).

Multi-Modal CI: Task 2 — Unknown Confounder Adjustment

- Suppose now that the unobserved U variables are not known, but are assumed to be available in image data I,



- Further, we do not want to recover U, but just want to adjust for it — how do we find a representation $f(I)$ that allows for such adjustment?

Input:

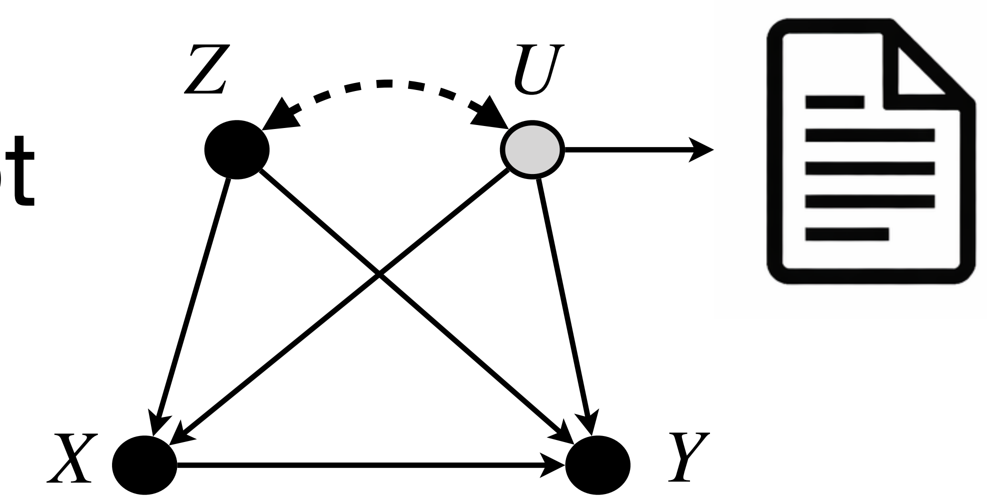
- Data (Z, X, Y)
- Text T
- Unknown and unavailable confounders U

Output:

$$P(y \mid do(x))$$

Multi-Modal CI: Task 3 — Unknown Confounder Recovery

- Suppose again that the unobserved U variables are not known, but are assumed to be available in text data T,



- Is there a way for us to recover the correct Us?

Input:

- Data (Z, X, Y)
- Text T
- Unknown and unavailable confounders U

Output:

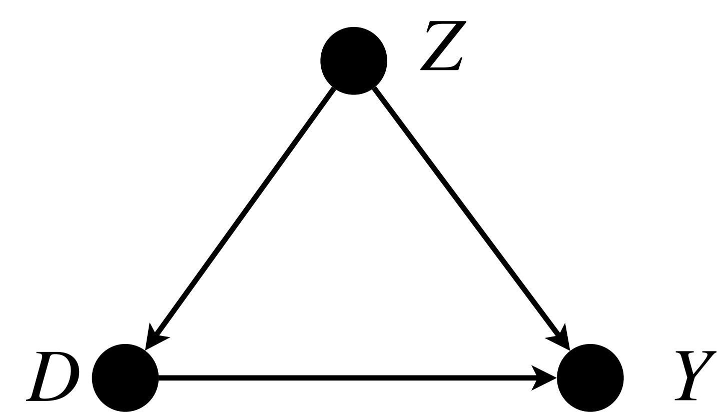
- list of confounders U

AI for Automated Decision-Making

- Treatment effect estimation is an important task, and AI tools allow treatment personalization based on known covariates,
- Generally, we may be interested in fully automated systems, which estimate treatment effects and then allocate the decision,
- Examples include glucose control, sepsis management, ventilator management, anesthesia depth, radiotherapy, …

Example: Allocation of Respirators

- We want to build an AI system in the ICU which automatically decides when to apply mechanical ventilation (D),
- Our goal is to optimize patients respiratory stability (Y),



- We are designing a treatment policy, that is, based on $Z = z$, we need to decide if $D = 0$ (don't treat) or $D = 1$ (treat).
- How to we make an optimal policy?

Optimal Decision-Making: Role of the CATE

- How is the problem solved? Expand the objective:

$$E[Y_D] = P(Y_D = 1)$$

$$= \sum_{x,z,w} P(Y_D = 1 \mid x, z, w) P(x, z, w) \qquad \textbf{(Law of Tot. Prob.)}$$

(Ignorability)

$$= \sum_{x,z,w} \Big[\underbrace{P(y_{d_1} \mid x, z, w) P(D = 1 \mid x, z, w)}_{\text{treated}} + \underbrace{P(y_{d_0} \mid x, z, w) P(D = 0 \mid x, z, w)}_{\text{untreated}} \Big] \cdot P(x, z, w),$$

and by using the fact that $P(D = 0 \mid x, z, w) \overset{(*)}{=} 1 - P(D = 1 \mid x, z, w)$, and defining $\Delta := E[Y_{d_1} - Y_{d_0} \mid x, z, w]$ (known as CATE or *benefit*), we get:

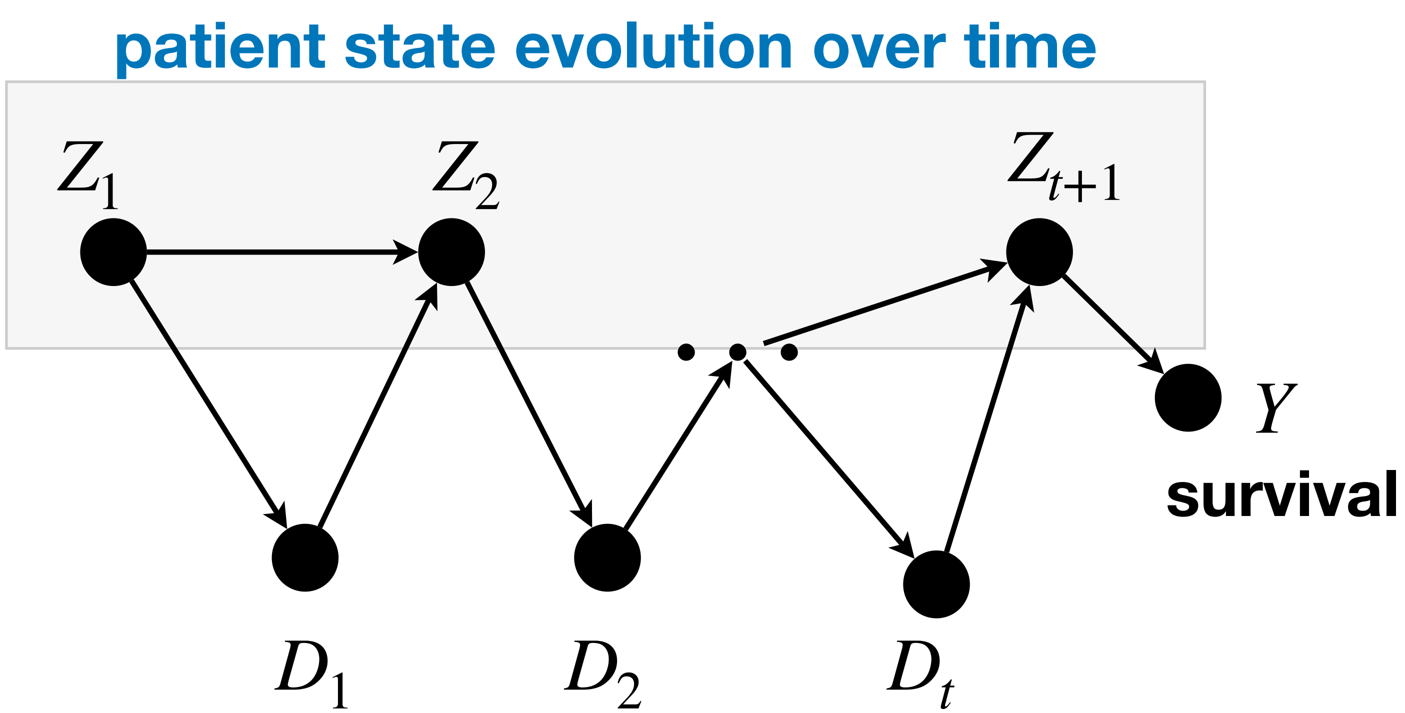
$$= P(y_{d_0} = 1) + \sum_{x,z,w} \boxed{P(D = 1 \mid x, z, w)} P(x, z, w) \Delta(x, z, w)\,.$$

Our choice

To optimize, treat those with highest Δ values

Multi-Step Decision-Making

- In clinical settings, decisions may be over multiple steps,
- For instance, consider the setting of radiotherapy for cancer: a radiation oncologist may treat a patient with different doses over multiple time steps, adapting the treatment to the patient's response:



Goal:

$$\pi = \operatorname{argmax} E_\pi[Y]$$

finding a policy that maximizes survival

Multi-Step Decision-Making

- In step t usually characterized by patient state Z_t, action A_t, and the reward in the step R_t,
- In our setting, the reward is only obtained in the end, if the patient survives; otherwise reward = 0,
- We define $v(z)$ as the optimal reward achieved for state z,
- Optimal solution governed by the Bellman equation:

$$v(z) := \max_d E[\underbrace{r(z, d, Z_{+1})}_{\text{expected reward of decision}} + \underbrace{\gamma v(Z_{+1})}_{\text{future reward (with discounting)}} \mid \underbrace{do(d), z}_{\text{in do(d), see(z) regime}}]$$

Multi-Step Decision-Making

- In step t usually characterized by patient state Z_t, action A_t, ar
- In ou d, if the patie
- We define $v(z)$ as the optimal reward achieved for state z,
- Optim :

In AI/ML, a whole field is focused on solving such problems: Reinforcement Learning

$v(z$]

RL problems are causal problems.

expected reward of decision

future reward (with discounting)

in do(d), see(z) regime

Dangers of RL in Observational Data

nature medicine

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature medicine › articles › article

Article | Published: 22 October 2018

The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care

Matthieu Komorowski, Leo A. Celi, Omar Badawi, Anthony C. Gordon & A. Aldo Faisal

Nature Medicine **24**, 1716–1720 (2018) | Cite this article

al

is not recorded perfectly— what happens?

Dangers of RL in Observational Data

nature medicine

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature

The algorithm falsely learns that avoiding treatment leads to better outcomes.

val

Article

The … al treatment strategies for sepsis in intensive care

Matthieu Komorowski, Leo A. Celi, Omar Badawi, Anthony C. Gordon & A. Aldo Faisal

Nature Medicine **24**, 1716–1720 (2018) | Cite this article

is not recorded perfectly— what happens?

Trustworthiness of AI Systems

- Suppose that using AI we can find a useful or optimal decision policy for a specific setting,
- What other challenges exist when deploying such a decision-making policy in practice?
- An important aspect of AI adoption is its <u>trustworthiness</u>:

Robustness | **Explainability** | **Fairness**

Robustness (Example)

- An endocrinologist has access to data on insulin administrations and blood glucose levels of her patients,

- She knows that hyperglycemia is bad, and sets a negative reward of -1 for it;
- further, she knows that hypoglycemia is bad, and thus also sets a negative reward of -1,
- She trains an AI algorithm to administer insulin and glucose for her patients, all of whom are adults.

Robustness (Example)

- An AI startup company starts deploying the algorithm, and sells it to different institutions,
- In one of the institutions, the patients are children instead of adults,
- For children the harm of hypoglycemia is far worse,
- The optimal algorithm should behave differently in this population, and may cause serious harm.

Robustness (Example)

- An AI startup company starts deploying the alg
- In pa adu
- For children t ia is far worse,
- The optimal a differently in this populatio us harm.

A human decision-maker would be able to respond adaptively, while an AI system may fail on unseen data.

Explanations (Example)

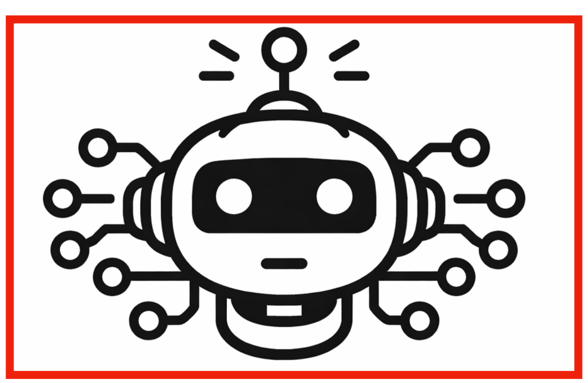
- A group of doctors in a hospital trains an AI system to decide whether to administer antibiotics to improve outcomes,
- Among the information provided to the system is treatment X, confounders Z, outcome Y (mortality), and other treatment information T,

- In this hospital, only patients with serious infections have their body fluids (culture) sampled, meaning that sampling S correlates with benefit from treatment.

Explanations (Example)

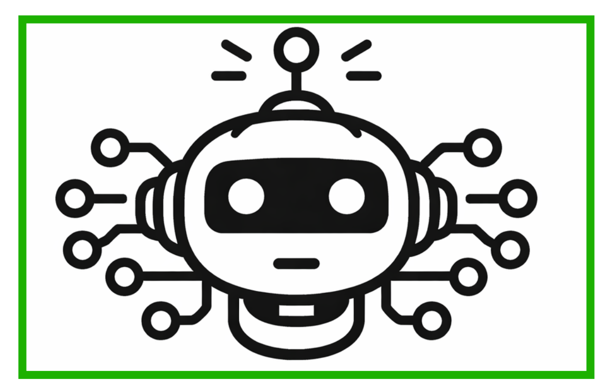
- The AI algorithm is then applied to another hospital in a different country,
- In this hospital, body fluid sampling is performed routinely, for every patient,
- As a consequence, antibiotics get oversubscribed,

- This is a robustness failure, but also a failure of explainability.



I made the decision to administer antibiotics

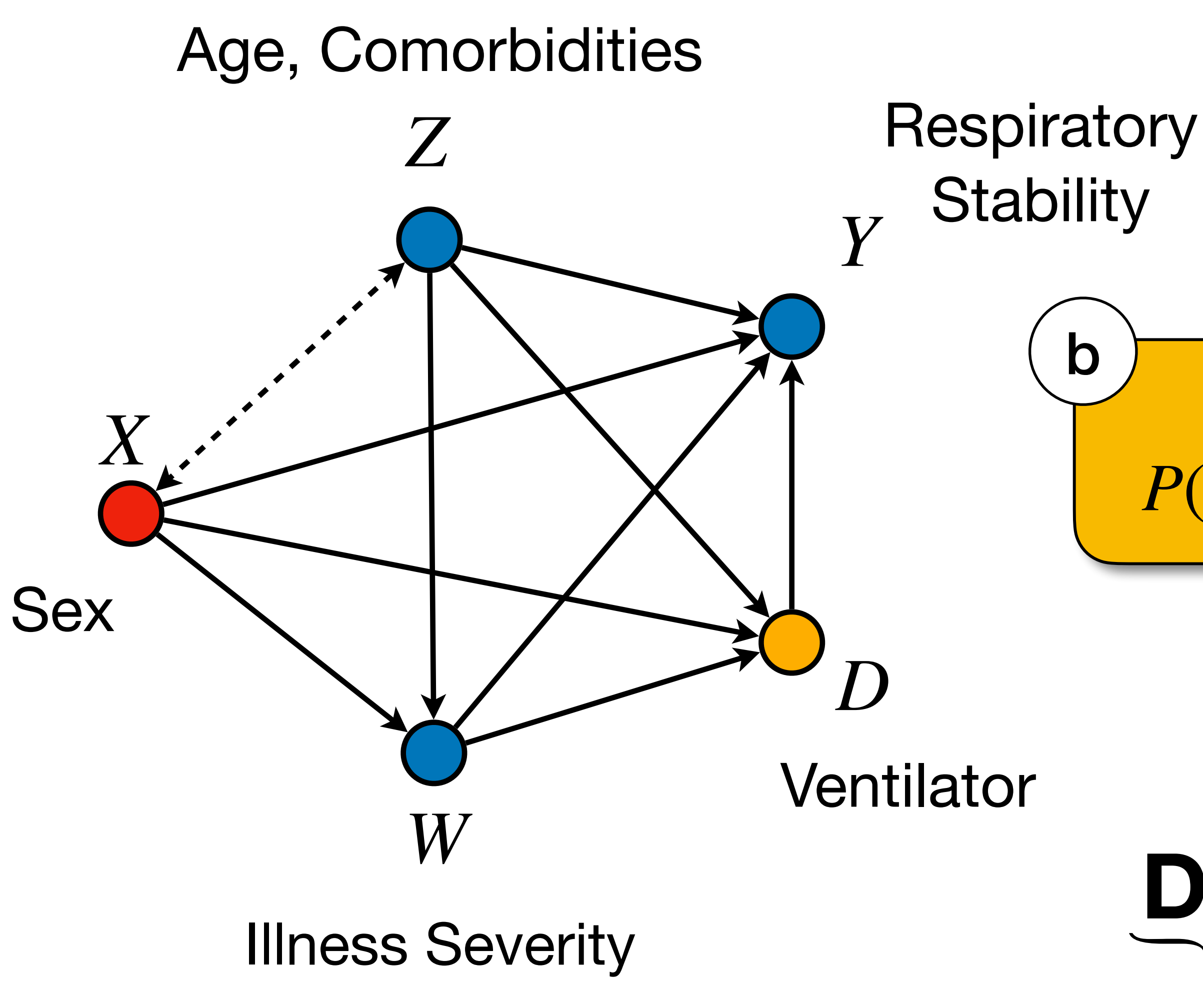
which is better? which makes it easier to avoid failure?



I made the decision to administer antibiotics *because blood culture was sampled*

Fairness (Example): Allocating Respirators

$E[\text{Respirator} \mid \ ] - E[\text{Respirator} \mid \ ] > 0.$



a

Decision based on

$\Delta = E[Y_{d_1} - Y_{d_0} \mid x, z, w]$

EP

$P(D = 1 \mid \Delta, \ ) = P(D = 1 \mid \Delta, \ )$

b

What if

$P(\Delta \mid x_0) \neq P(\Delta \mid x_1)$**?**

c

Perform a causal explanation

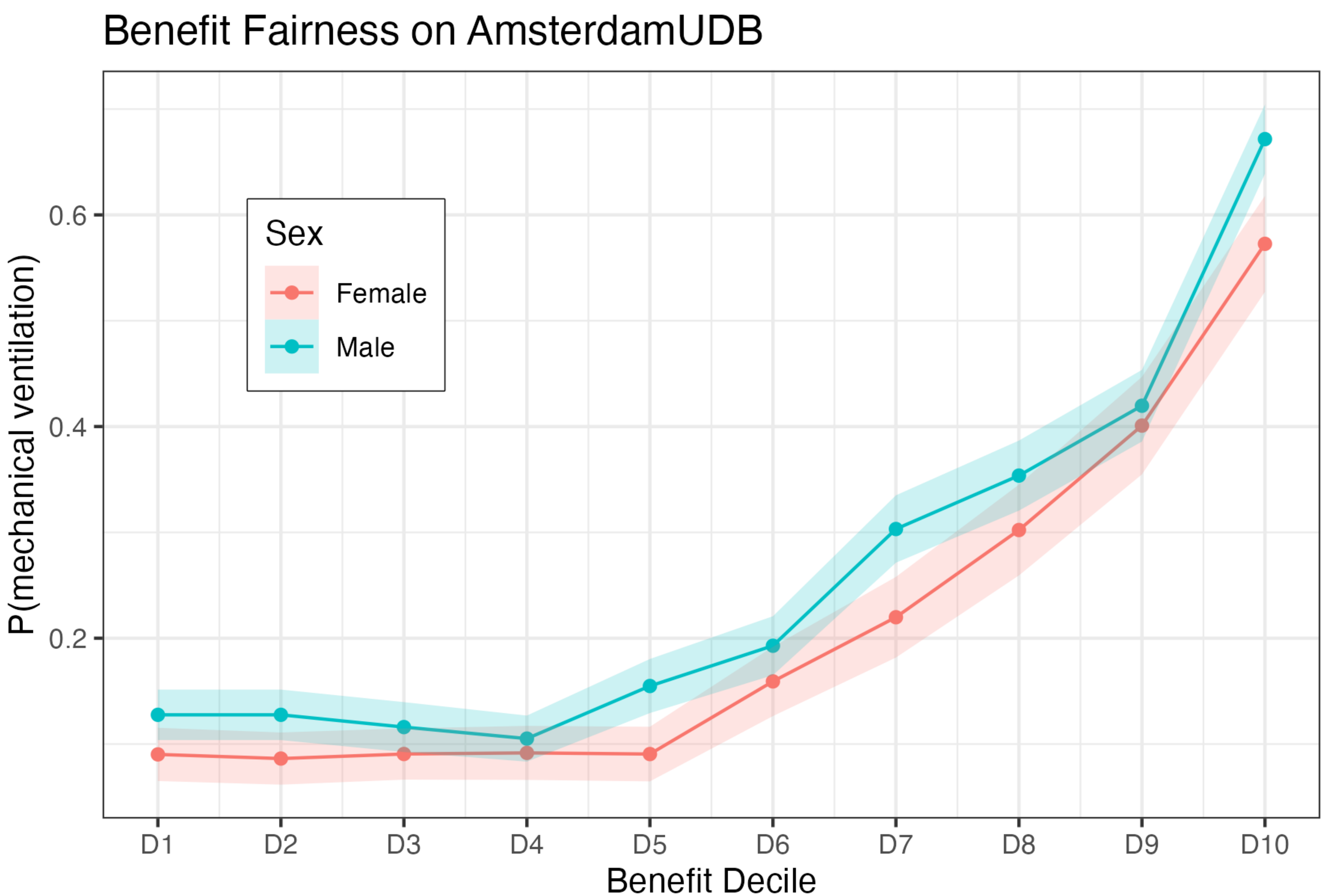
$$E[\Delta \mid x_1] - E[\Delta \mid x_0] = \underbrace{\textbf{DE}}_{X \rightarrow \Delta} + \underbrace{\textbf{IE}}_{X \rightarrow W \rightarrow \Delta} + \underbrace{\textbf{SE}}_{X \leftarrow\rightarrow Z \rightarrow \Delta}$$

women have a lower benefit biologically

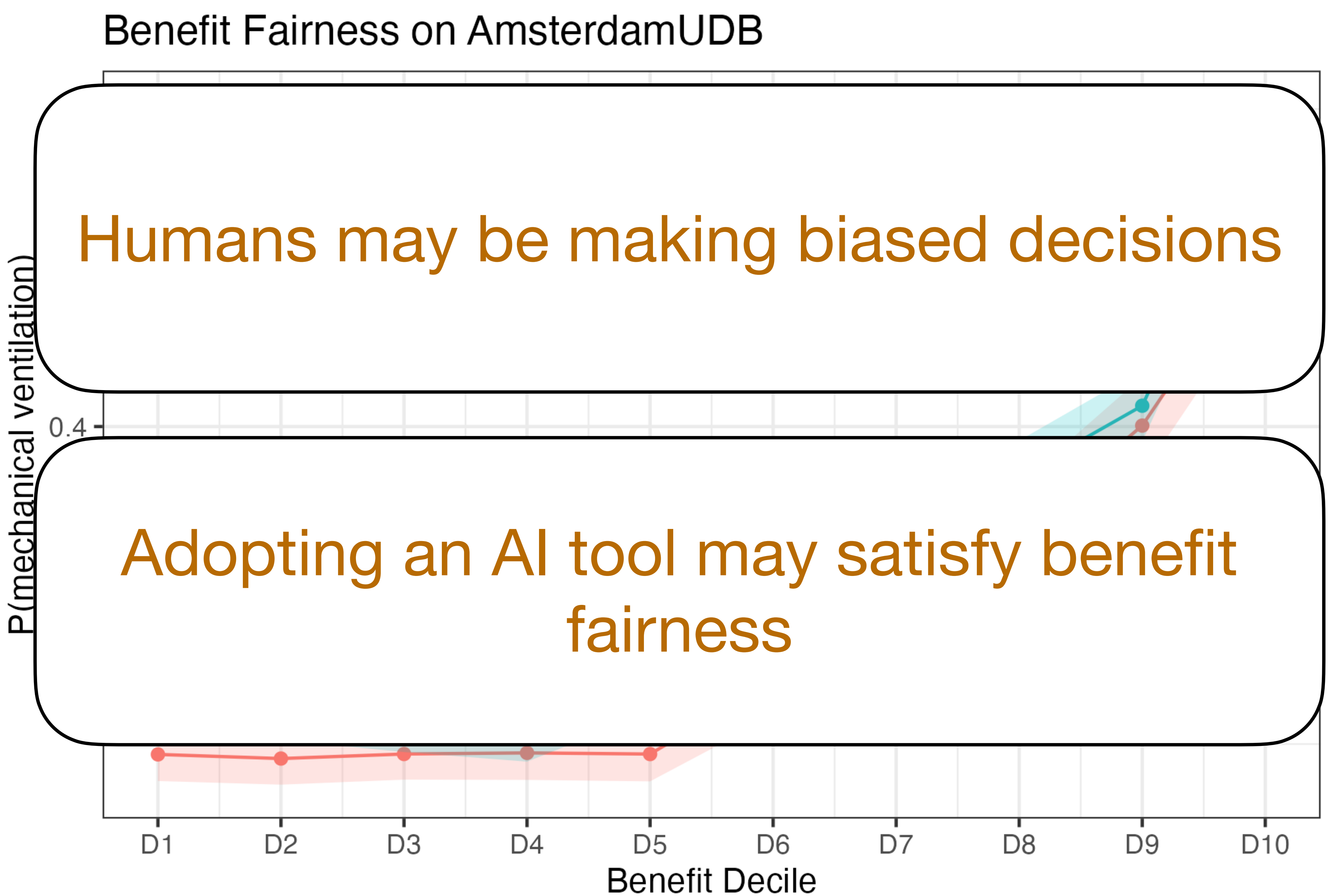
women have a lower illness severity

women in ICU are younger and thus benefit less

Causal Health Equity (Task 3): Allocating Respirators - Results



Causal Health Equity (Task 3): Allocating Respirators - Results

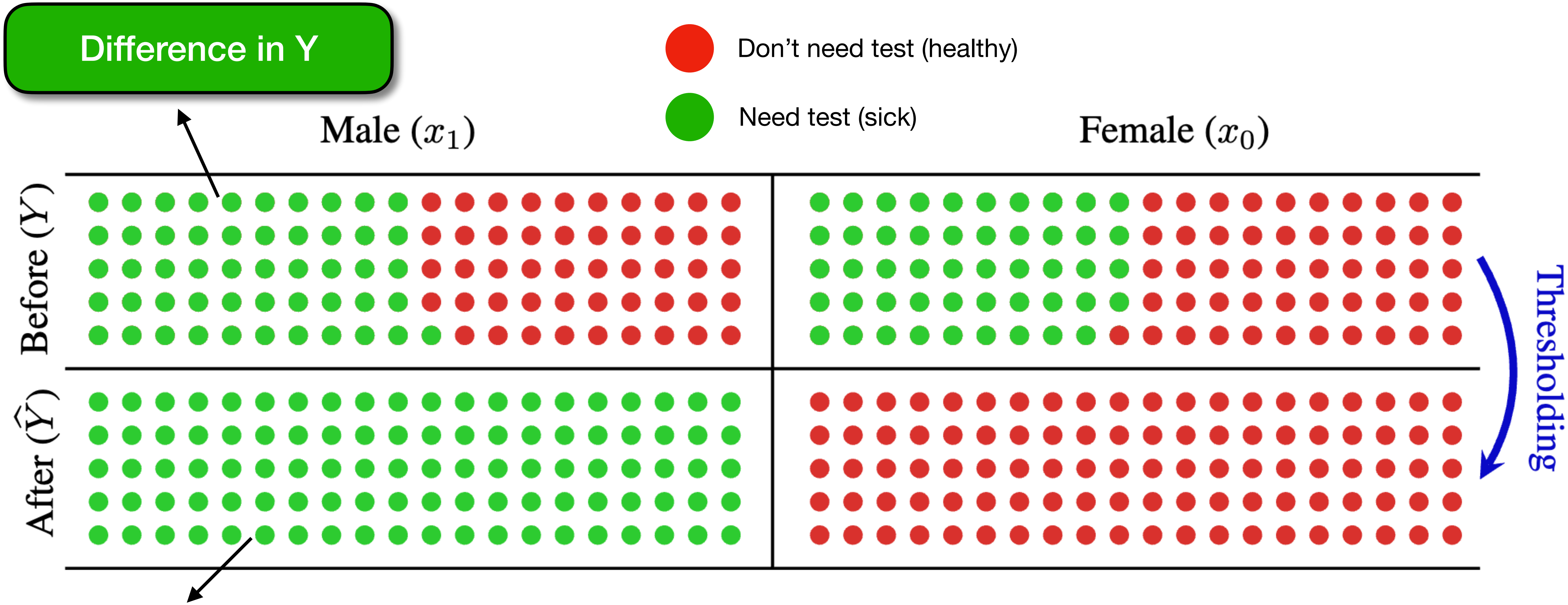


Humans may be making biased decisions

Adopting an AI tool may satisfy benefit fairness

Can AI Exacerbate Disparities?

Difference in Y

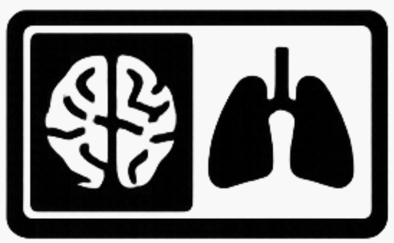


Difference in $\hat{Y}$

A small disparity in Y gives a large disparity in $\hat{Y}$!

Causal AI for Health Data: Key Takeaways

Modern data can be complex and multi-modal



radiology imaging

waveform data

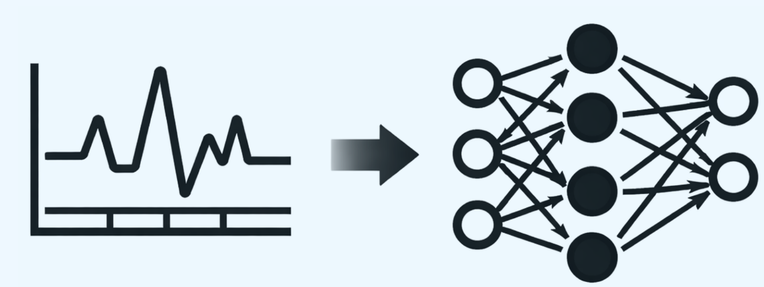
patient text notes

omics

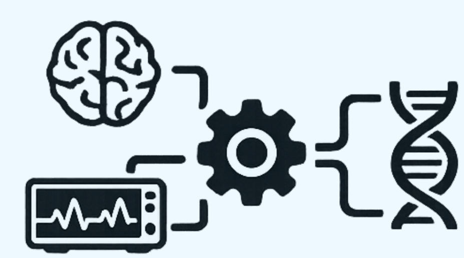
EHR data

opens the door for …

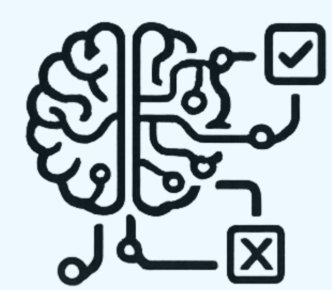
Predictive Analytics



Multi-Modal CI



AI Decision-Making



but requires taking care of trustworthy aspects:

- **Robustness**
- **Explainability**
- **Fairness**