Generalized Adjustment Under Confounding and Selection Biases

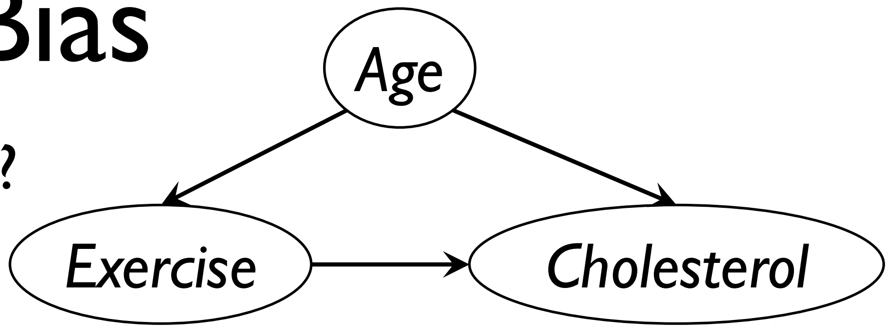
Juan D. Correa Jin Tian Elias Bareinboim

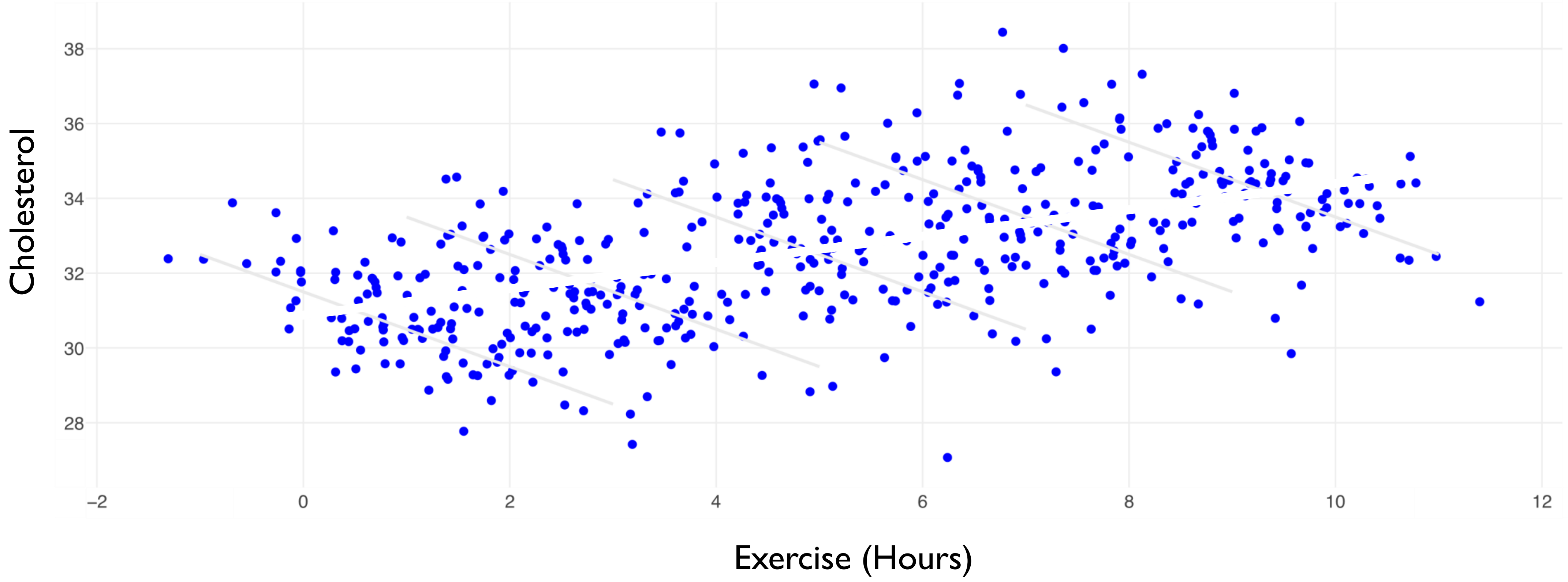
AAAI
New Orleans, 2018

Challenge 1: Confounding Bias

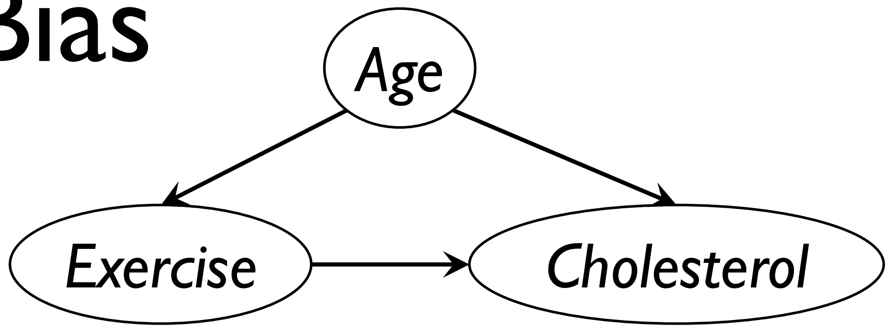
What's the causal effect of *Exercise* on *Cholesterol*?

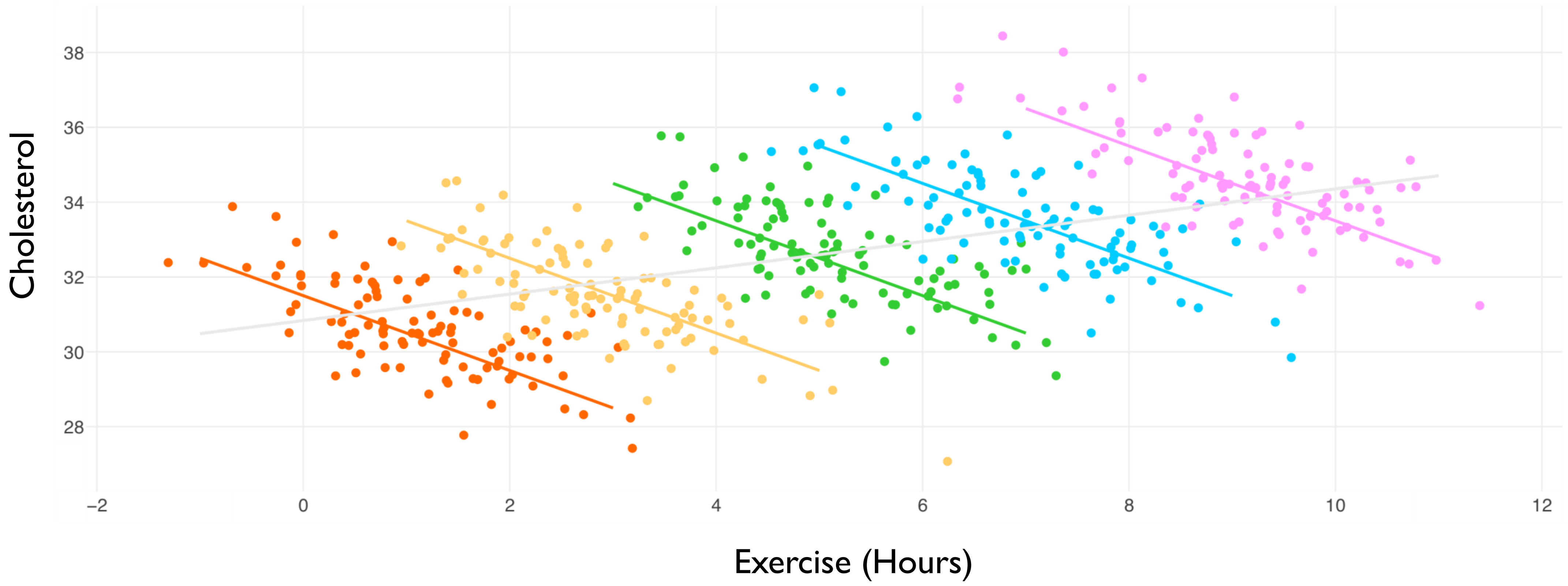
What about $P(cholesterol \mid exercise)$?



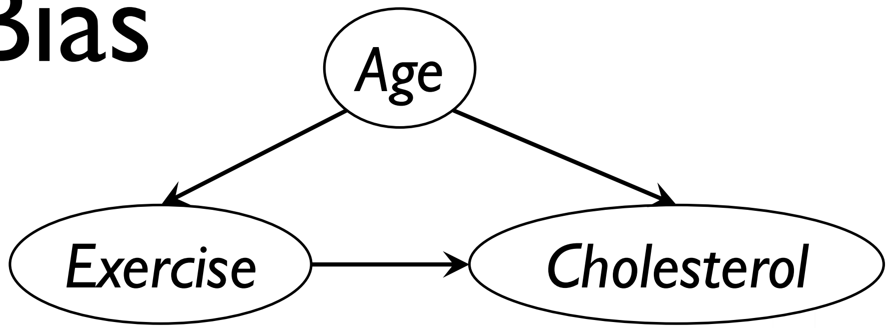


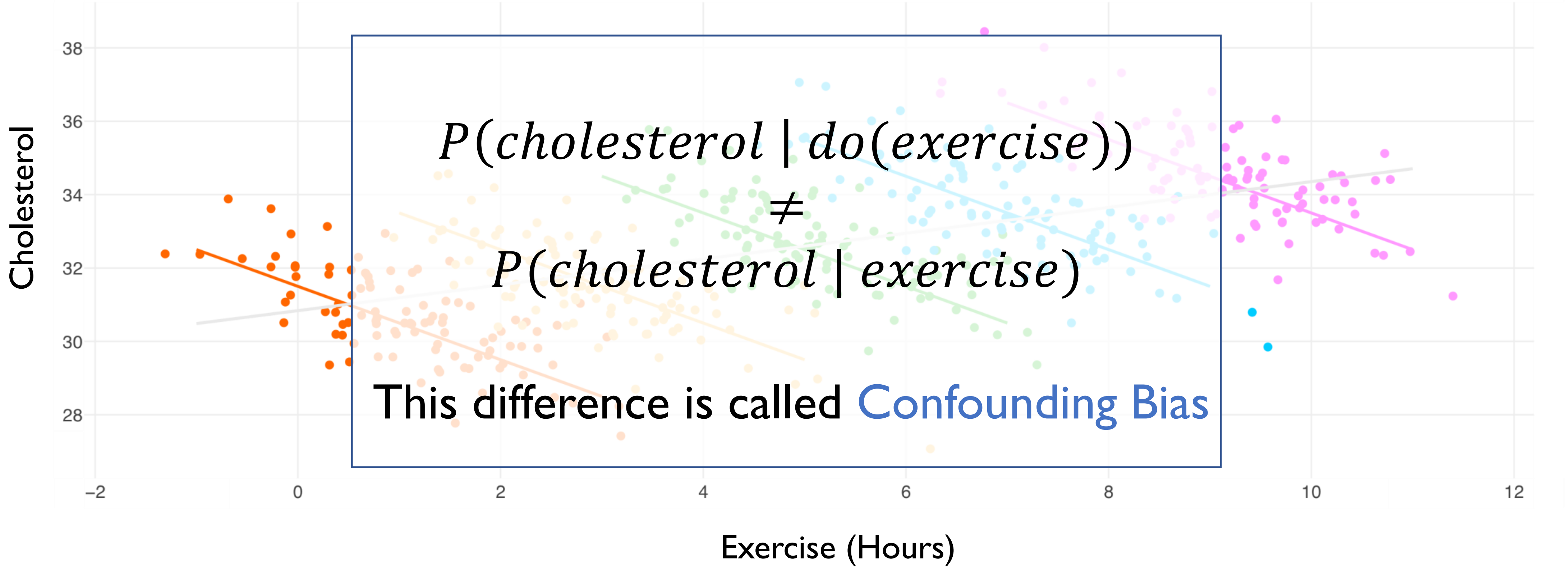
Challenge 1: Confounding Bias





Challenge 1: Confounding Bias



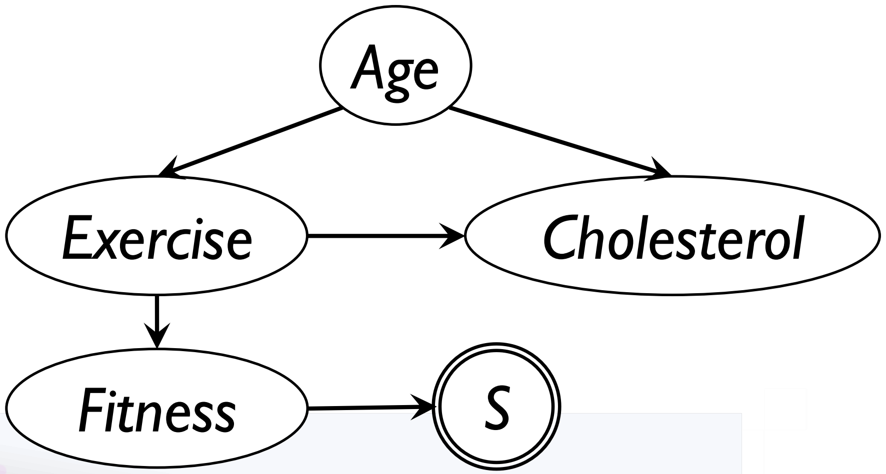


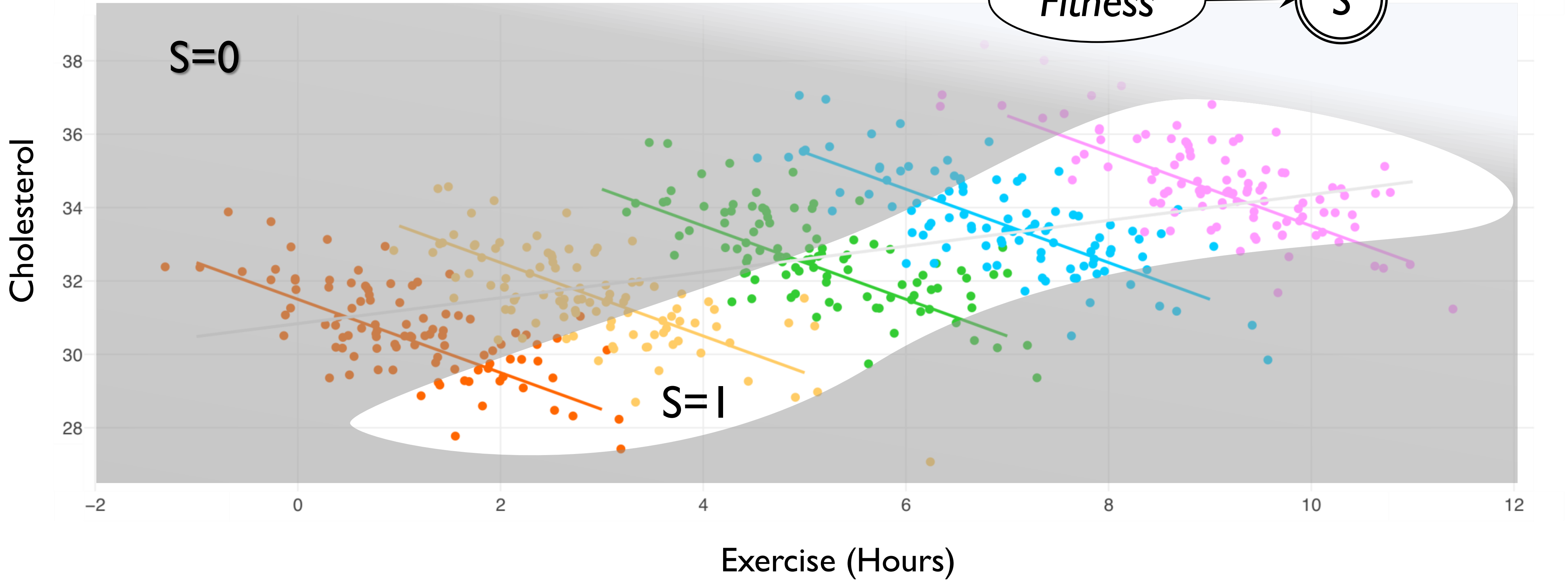
$$P(cholesterol \mid do(exercise)) \neq P(cholesterol \mid exercise)$$

This difference is called Confounding Bias

Challenge 2: Selection Bias

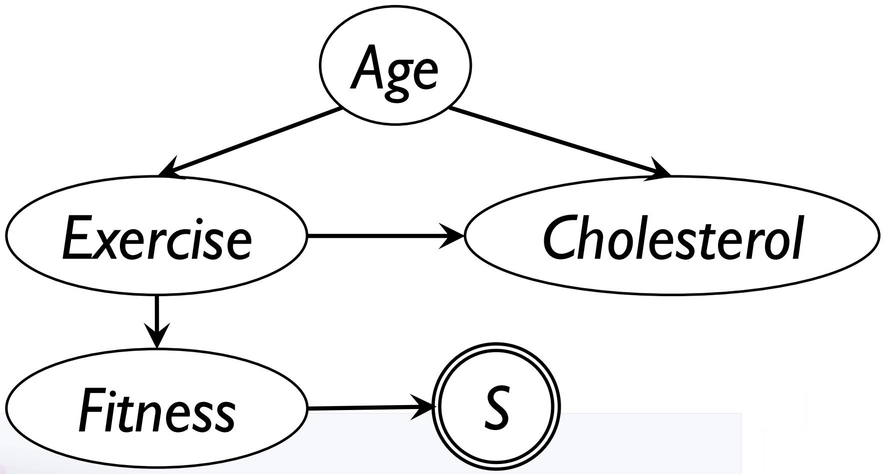
Variables in the system affect the inclusion of units in the sample

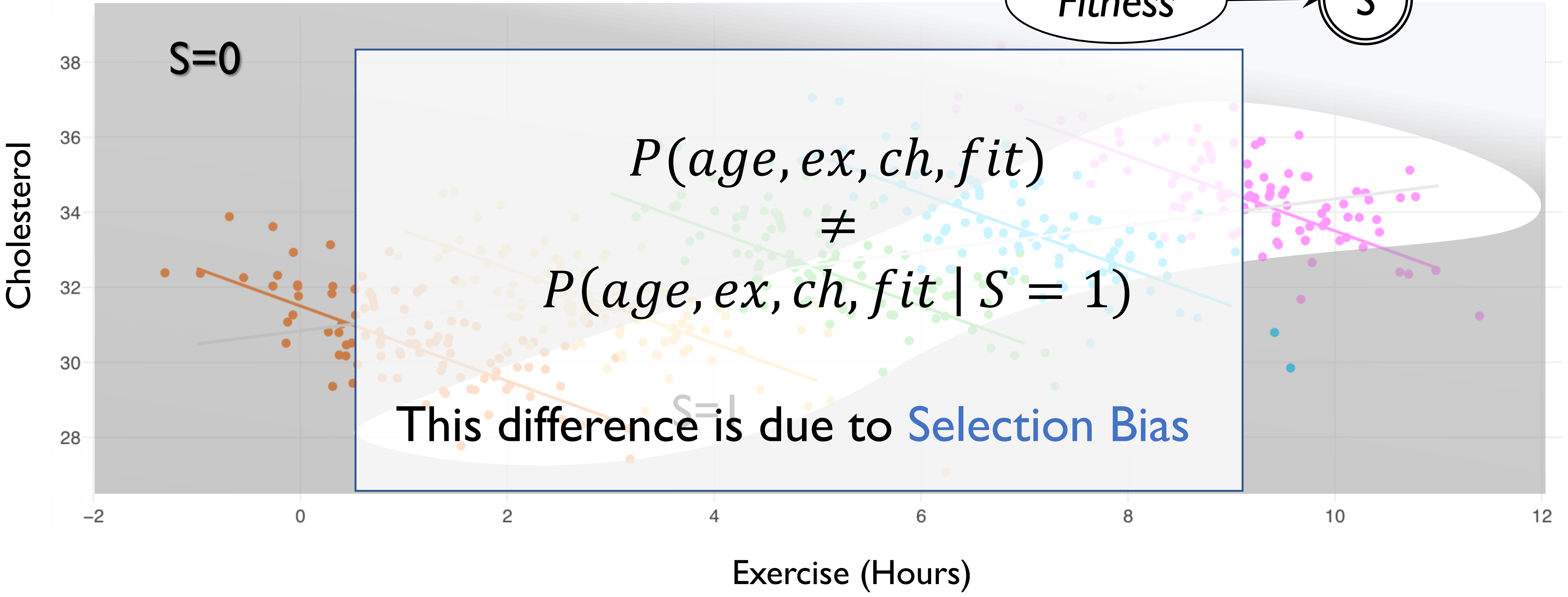




Challenge 2: Selection Bias

Variables in the system affect the inclusion of units in the sample





$$P(age, ex, ch, fit) \neq P(age, ex, ch, fit \mid S = 1)$$

This difference is due to Selection Bias

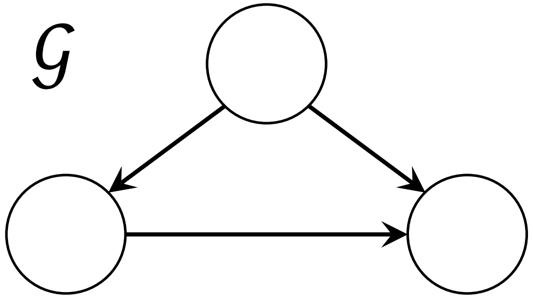
Current literature

	No Confounding	Confounding
No Selection	Association = Causation No control	Backdoor Criterion [Pearl '93] Extended Backdoor [Pearl, Paz '10] Adjustment Criterion [Shpitser, VanderWeele, Robins '10]
Selection	Controlling Selection Bias [Bareinboim and Pearl '12] Recovering from Selection Bias in Causal and Statistical Inference [Bareinboim, Tian, Pearl '14]	Selection Backdoor [Bareinboim, Tian, Pearl '14] **Generalized Adjustment** [Correa, Tian, Bareinboim '18]

Problem 1

Given:

$\mathcal{G}$



Variables
$\boldsymbol{X}, \boldsymbol{Y}$

			S	$P(\boldsymbol{v} \vert S = 1)$
			1	…
			1	…
			1	…

P_1

	$P(\boldsymbol{t})$
	…
	…
	…

P_2

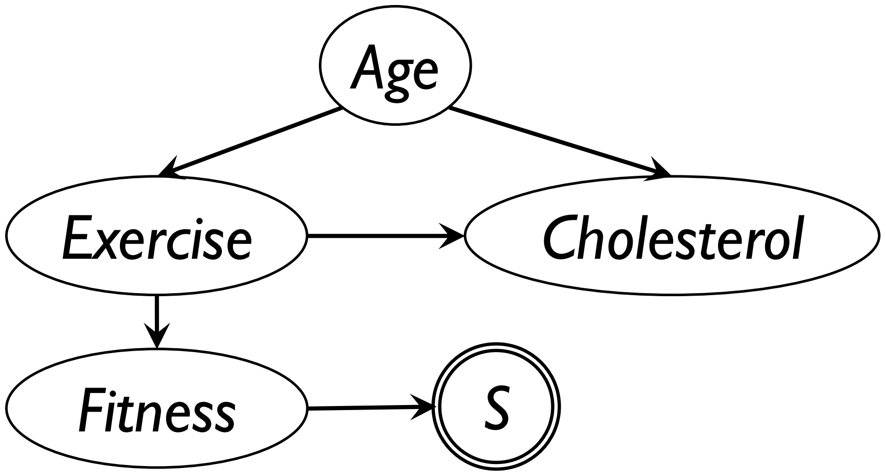
Is there a function f such that

$$P(\boldsymbol{y} | do(\boldsymbol{x})) = f(P_1, P_2)$$

?

Backdoor Adjustment

[Pearl '93; Bareinboim, Tian, Pearl '14]



Control de biases by reweighing $P(ch|ex)$ in each subpopulation defined by a set of covariates

$$P(ch|do(ex)) = \sum_{age,fit} \underbrace{P(ch|ex, age, fit, S=1)\, P(age|fit, S=1)}_{\text{Biased Data}} \underbrace{P(fit)}_{\text{Unbiased Data}}$$

Here $(\underbrace{\{age, fit\}}_{\text{Biased}}, \underbrace{\{fit\}}_{\text{Unbiased}})$ is called and **adjustment pair**

Result 1: Graphical Criterion

Definition 8: A pair $\left(Z, Z^T\right)$ satisfies the criterion if:

a) No element in $\boldsymbol{Z}$ is a descendant in $G_{\overline{\boldsymbol{X}}}$ of any $W \notin \boldsymbol{X}$ lying on a proper causal path from $\boldsymbol{X}$ to $\boldsymbol{Y}$

b) All proper non-causal paths in $\mathcal{G}$ from $\boldsymbol{X}$ to $\boldsymbol{Y}$ are blocked by $\boldsymbol{Z}$ and S

c) $\boldsymbol{Z}^{\boldsymbol{T}}$ d-separates $\boldsymbol{Y}$ from S in the proper backdoor graph, i.e. $\left(\boldsymbol{Y} \perp\!\!\!\perp S \mid \boldsymbol{Z}^{\boldsymbol{T}}\right)_{G_{\boldsymbol{XY}}^{pbd}}$

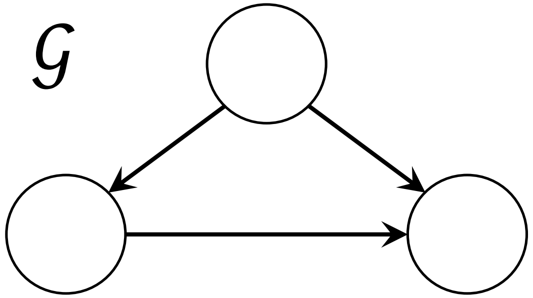
Theorem 1

$\left(\boldsymbol{Z}, \boldsymbol{Z}^{\boldsymbol{T}}\right)$ is an admissible adjustment pair *if* and *only if* **definition 8** is satisfied.

Problem 2

Given:

$\mathcal{G}$



Variables
$\boldsymbol{X}, \boldsymbol{Y}$

			S	$P(\boldsymbol{v} \mid S = 1)$
			1	…
			1	…
			1	…

P_1

	$P(\boldsymbol{t})$
	…
	…
	…

P_2

Can we list all possible pairs $(\boldsymbol{Z}, \boldsymbol{Z}^T)$ that are admissible for estimating $P(\boldsymbol{y} \mid do(\boldsymbol{x}))$?

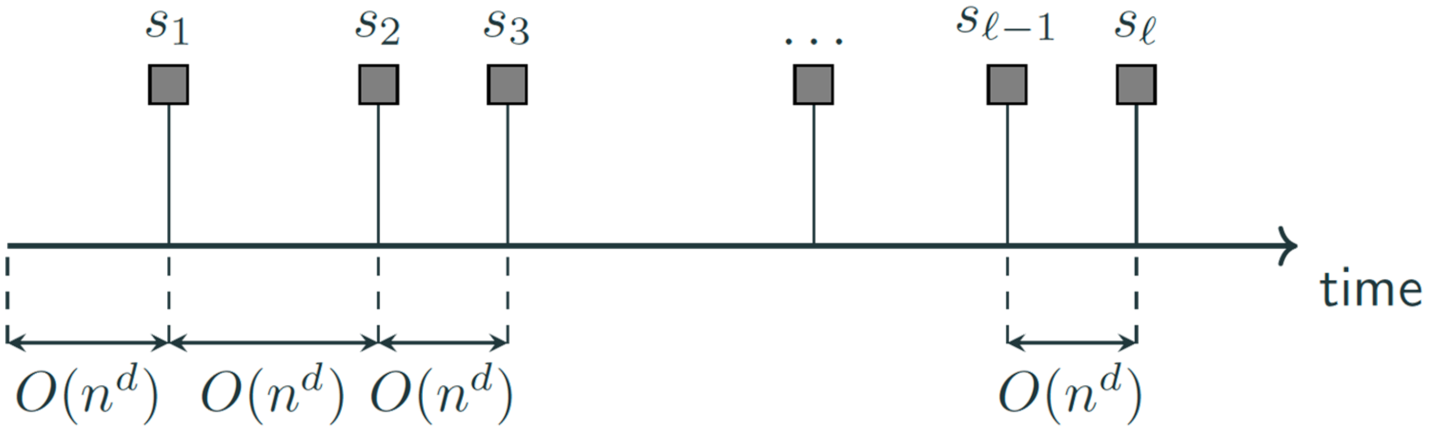
There could be exponentially many admissible pairs!

Result 2: How to list adjustment pairs?

Algorithm 1 List Admissible Sets

Given a causal diagram $\mathcal{G}$ and disjoint sets of variables $\boldsymbol{X}, \boldsymbol{Y}$;

list all adjustment pairs for the effect $P(\boldsymbol{y} \mid do(\boldsymbol{x}))$ ***polynomial delay*** [Takata '10].



- Key Idea: Reduce the problem to find vertex separators in an induced undirected graph [van der Zander, Liskiewicz, Textor '14]

Problem 3

- We have selected an admissible pair $(\boldsymbol{Z}, \boldsymbol{Z^T})$
- How hard is to estimate the effect from data?

$$P(\boldsymbol{y}|do(\boldsymbol{x})) = \sum_{\boldsymbol{Z}} P(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{z}, S=1) P\left(\boldsymbol{z} \setminus \boldsymbol{z^T} \middle| \boldsymbol{z^T}, S=1\right) P(\boldsymbol{z^T})$$

- The naive approach takes time proportional to the dimensionality of $\boldsymbol{Z}$

Result 3: Extend Inverse Probability Weighting

$$P(\boldsymbol{y}|do(\boldsymbol{x})) = \sum_{\boldsymbol{z}} P(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{z}, S = 1) P\left(\boldsymbol{z} \setminus \boldsymbol{z}^{\boldsymbol{T}} \middle| \boldsymbol{z}^{\boldsymbol{T}}, S = 1\right) P(\boldsymbol{z}^{\boldsymbol{T}})$$

$$= \sum_{\boldsymbol{z}} \frac{P(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{z}|S = 1)}{P(\boldsymbol{x}|\boldsymbol{z}, S = 1)} \frac{P(\boldsymbol{z}^{\boldsymbol{T}})}{P(\boldsymbol{z}^{\boldsymbol{T}}|S = 1)}$$

- $P(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{z}|S = 1)$ are entries from the **Biased Joint Distribution**!
- $1/P(\boldsymbol{x}|\boldsymbol{z}, S = 1)$ can be approximated by a "**Biased Propensity Score**"
- $\frac{P(\boldsymbol{z}^{\boldsymbol{T}})}{P(\boldsymbol{z}^{\boldsymbol{T}}|S=1)} = \frac{P(S=1)}{P(S=1|\boldsymbol{z}^{\boldsymbol{T}})}$ is known as "**Inverse Probability-of-selection Weight**"
- Expression may be estimated from data in time proportional to the number of samples

Summary

1. Complete characterization of admissible adjustment pairs (under both confounding and selection bias)
2. Algorithm to list all admissible adjustment pairs with polynomial delay
3. Extension of Inverse Probability Weighting compatible with generalized adjustment

Thanks!