

Machine Learning for Healthcare

Causal Inference Part 1

David Sontag

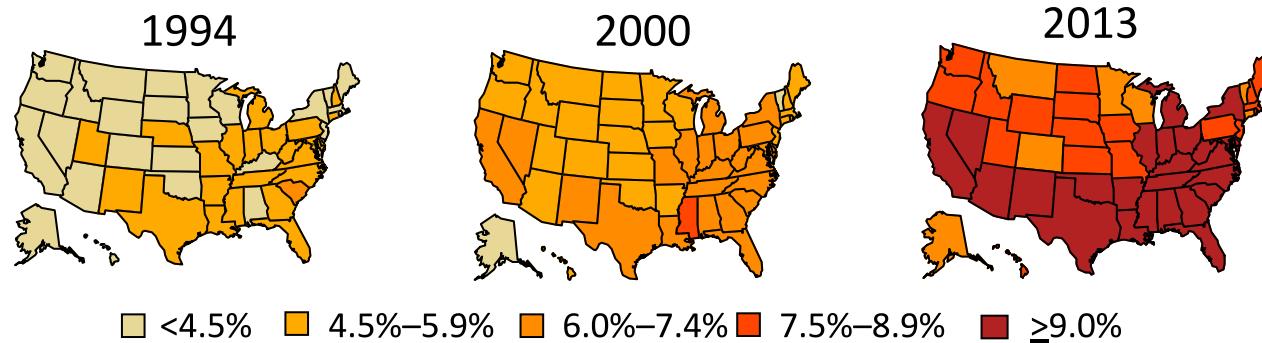


INSTITUTE FOR MEDICAL
ENGINEERING & SCIENCE



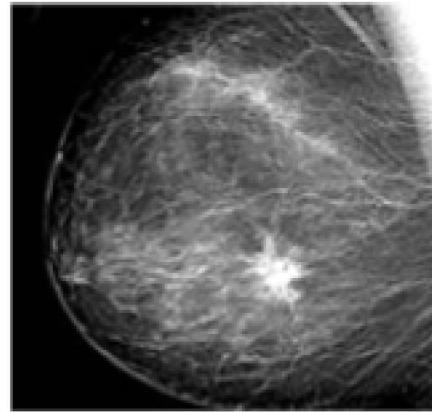
HEALTH SCIENCES
& TECHNOLOGY

Does gastric bypass surgery prevent onset of diabetes?

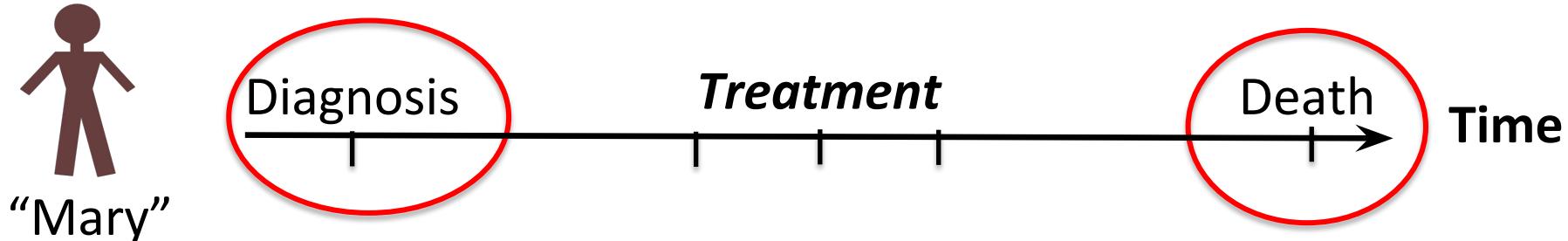


- In Lecture 4 & PS2 we used machine learning for early detection of Type 2 diabetes
- Health system doesn't want to know how to predict diabetes – they want to know how to *prevent it*
- Gastric bypass surgery is the highest negative weight (9th most predictive feature)
 - Does this mean it would be a good intervention?

What is the likelihood this patient, with breast cancer, will survive 5 years?

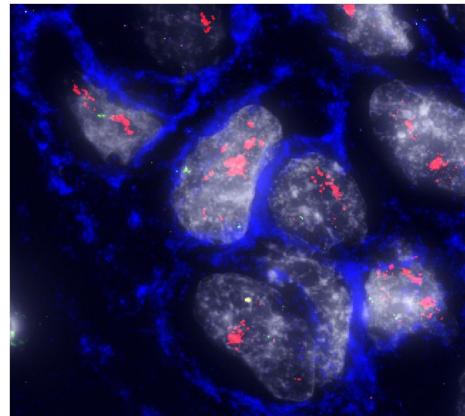


- Such predictive models widely used to stage patients. Should we initiate treatment? How aggressive?
- What could go wrong if we trained to predict survival, and then used to guide patient care?



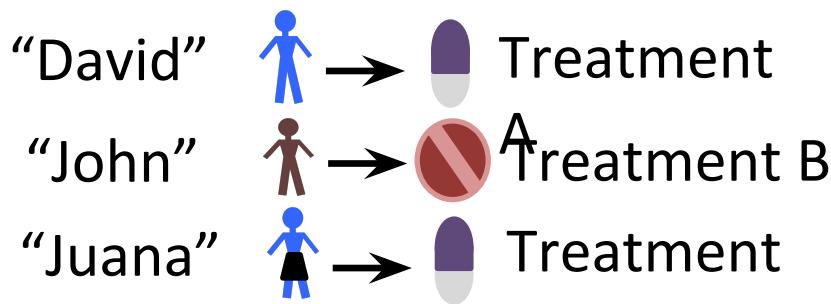
A long survival time may be because of treatment!

What treatment should we give this patient?



Expansion pathology
(image from Andy Beck)

- People respond differently to treatment
- Goal: use data from other patients and their journeys to guide future treatment decisions
- What could go wrong if we trained to predict (past) treatment decisions?



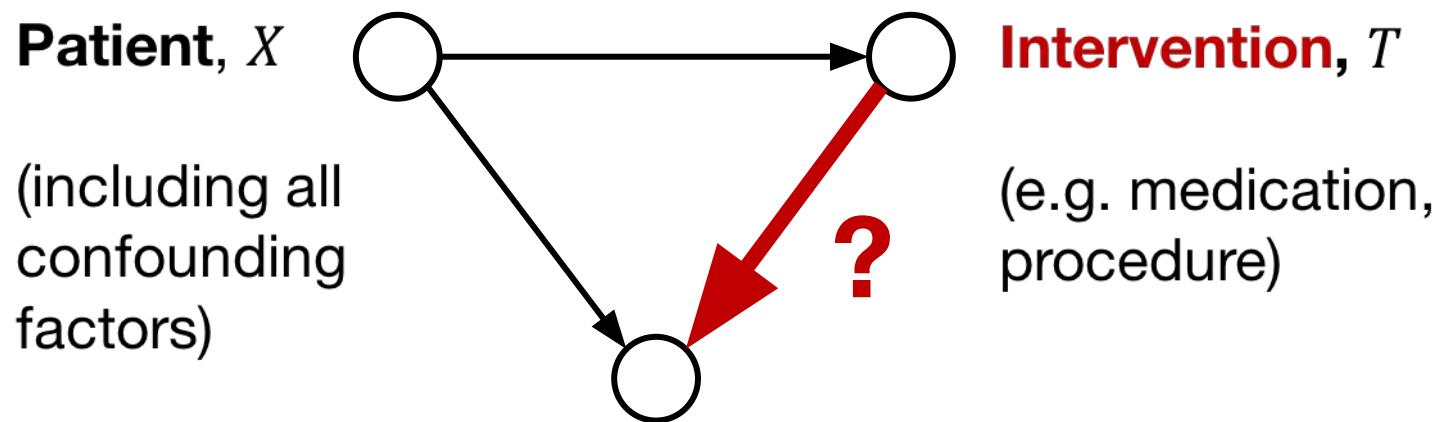
**Best this can do is
match current
medical practice!**

Does smoking cause lung cancer?



- Doing a randomized control trial is unethical
- Could we simply answer this question by comparing $\text{Pr}(\text{lung cancer} \mid \text{smoker})$ vs $\text{Pr}(\text{lung cancer} \mid \text{nonsmoker})$?
- No! Answering such questions from observational data is difficult because of *confounding*

To properly answer, need to formulate as *causal* questions:



High dimensional

Observational data

Potential Outcomes Framework (Rubin-Neyman Causal Model)

- Each unit (individual) x_i has two potential outcomes:
 - $Y_0(x_i)$ is the potential outcome had the unit not been treated:
“control outcome”
 - $Y_1(x_i)$ is the potential outcome had the unit been treated:
“treated outcome”
- Conditional average treatment effect for unit i :
$$CATE(x_i) = \mathbb{E}_{Y_1 \sim p(Y_1|x_i)} [Y_1|x_i] - \mathbb{E}_{Y_0 \sim p(Y_0|x_i)} [Y_0|x_i]$$
- Average Treatment Effect:
$$ATE := \mathbb{E}[Y_1 - Y_0] = \mathbb{E}_{x \sim p(x)} [CATE(x)]$$

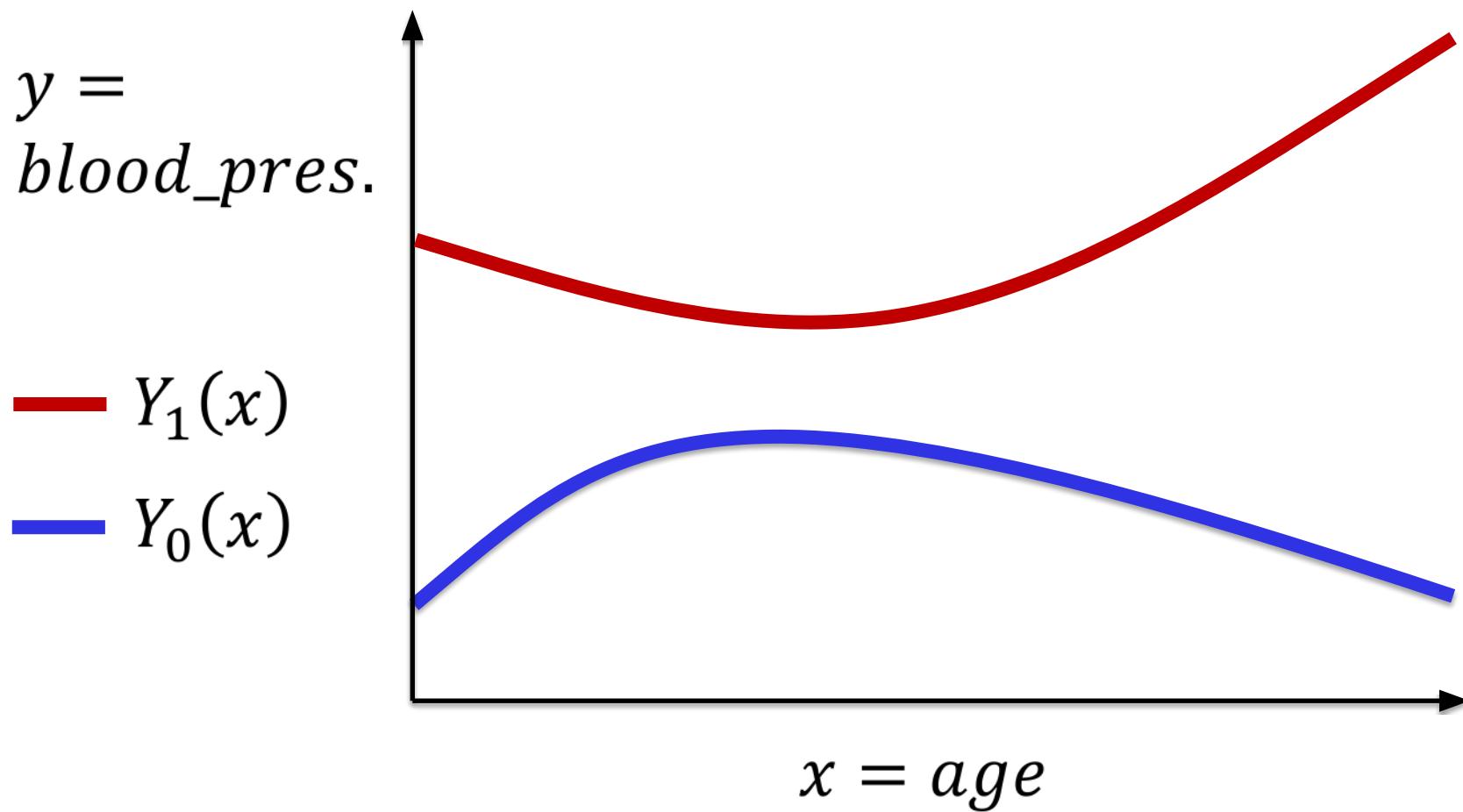
Potential Outcomes Framework (Rubin-Neyman Causal Model)

- Each unit (individual) x_i has two potential outcomes:
 - $Y_0(x_i)$ is the potential outcome had the unit not been treated:
“control outcome”
 - $Y_1(x_i)$ is the potential outcome had the unit been treated:
“treated outcome”
- Observed factual outcome:
$$y_i = t_i Y_1(x_i) + (1 - t_i) Y_0(x_i)$$
- Unobserved counterfactual outcome:
$$y_i^{CF} = (1 - t_i) Y_1(x_i) + t_i Y_0(x_i)$$

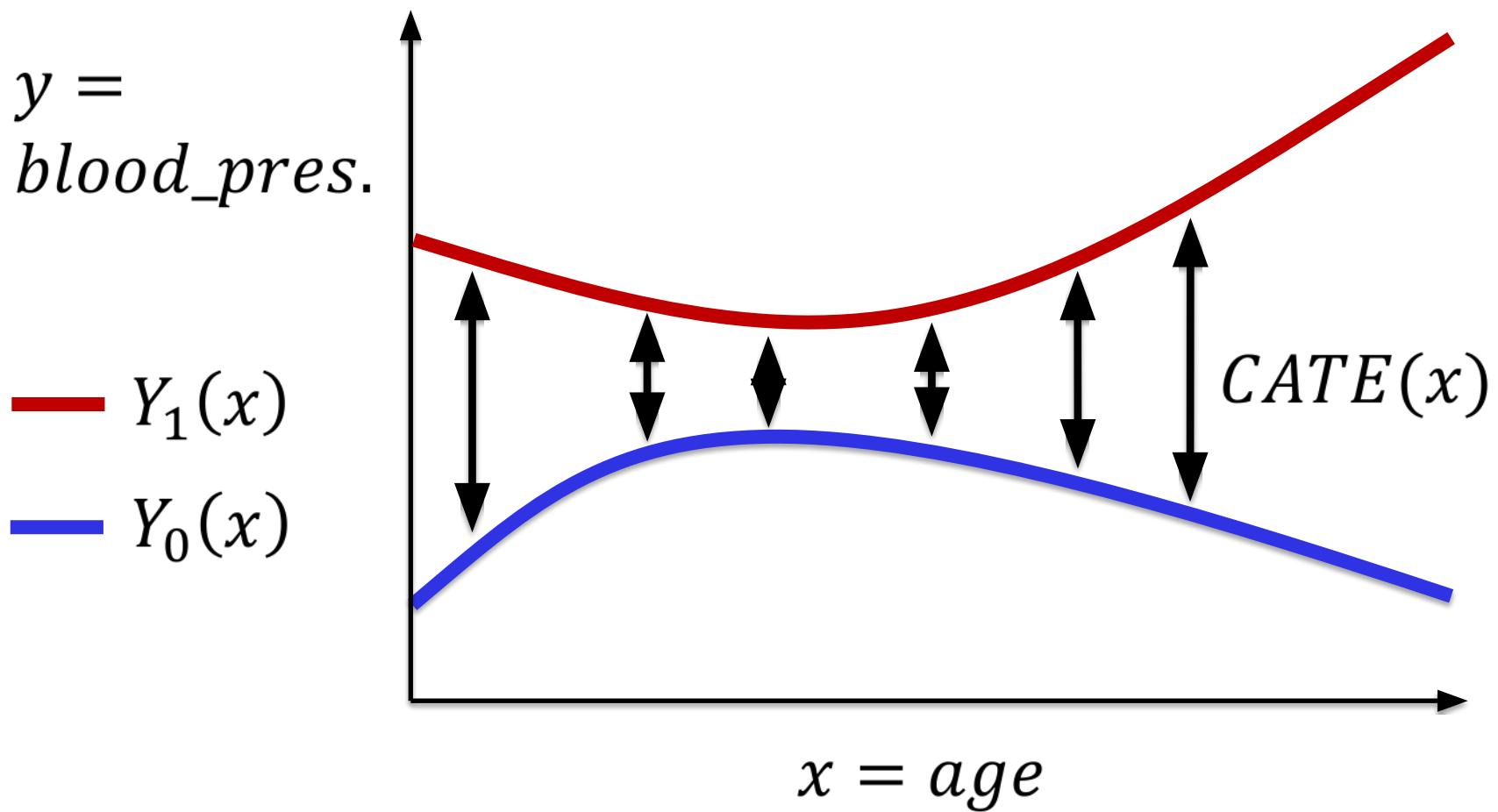
“The fundamental problem of causal inference”

We only ever observe one of the two outcomes

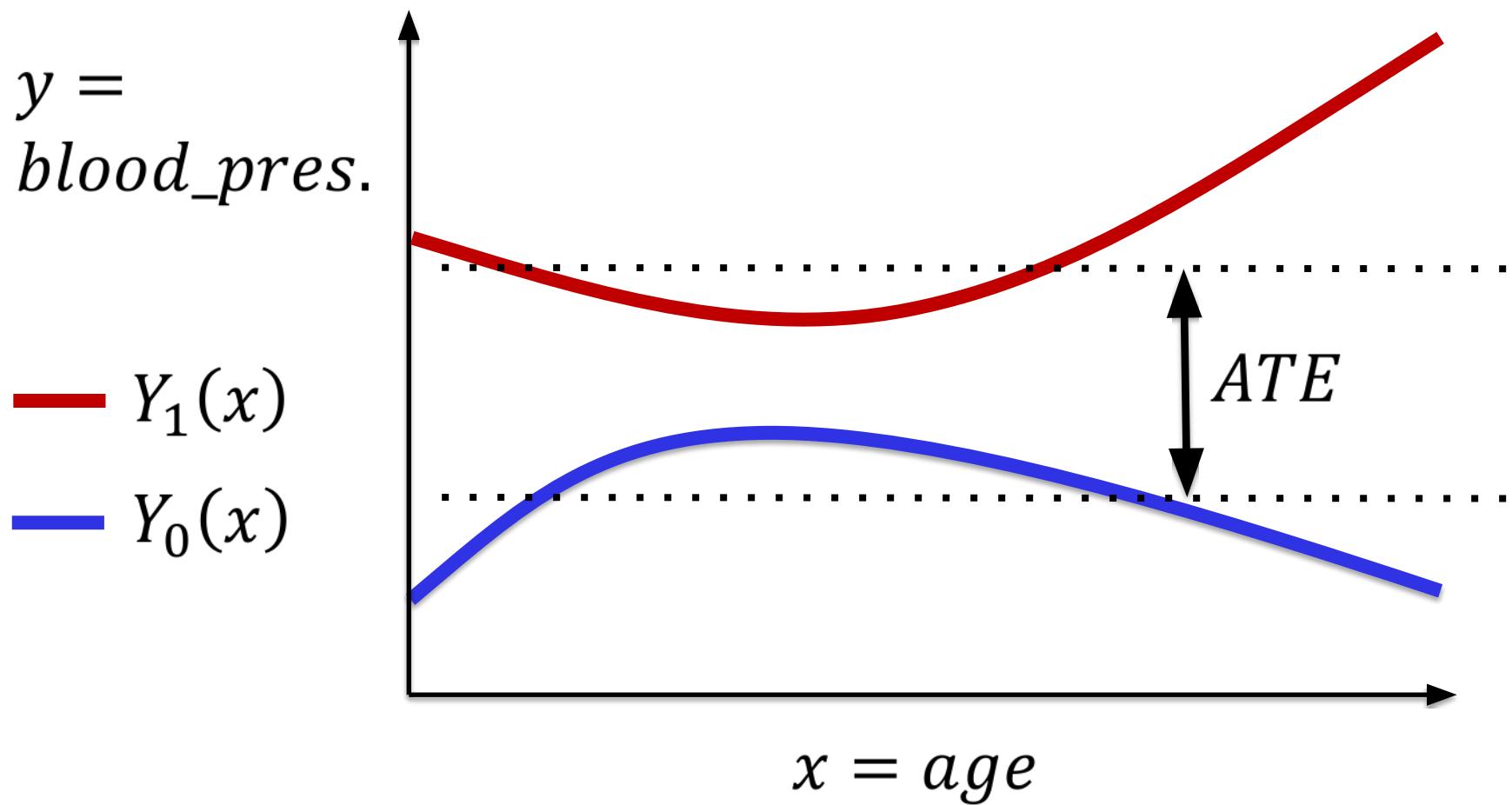
Example – Blood pressure and age



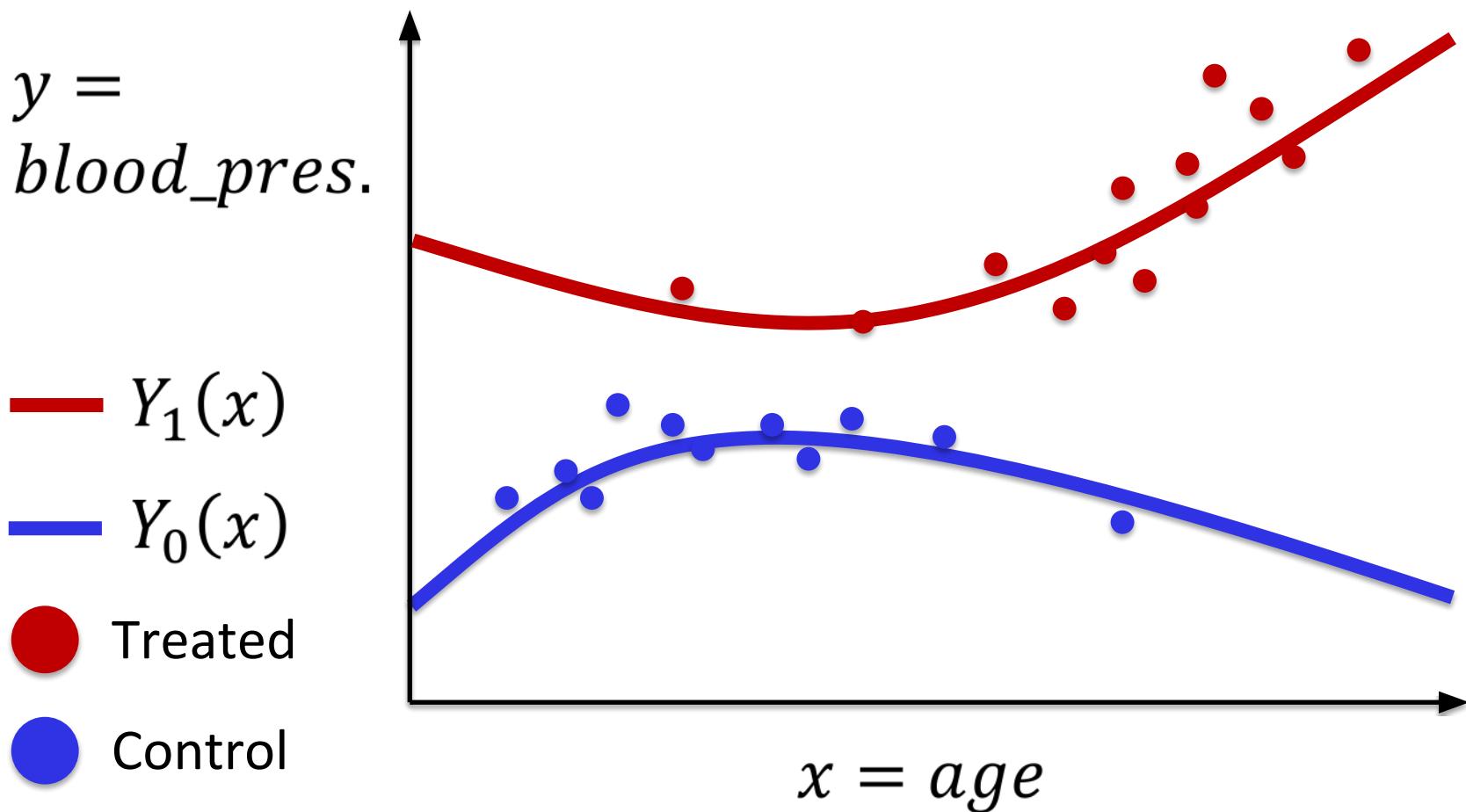
Blood pressure and age



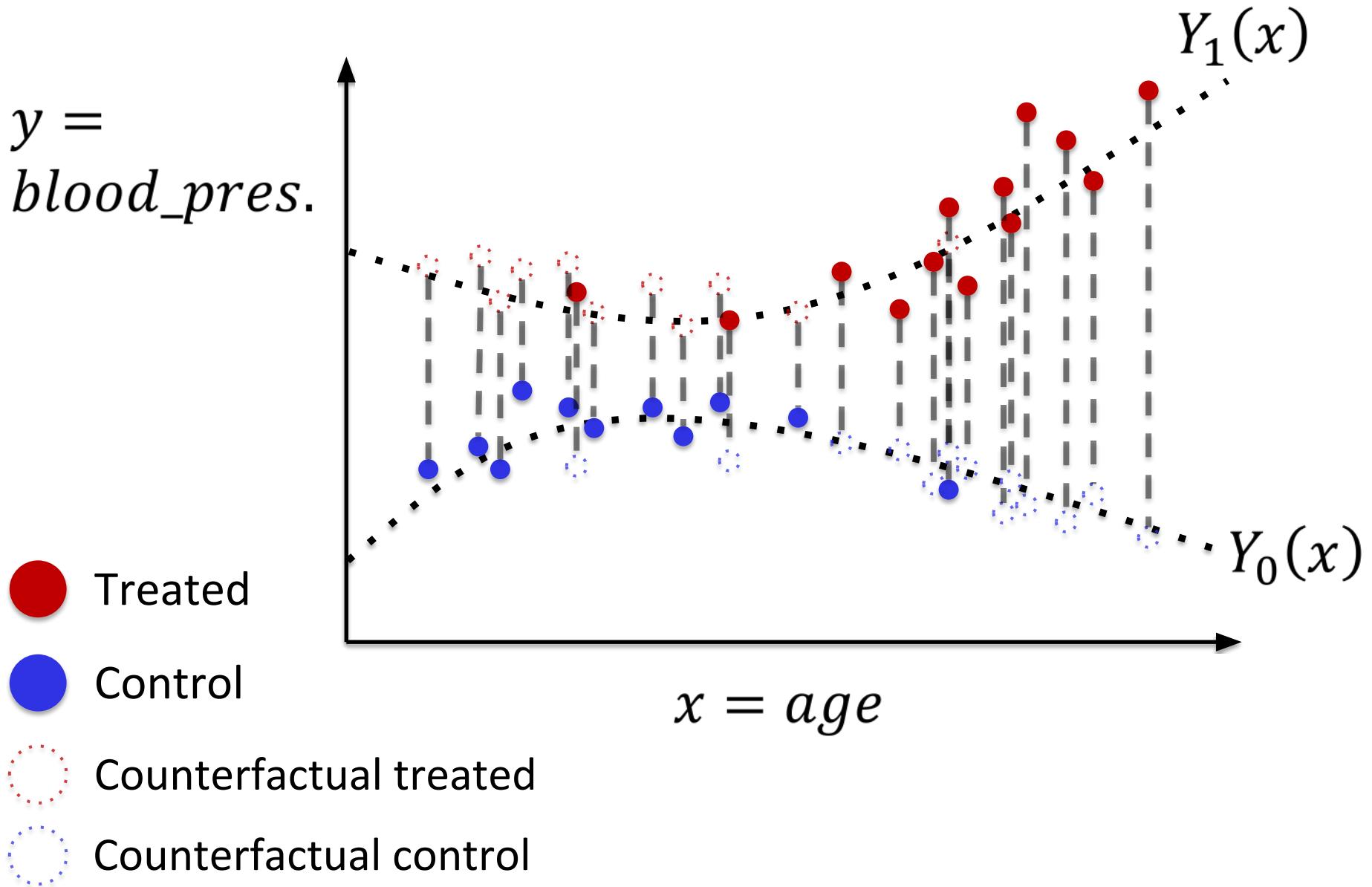
Blood pressure and age



Blood pressure and age



Blood pressure and age



(age, gender, exercise,treatment)	Sugar levels <i>had they received</i> medication A	Sugar levels <i>had they received</i> medication B	Observed sugar levels
(45, F, 0, A)	6	5.5	6
(45, F, 1, B)	7	6.5	6.5
(55, M, 0, A)	7	6	7
(55, M, 1, B)	9	8	8
(65, F, 0, B)	8.5	8	8
(65, F, 1, A)	7.5	7	7.5
(75, M, 0, B)	10	9	9
(75, M, 1, A)	8	7	8

(Example from Uri Shalit)

(age, gender, exercise)	Sugar levels <i>had they received</i> medication A	Sugar levels <i>had they received</i> medication B	Observed sugar levels
(45, F, 0)	6	5.5	6
(45, F, 1)	7	6.5	6.5
(55, M, 0)	7	6	7
(55, M, 1)	9	8	8
(65, F, 0)	8.5	8	8
(65, F, 1)	7.5	7	7.5
(75, M, 0)	10	9	9
(75, M, 1)	8	7	8

(Example from Uri Shalit)

(age, gender, exercise)	Y_0 : Sugar levels <i>had they received</i> medication A	Y_1 : Sugar levels <i>had they received</i> medication B	Observed sugar levels
(45, F, 0)	6	5.5	6
(45, F, 1)	7	6.5	6.5
(55, M, 0)	7	6	7
(55, M, 1)	9	8	8
(65, F, 0)	8.5	8	8
(65, F, 1)	7.5	7	7.5
(75, M, 0)	10	9	9
(75, M, 1)	8	7	8

(Example from Uri Shalit)

(age,gender, exercise)	Sugar levels <i>had they received</i> medication A	Sugar levels <i>had they received</i> medication B	Observed sugar levels
(45, F, 0)	6	5.5	6
(45, F, 1)	7	6.5	6.5
(55, M, 0)	7	6	7
(55, M, 1)	9	8	8
(65, F, 0)	8.5	8	8
(65, F, 1)	7.5	7	7.5
(75, M, 0)	10	9	9
(75, M, 1)	8	7	8

$$\text{mean}(\text{sugar} \mid \text{medication B}) - \text{mean}(\text{sugar} \mid \text{medication A}) = ?$$

$$\text{mean}(\text{sugar} \mid \text{had they received B}) - \text{mean}(\text{sugar} \mid \text{had they received A}) = ?$$

(Example from Uri Shalit)

(age,gender, exercise)	Sugar levels <i>had they received</i> medication A	Sugar levels <i>had they received</i> medication B	Observed sugar levels
(45, F, 0)	6	5.5	6
(45, F, 1)	7	6.5	6.5
(55, M, 0)	7	6	7
(55, M, 1)	9	8	8
(65, F, 0)	8.5	8	8
(65, F, 1)	7.5	7	7.5
(75, M, 0)	10	9	9
(75, M, 1)	8	7	8

$$\text{mean}(\text{sugar} | \text{medication B}) - \text{mean}(\text{sugar} | \text{medication A}) = 7.875 - 7.125 = 0.75$$

$$\text{mean}(\text{sugar} | \text{had they received B}) - \text{mean}(\text{sugar} | \text{had they received A}) = 7.125 - 7.875 = -0.75$$

(Example from Uri Shalit)

Typical assumption – no unmeasured confounders

Y_0, Y_1 : potential outcomes for control and treated

x : unit covariates (features)

T: treatment assignment

We assume:

$$(Y_0, Y_1) \perp\!\!\!\perp T \mid x$$

The potential outcomes are independent of treatment assignment, conditioned on covariates x

Typical assumption – no unmeasured confounders

Y_0, Y_1 : potential outcomes for control and treated

x : unit covariates (features)

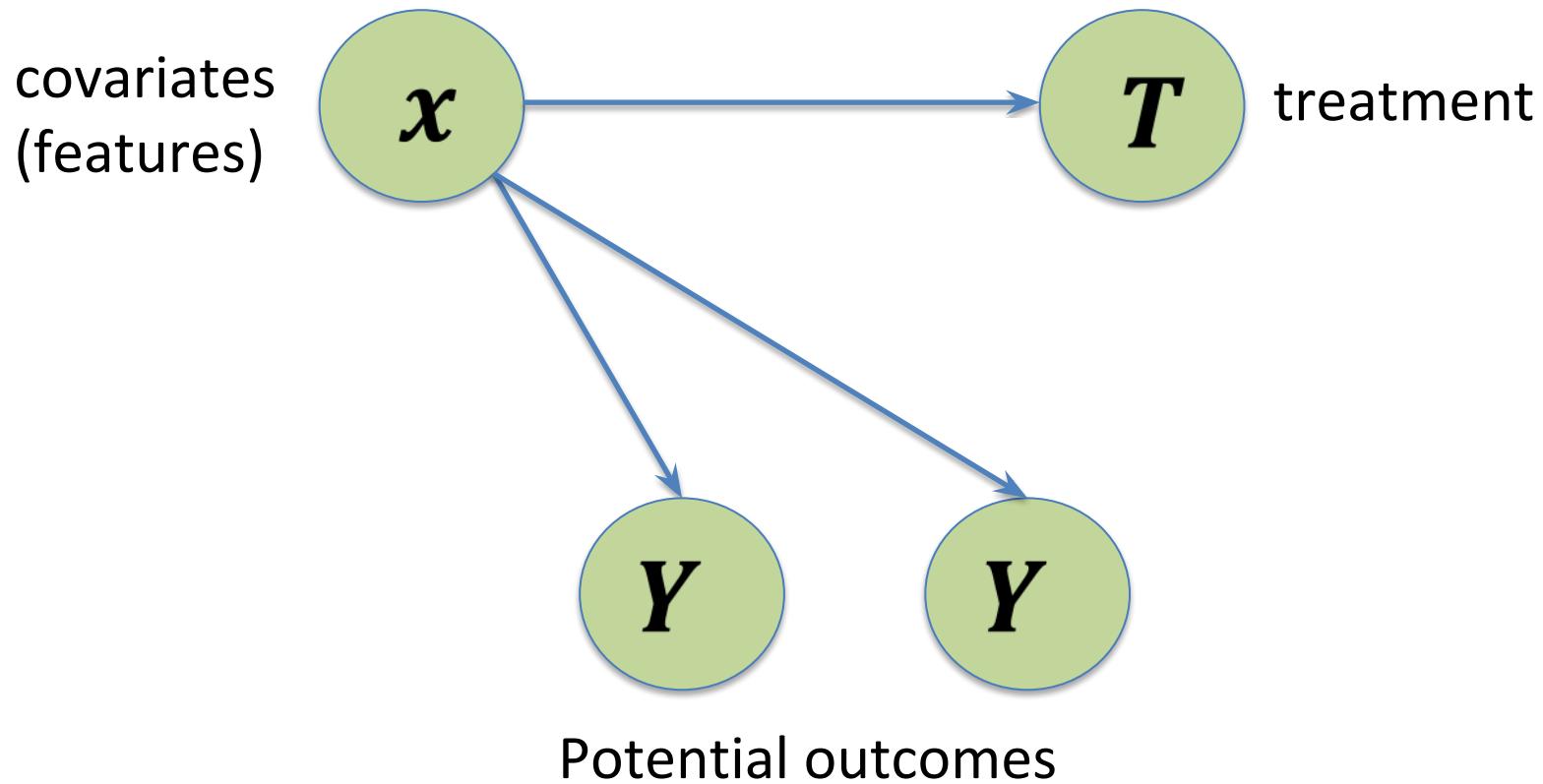
T: treatment assignment

We assume:

$$(Y_0, Y_1) \perp\!\!\!\perp T \mid x$$

Ignorability

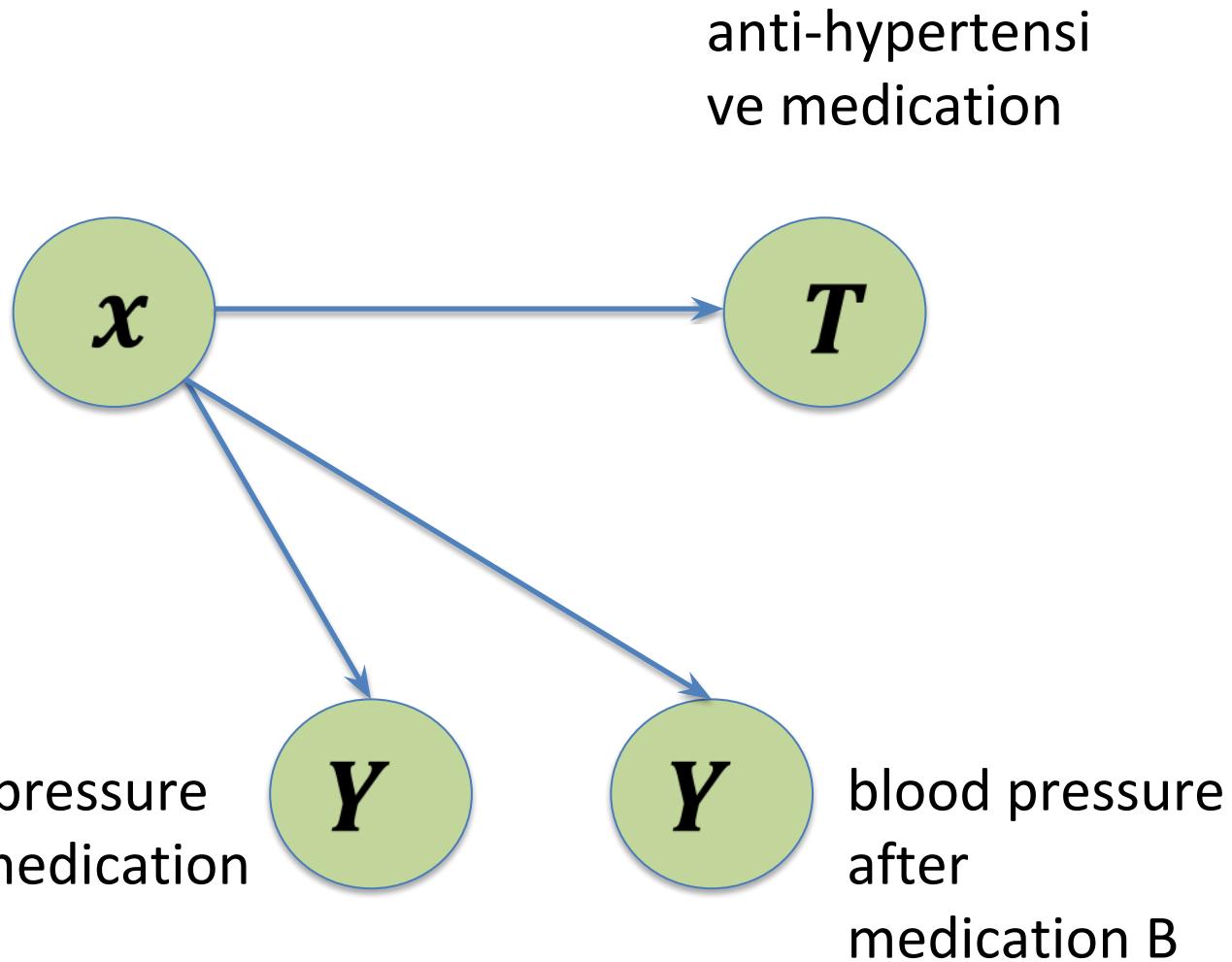
Ignorability



$$(Y_0, Y_1) \perp\!\!\!\perp T \mid x$$

Ignorability

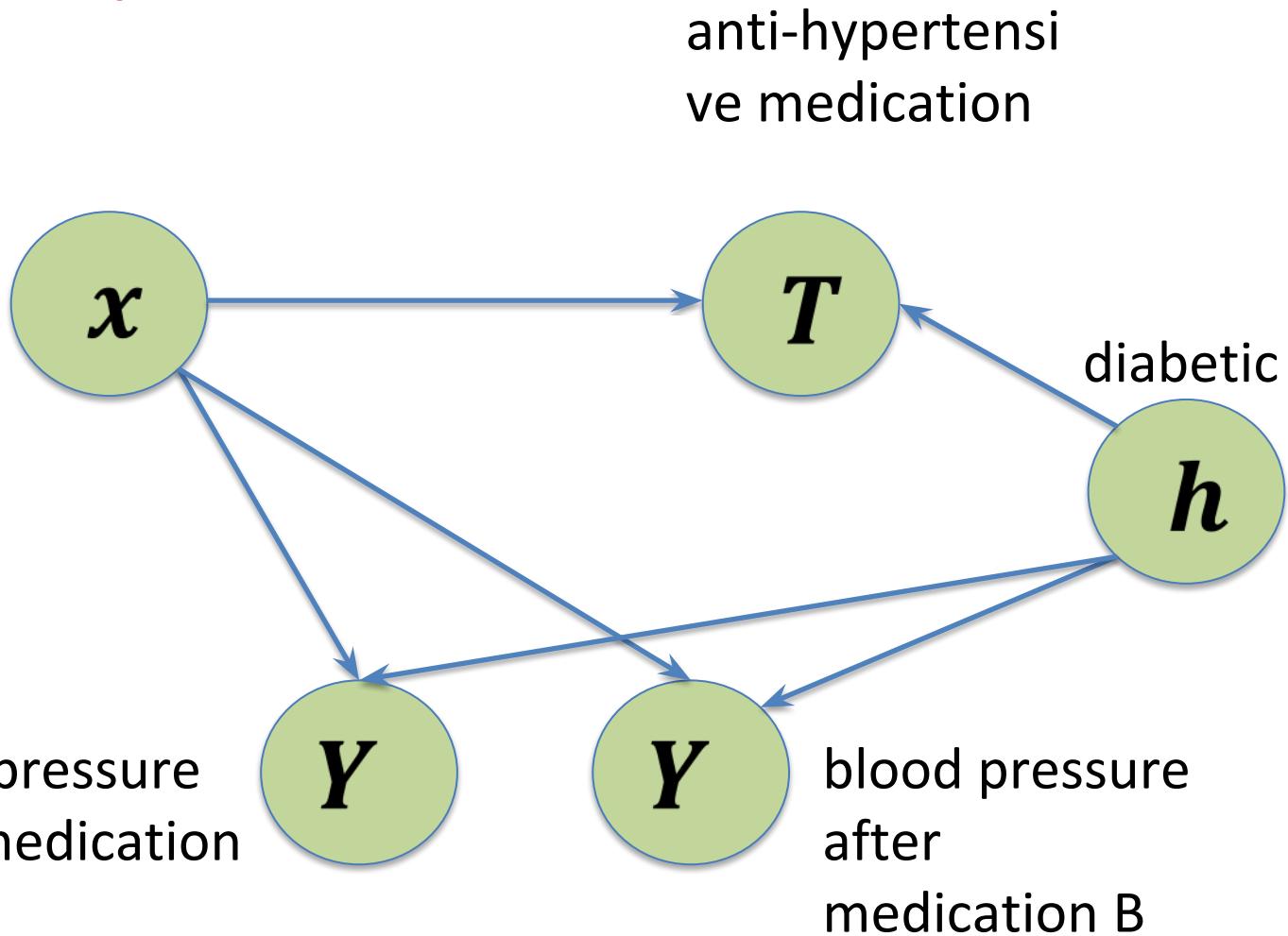
age, gender,
weight, diet,
heart rate at
rest,...



$$(Y_0, Y_1) \perp\!\!\!\perp T \mid x$$

No Ignorability

age, gender,
weight, diet,
heart rate at
rest,...



$$(Y_0, Y_1) \not\perp\!\!\!\perp T \mid x$$

Typical assumption – common support

Y_0, Y_1 : potential outcomes for control and treated

x : unit covariates (features)

T : treatment assignment

We assume:

$$p(T = t | X = x) > 0 \quad \forall t, x$$

Framing the question

1. Where could we go to for data to answer these questions?
2. What should X , T , and Y be to satisfy ignorability?
3. What is the specific causal inference question that we are interested in?
4. Are you worried about common support?

Outline for lecture

- How to recognize a causal inference problem
- Potential outcomes framework
 - Average treatment effect (ATE)
 - Conditional average treatment effect (CATE)
- **Algorithms for estimating ATE and CATE**

Average Treatment Effect

The expected causal effect of T on Y :

$$ATE := \mathbb{E}[Y_1 - Y_0]$$

Average Treatment Effect – the adjustment formula

- Assuming ignorability, we will derive the *adjustment formula* (Hernán & Robins 2010, Pearl 2009)
- The adjustment formula is extremely useful in causal inference
- Also called *G-formula*

Average Treatment Effect

The expected causal effect of T on Y :

$$ATE := \mathbb{E}[Y_1 - Y_0]$$

Average Treatment Effect

The expected causal effect of T on Y :

$$ATE := \mathbb{E} [Y_1 - Y_0]$$

$$\mathbb{E} [Y_1] = \text{law of total expectation}$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_1 \sim p(Y_1|x)} [Y_1|x]] =$$

Average Treatment Effect

The expected causal effect of T on Y :

$$ATE := \mathbb{E} [Y_1 - Y_0]$$

$$\mathbb{E} [Y_1] = \mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_1 \sim p(Y_1|x)} [Y_1 | x]] = \mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_1 \sim p(Y_1|x)} [Y_1 | x, T = 1]] =$$

ignorability
 $(Y_0, Y_1) \perp\!\!\!\perp T | x$

Average Treatment Effect

The expected causal effect of T on Y :

$$ATE := \mathbb{E} [Y_1 - Y_0]$$

$$\mathbb{E} [Y_1] =$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_1 \sim p(Y_1|x)} [Y_1 | x]] =$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_1 \sim p(Y_1|x)} [Y_1 | x, T = 1]] =$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E} [Y_1 | x, T = 1]]$$

shorter notation

Average Treatment Effect

The expected causal effect of T on Y :

$$ATE := \mathbb{E} [Y_1 - Y_0]$$

$$\mathbb{E} [Y_0] =$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_0 \sim p(Y_0|x)} [Y_0|x]] =$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E}_{Y_0 \sim p(Y_0|x)} [Y_0|x, T = 1]] =$$

$$\mathbb{E}_{x \sim p(x)} [\mathbb{E} [Y_0|x, T = 0]]$$

The adjustment formula

Under the assumption of ignorability, we have that:

$$ATE = \mathbb{E}[Y_1 - Y_0] = \\ \mathbb{E}_{x \sim p(x)} [\mathbb{E}[Y_1 | x, T = 1] - \mathbb{E}[Y_0 | x, T = 0]]$$

$$\left. \begin{array}{l} \mathbb{E}[Y_1 | x, T = 1] \\ \mathbb{E}[Y_0 | x, T = 0] \end{array} \right\}$$

Quantities we
can estimate
from data

The adjustment formula

Under the assumption of ignorability, we have that:

$$ATE = \mathbb{E} [Y_1 - Y_0] =$$

$$\mathbb{E}_{x \sim p(x)} [\textcolor{red}{\mathbb{E} [Y_1 | x, T = 1]} - \textcolor{blue}{\mathbb{E} [Y_0 | x, T = 0]}]$$

$$\left. \begin{array}{l} \mathbb{E} [Y_0 | x, T = 1] \\ \mathbb{E} [Y_1 | x, T = 0] \\ \mathbb{E} [Y_0 | x] \\ \mathbb{E} [Y_1 | x] \end{array} \right\}$$

Quantities we
cannot directly
estimate from data

The adjustment formula

Under the assumption of ignorability, we have that:

$$ATE = \mathbb{E} [Y_1 - Y_0] =$$

$$\mathbb{E}_{x \sim p(x)} [\underbrace{\mathbb{E} [Y_1 | x, T = 1] - \mathbb{E} [Y_0 | x, T = 0]}_{\text{}}]$$

$$\left. \begin{array}{l} \mathbb{E} [Y_1 | x, T = 1] \\ \mathbb{E} [Y_0 | x, T = 0] \end{array} \right\}$$

Quantities we
can estimate
from data

Many methods!

Covariate adjustment

Propensity score re-weighting

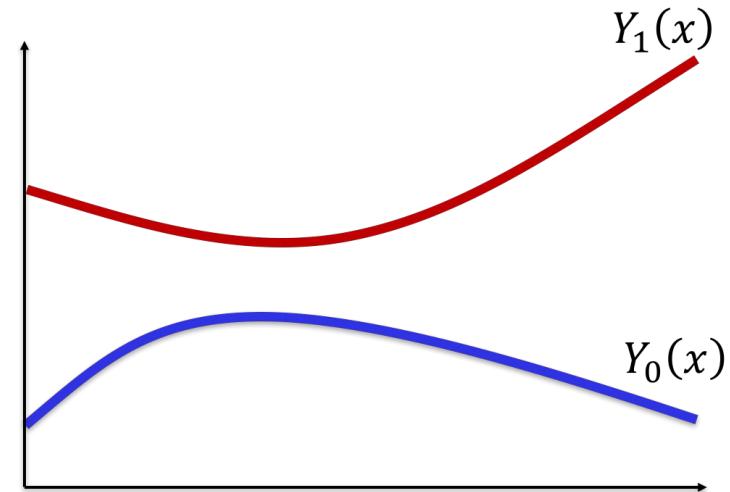
Doubly robust estimators

Matching

...

Covariate adjustment

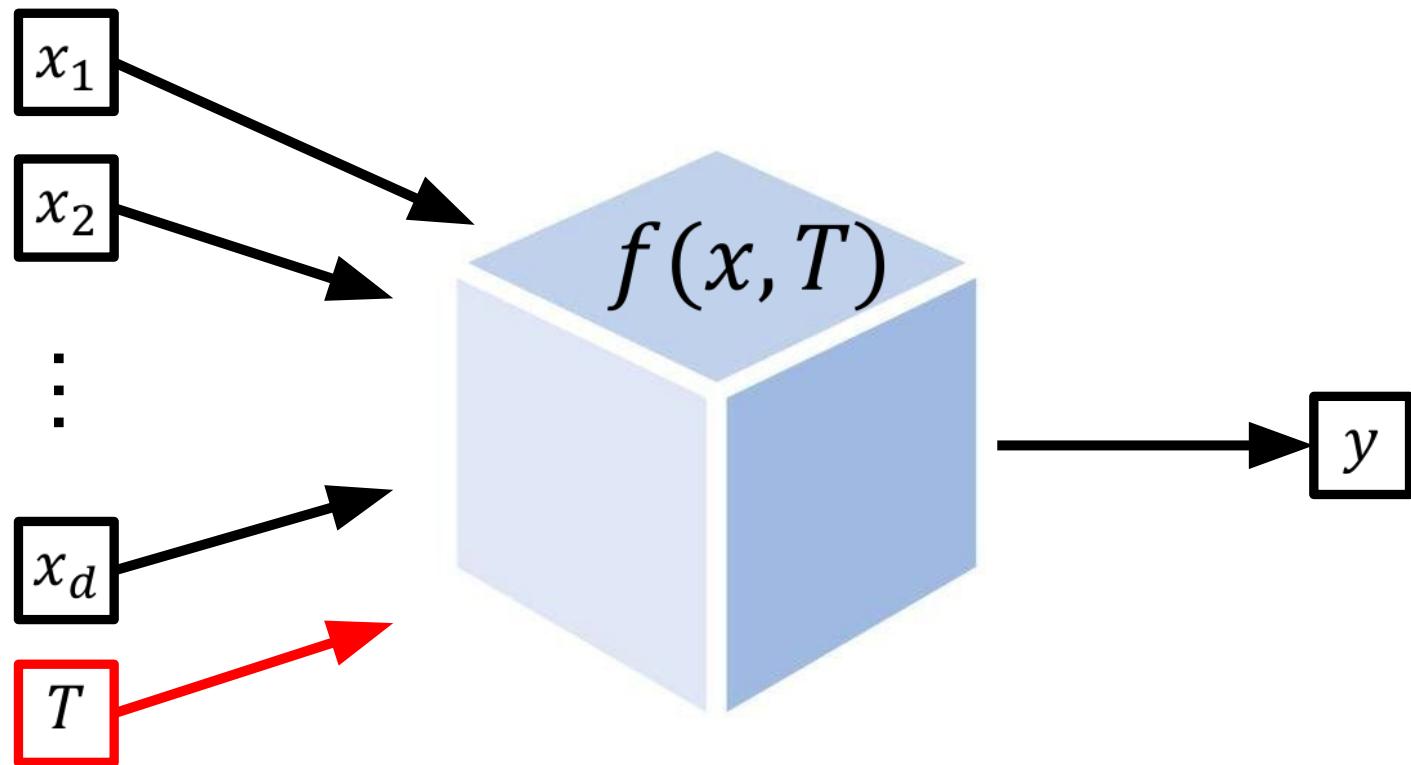
- Explicitly model the relationship between treatment, confounders, and outcome
- Also called “Response Surface Modeling”
- Used for both ITE and ATE
- A regression problem



Covariates
(Features)

Regression
model

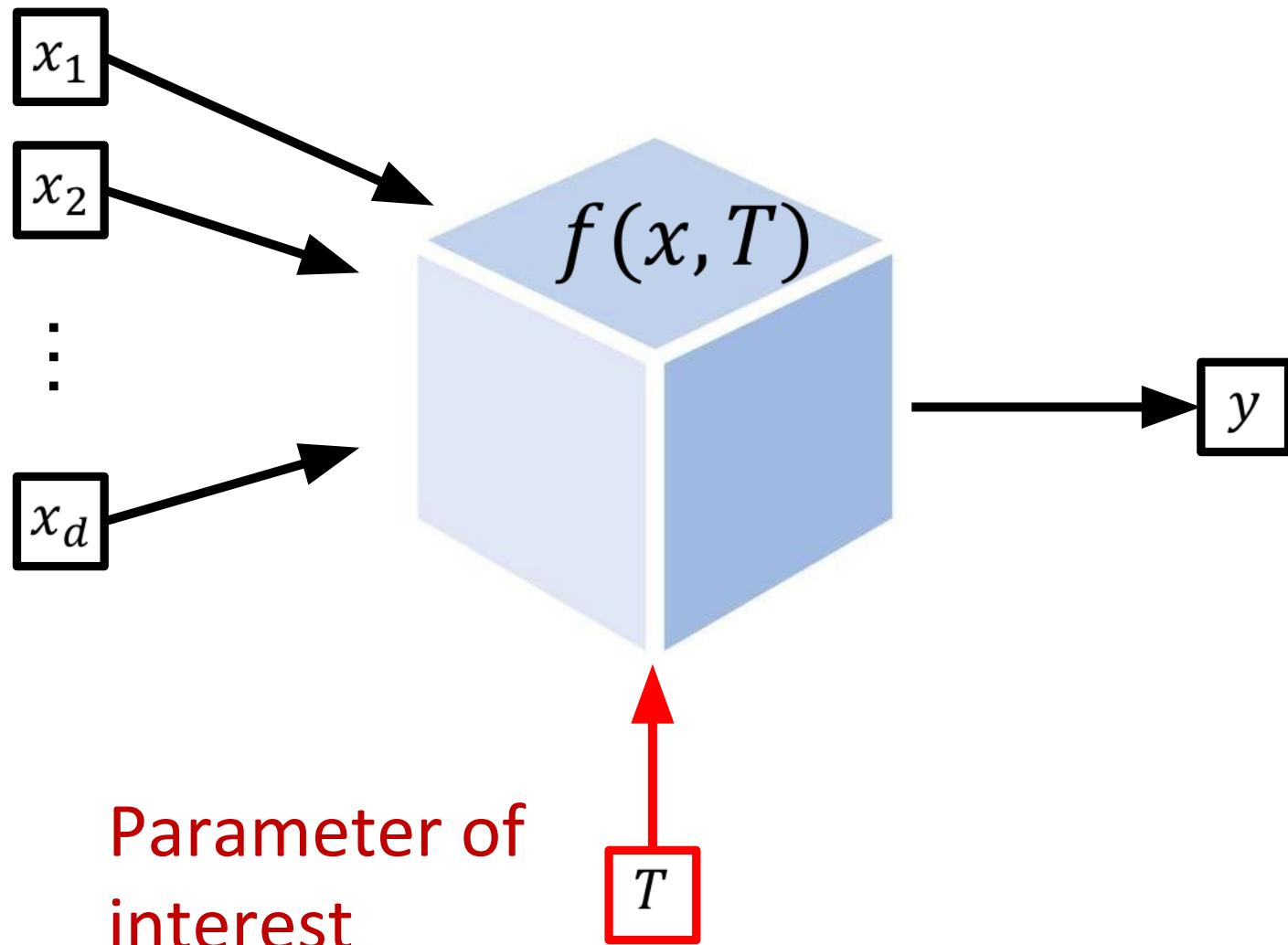
Outcome



Nuisance
Parameters

Regression
model

Outcome



Parameter of
interest

Covariate adjustment (parametric g-formula)

- Explicitly model the relationship between treatment, confounders, and outcome
- Under ignorability, the expected causal effect of T on Y :

$$\mathbb{E}_{x \sim p(x)} [\textcolor{red}{\mathbb{E}[Y_1 | T = 1, x]} - \textcolor{blue}{\mathbb{E}[Y_0 | T = 0, x]}]$$

- Fit a model $f(x, t) \approx \mathbb{E}[Y_t | T = t, x]$

$$\widehat{ATE} = \frac{1}{n} \sum_{i=1}^n f(x_i, 1) - f(x_i, 0)$$

Covariate adjustment (parametric g-formula)

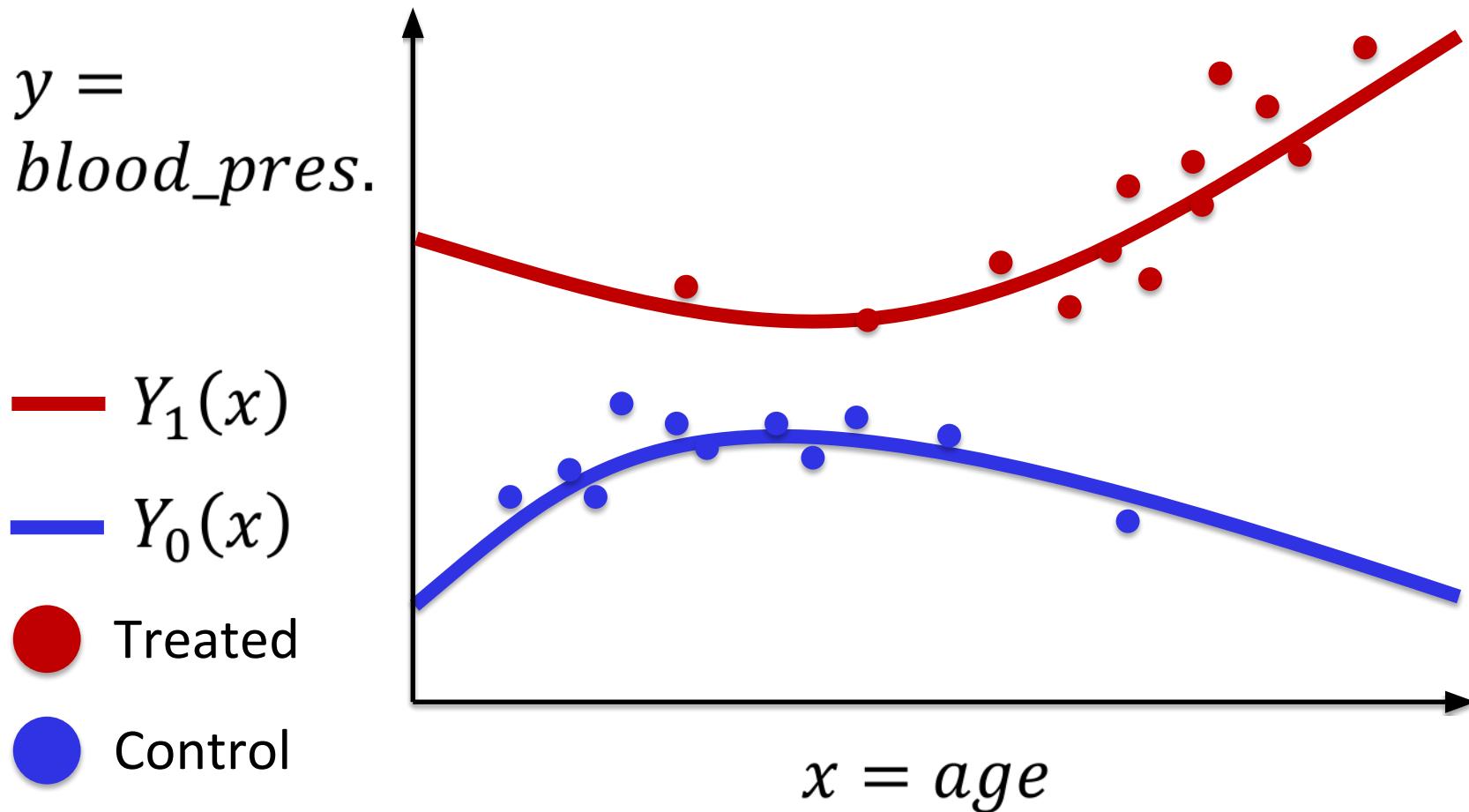
- Explicitly model the relationship between treatment, confounders, and outcome
- Under ignorability, the expected causal effect of T on Y :

$$\mathbb{E}_{x \sim p(x)} \left[\mathbb{E}[Y_1 | T = 1, x] - \mathbb{E}[Y_0 | T = 0, x] \right]$$

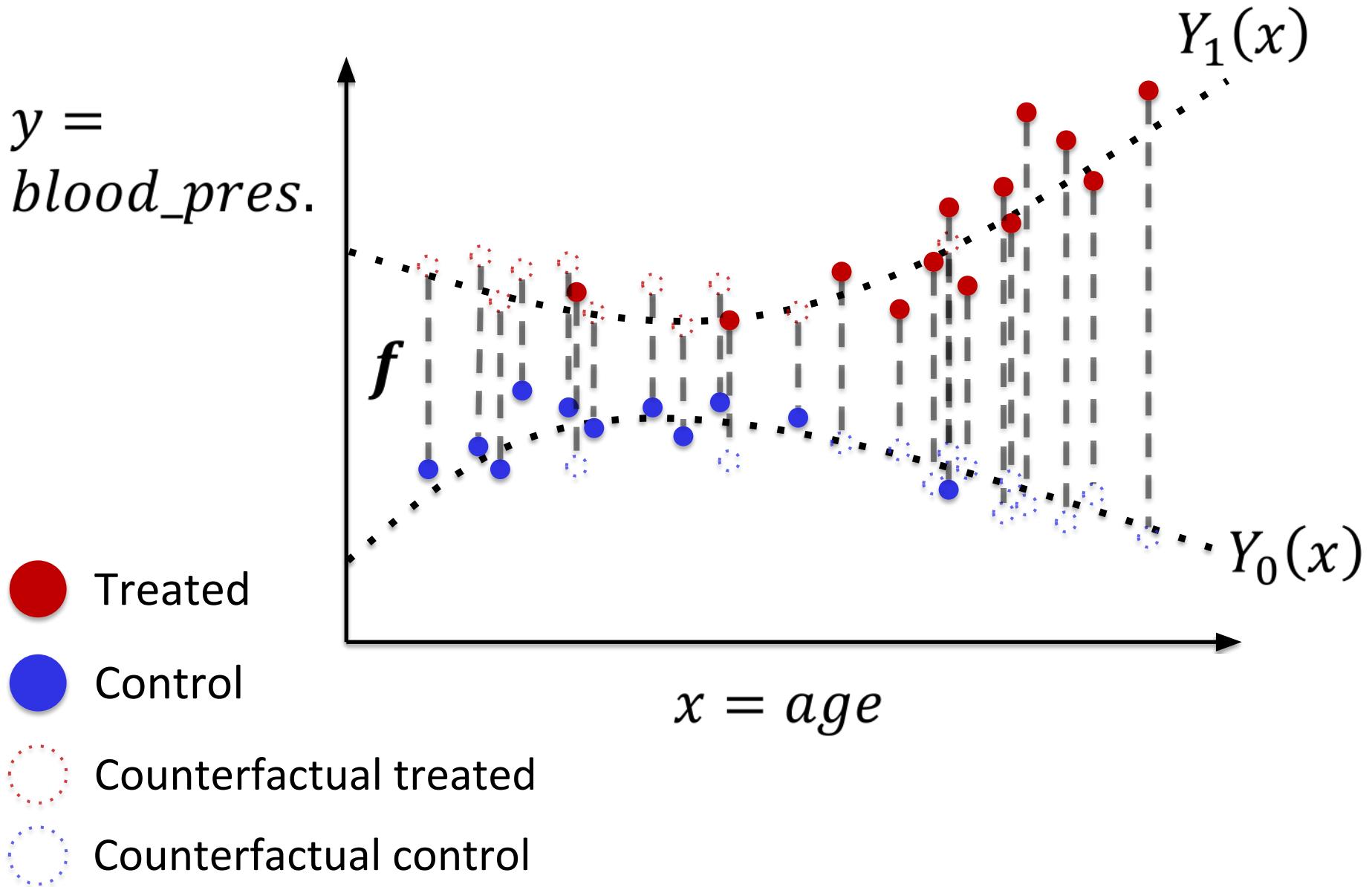
- Fit a model $f(x, t) \approx \mathbb{E}[Y_t | T = t, x]$

$$\widehat{CATE}(\nu.) = f(\nu. 1) - f(\nu. 0)$$

Covariate adjustment



Covariate adjustment



Example of how covariate adjustment fails when there is no overlap

