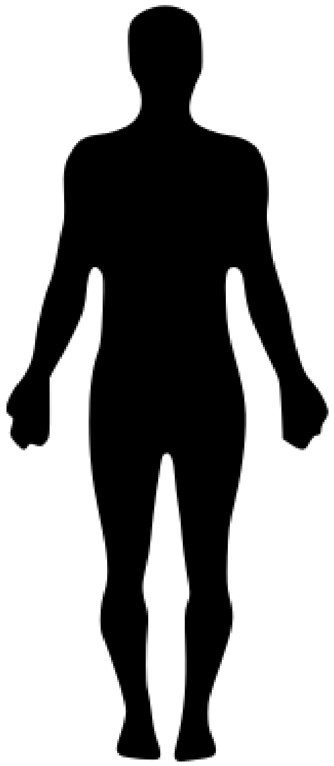
Learning Treatment Effects from Multiple Data

Sofia Triantafillou

Assistant Professor, Department of Mathematics and Applied Mathematics, University of Crete, Greece

Affiliated Researcher, Institute of Applied and Computational Mathematics, FORTH-GREECE

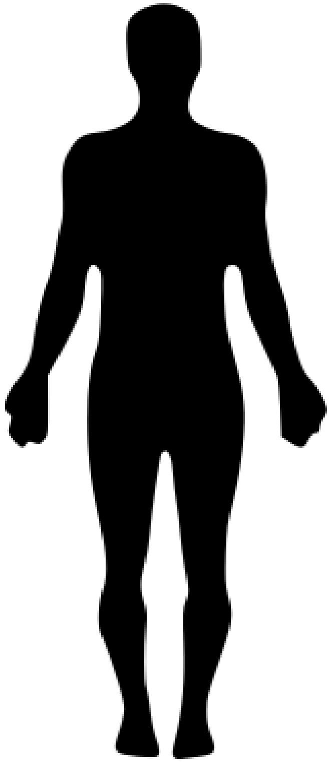
Association



You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD?

Association

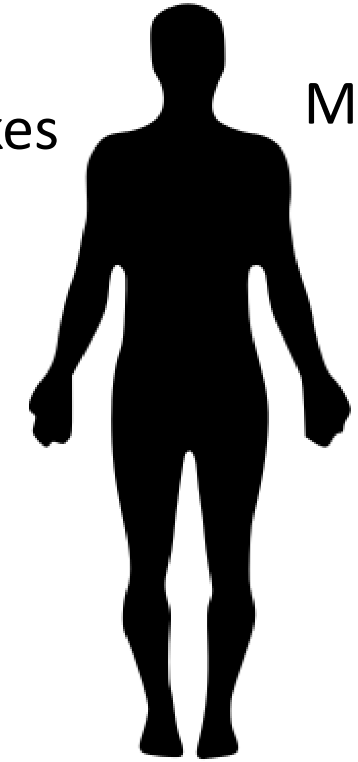


You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD?

P(COPD) = 4.8%

Association

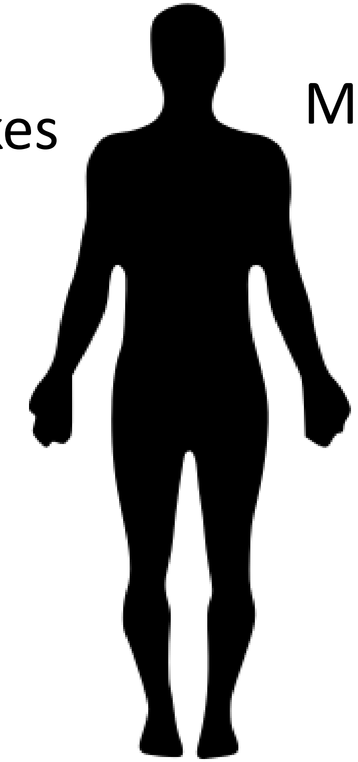


You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD given that they have blonde hair, a mild cough, and they smoke?

P(COPD) = 4.8%

Association



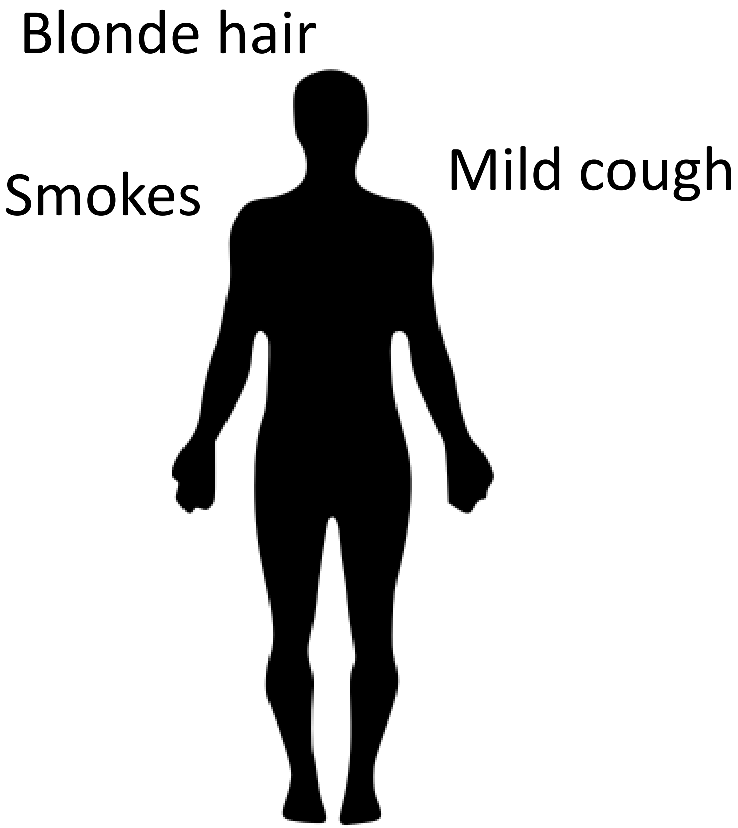
You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD given that they have blonde hair, a mild cough, and they smoke?

P(COPD) = 4.8%

P(COPD|Blonde hair) = 4.8%

Association



You are a pulmonologist, and you are offering free screenings in your community

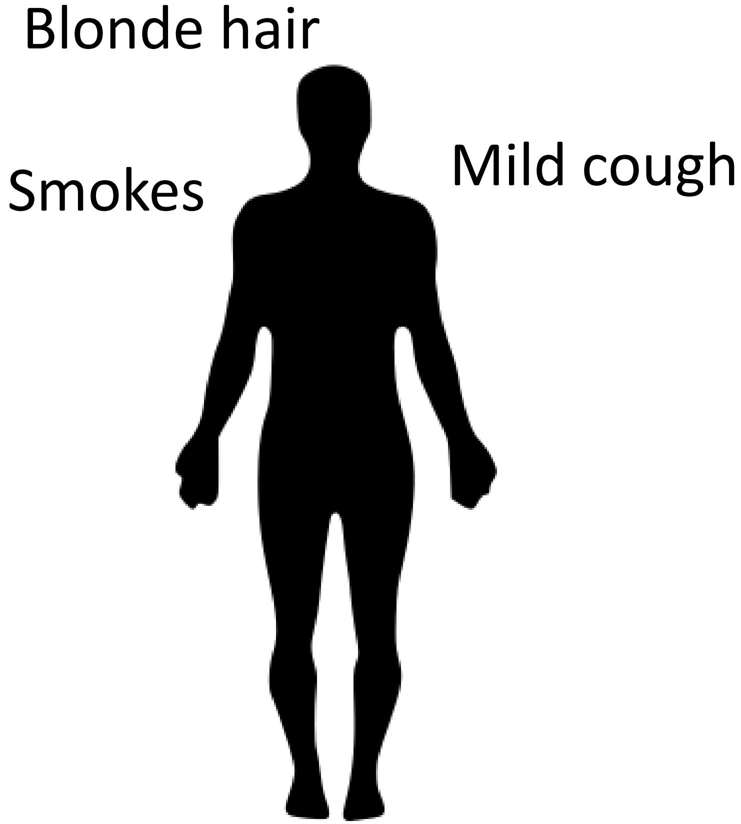
What is the probability that they have COPD given that they have blonde hair, a mild cough, and they smoke?

P(COPD) = 4.8%

P(COPD|Blonde hair) = 4.8%

P(COPD|Cough) = 7.2%

Association



You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD given that they have blonde hair, a mild cough, and they smoke?

P(COPD) = 4.8%

P(COPD|Blonde hair) = 4.8%

P(COPD|Cough) = 7.2%

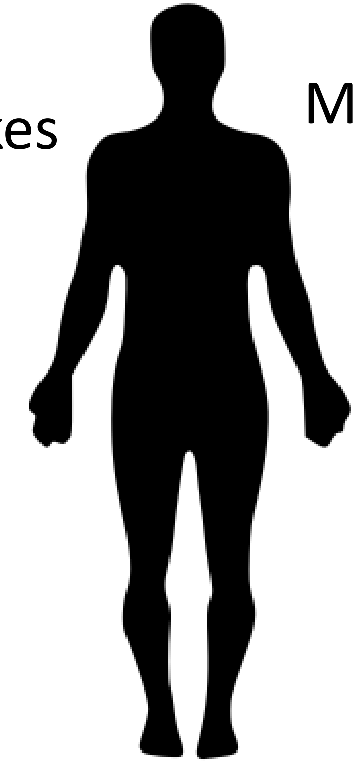
P(COPD|Smoking) = 15.2%

Association

Blonde hair

Smokes

Mild cough



You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD given that they have blonde hair, a mild cough, and they smoke?

P(COPD) = 4.8%

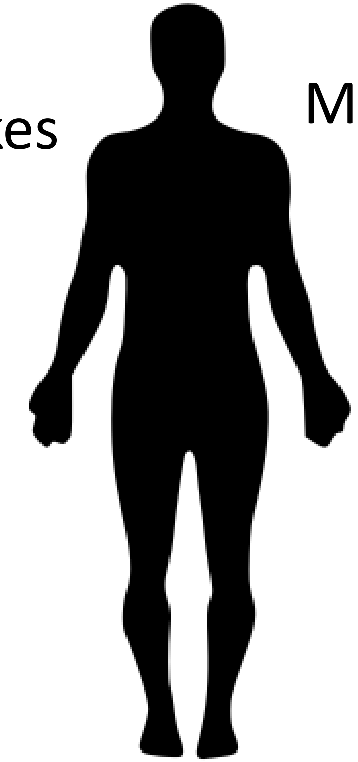
P(COPD|Blonde hair) = 4.8%

P(COPD|Cough) = 7.2%

P(COPD|Smoking) = 15.2%

P(COPD|Smoking, Cough) = 25.3%

Association



You are a pulmonologist, and you are offering free screenings in your community

What is the probability that they have COPD given that they have blonde hair, a mild cough, and they smoke?

P(COPD) = 4.8%

P(COPD|Blonde hair) = 4.8%

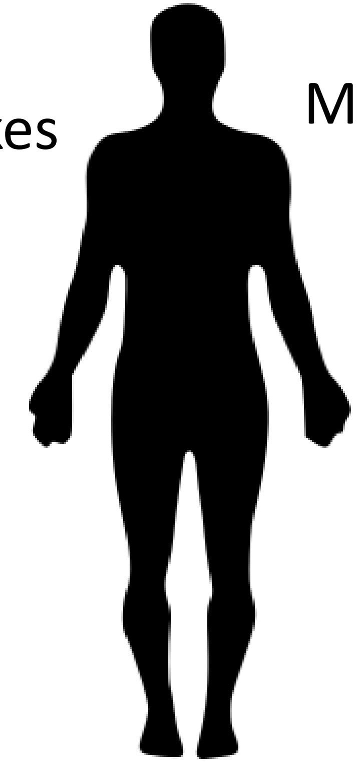
P(COPD|Cough) = 7.2%

P(COPD|Smoking) = 15.2%

P(COPD|Smoking, Cough) = 25.3%

Association can improve your prediction for a variable of interest.

Association vs Causality



- Had they not been smoking, would they have gotten COPD?
- If you make them stop smoking, will they stop having COPD?
- If you give them anti-cough medication, will they stop having COPD?

Why Causality

Association is NOT causation.

We can use association to predict what will happen if we observe something.

We need causation:

- To understand **why** something happened.
- To predict **what will happen** if **we intervene** and change something.

Association is NOT causality

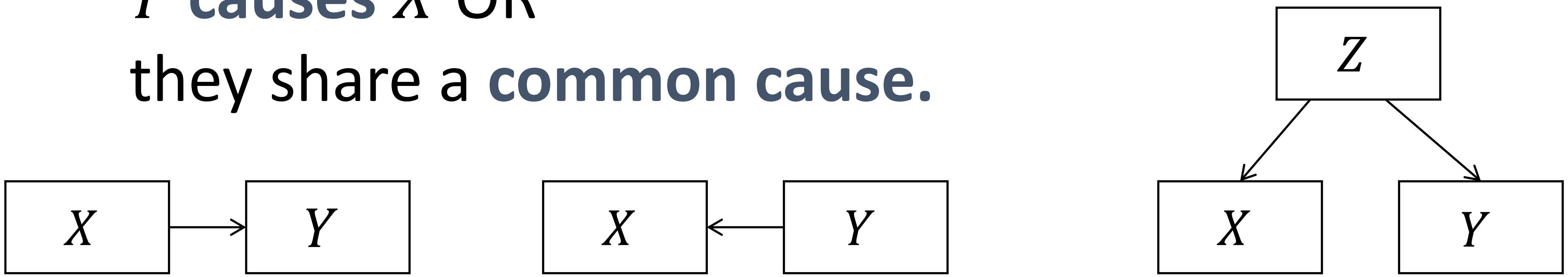
- In this talk (and often in science), we are talking about probabilistic causality: Setting the value of the cause changes the probability distribution of the effect (outcome).

- Association: What is $P(Y= y|X =x)$?
- Causality: What is $P(Y=y|do(X=x))$?

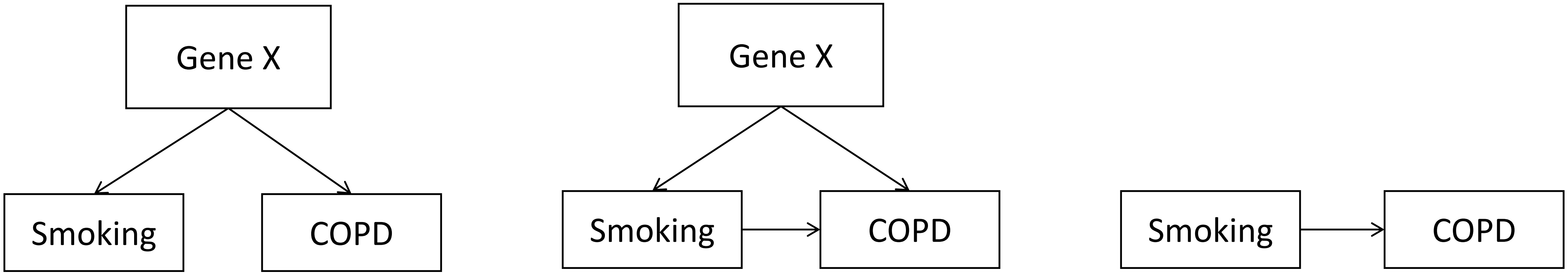
BUT: No association without causation

- Reichenbach's common cause principle:

If X and Y are correlated,
X **causes** Y OR
Y **causes** X OR
they share a **common cause.**

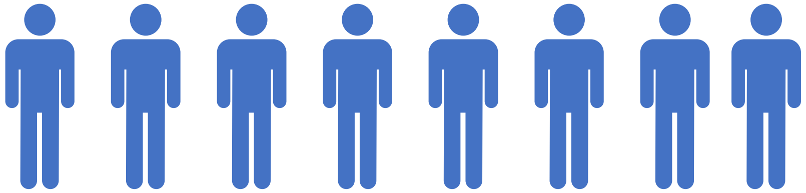


Learning Causality



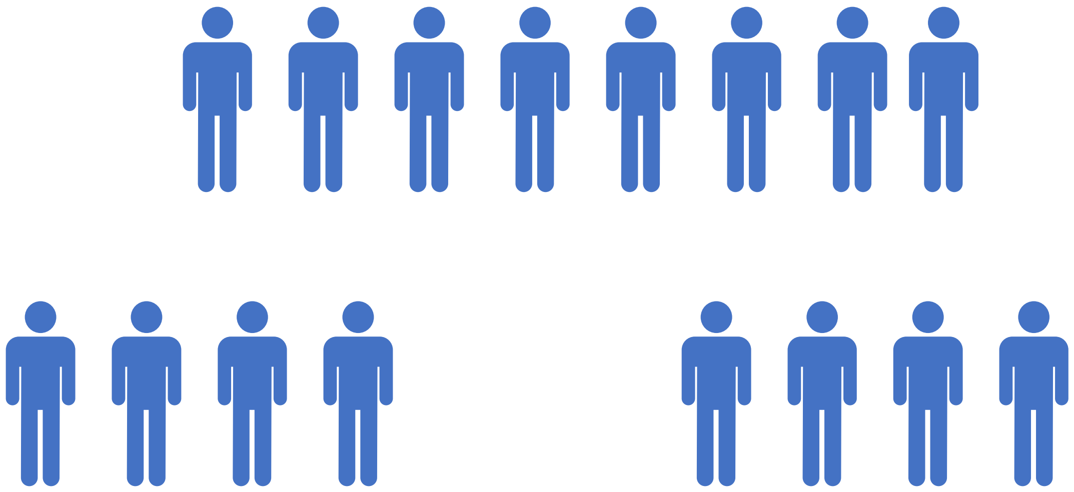
Example: Does smoking cause chronic obstructive pulmonary disease (COPD)?

Learning Causality: Randomized Control Trial



1. Take a **random** sample from your population.

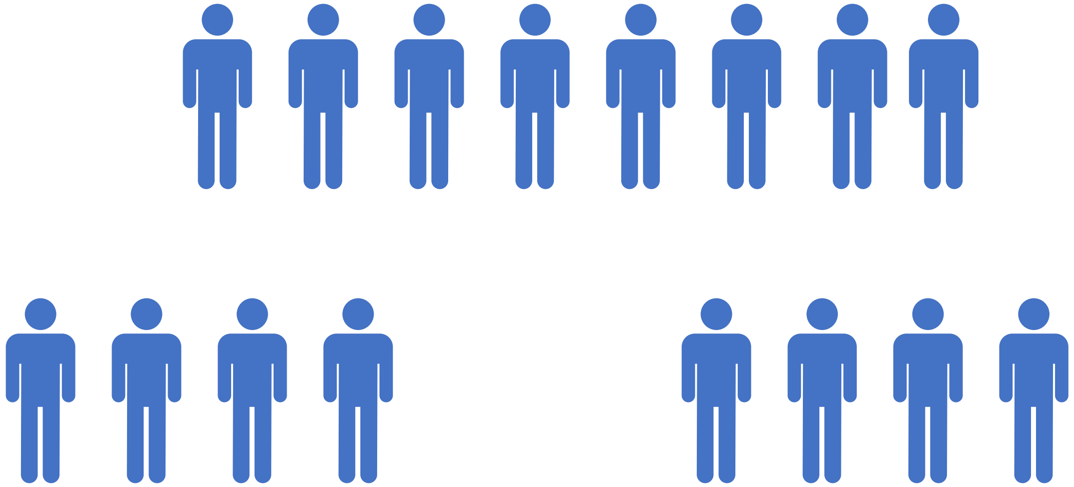
Learning Causality: Randomized Control Trial



1. Take a **random** sample from your population.

2. Randomly split them in **control** and **treatment** groups.

Learning Causality: Randomized Control Trial

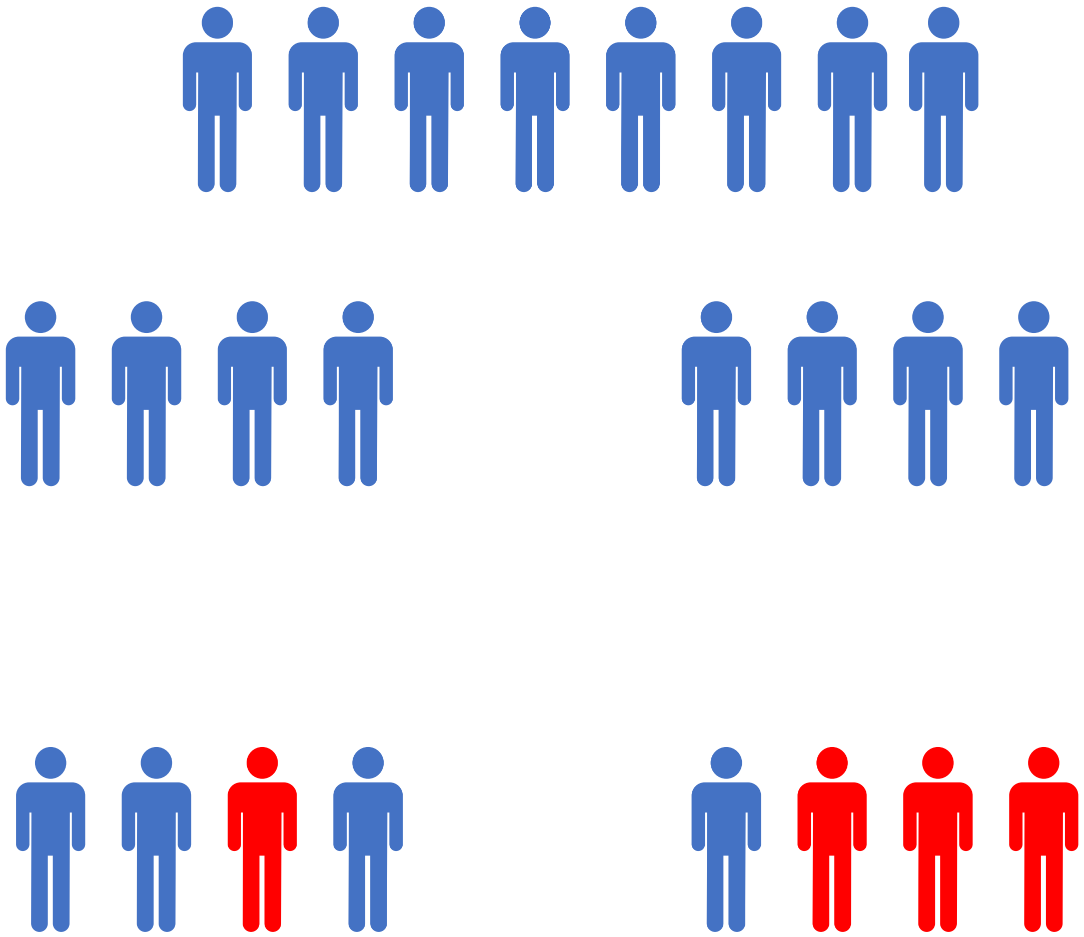


1. Take a **random** sample from your population.

2. Randomly split them in **control** and **treatment** groups.

3. Force control group **not to smoke**, force treatment group to **smoke.**

Learning Causality: Randomized Control Trial



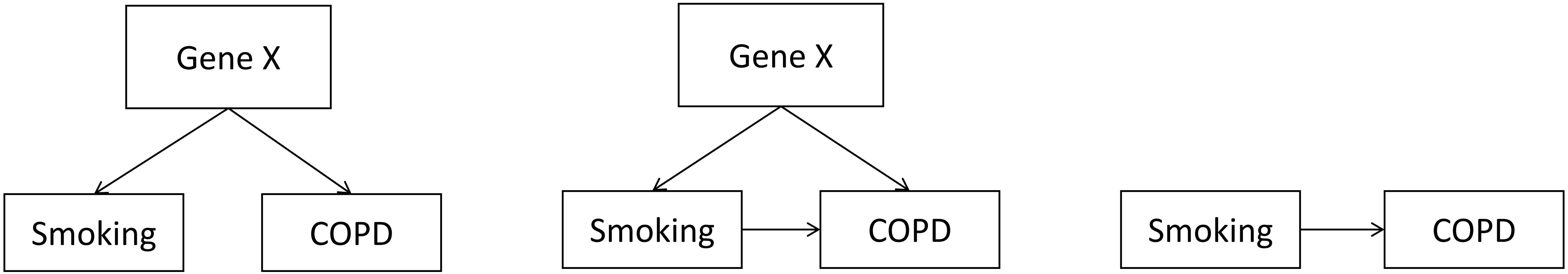
1. Take a **random** sample from your population.

2. Randomly split them in **control** and **treatment** groups.

3. Force control group **not to smoke**, force treatment group to **smoke.**

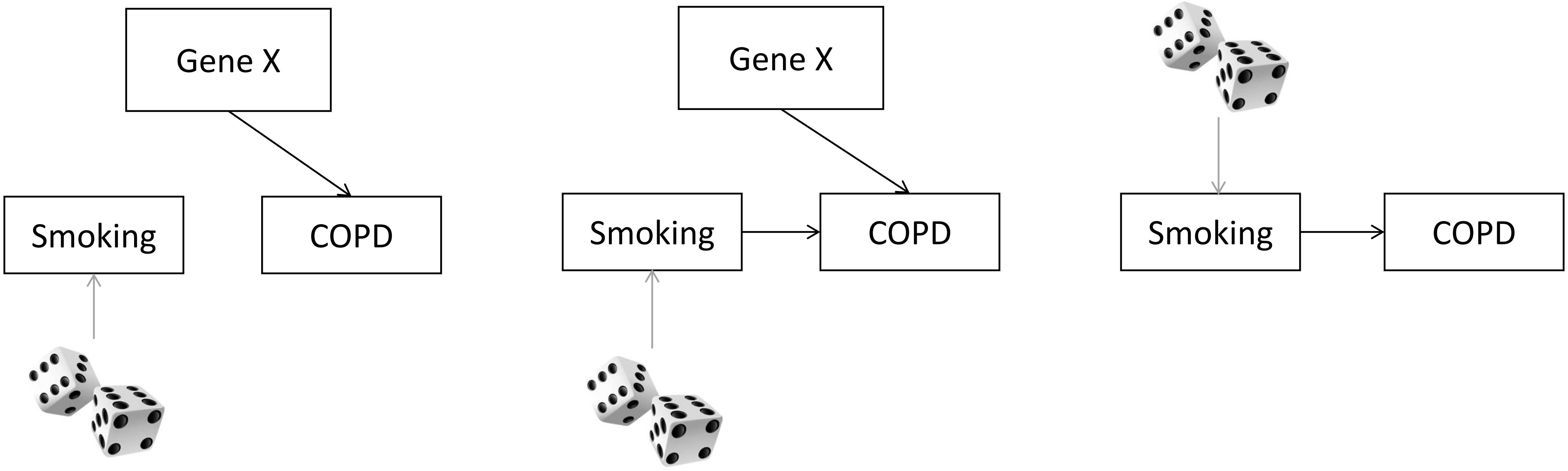
4. Wait a few years to see who have developed **COPD**.

Randomization



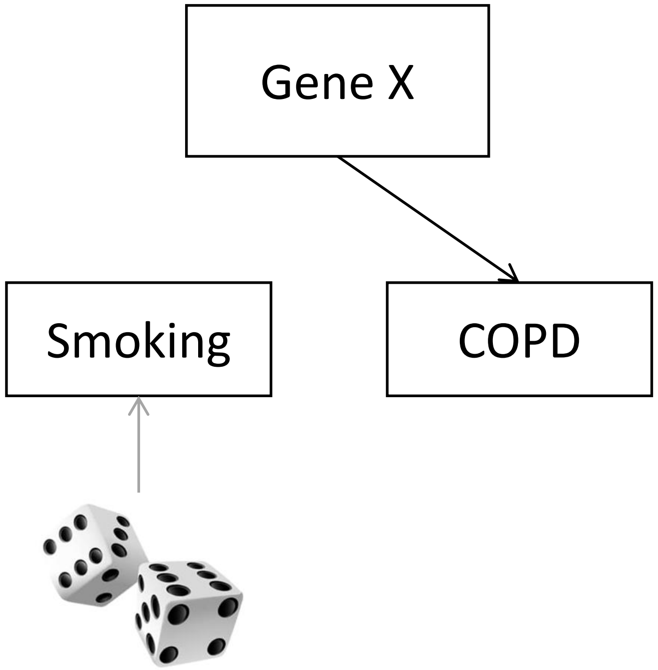
Models considered possible

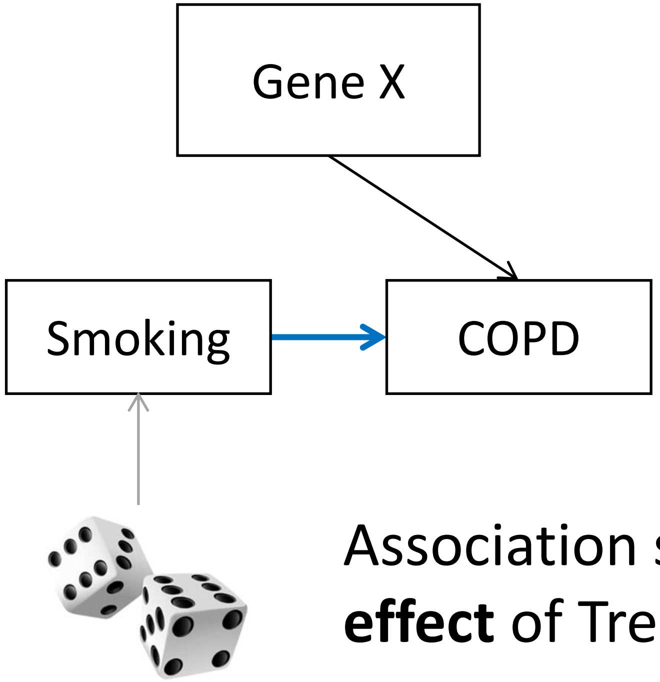
Randomization

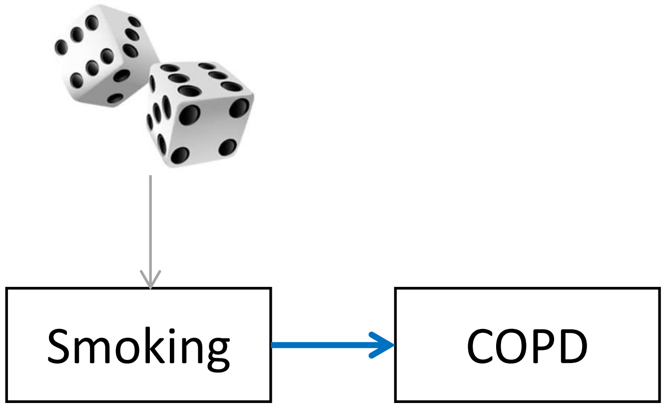


Randomization

Association persists only when the relationship is causal.

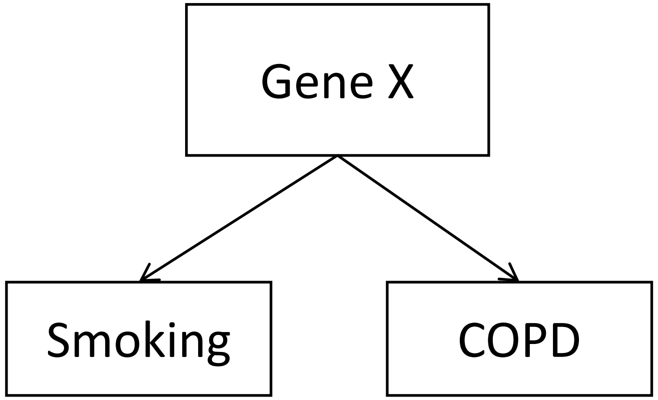






Association strength quantifies the **causal effect** of Treatment on Outcome

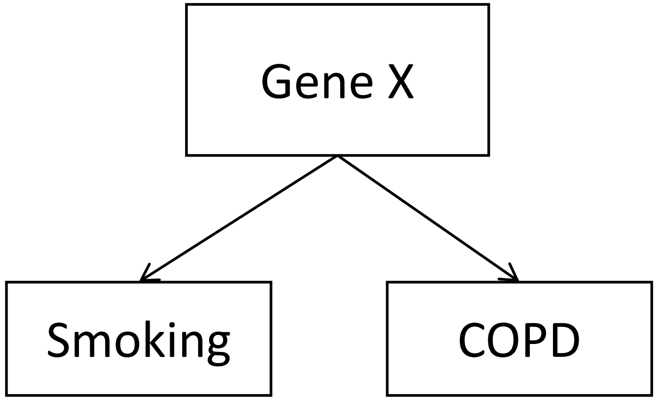
Causal Graphical Models



Causal Markov Condition:
A variable is independent of its non-effects given its direct causes

Causal Graphical Models

Causal graph



Probability distribution

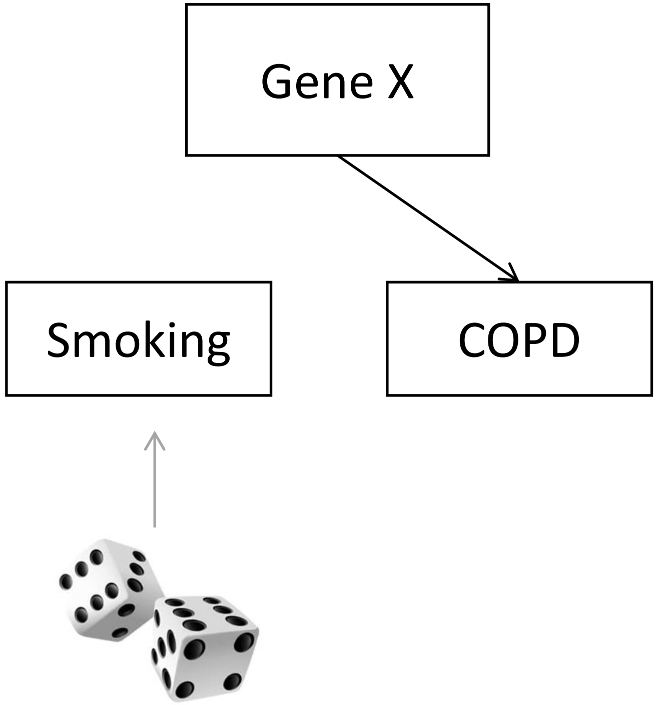
P(Smoking, Gene X, COPD) =
P(Gene X)*P(Smoking|Gene X)*P(COPD|Gene X)

Gene X mutation	Smoking	COPD
No	Yes	No
Yes	No	Yes
No	No	No

Causal Markov Condition:
A variable is independent of its non-effects given its direct causes

Causal Graphical Models

Causal graph



Probability distribution

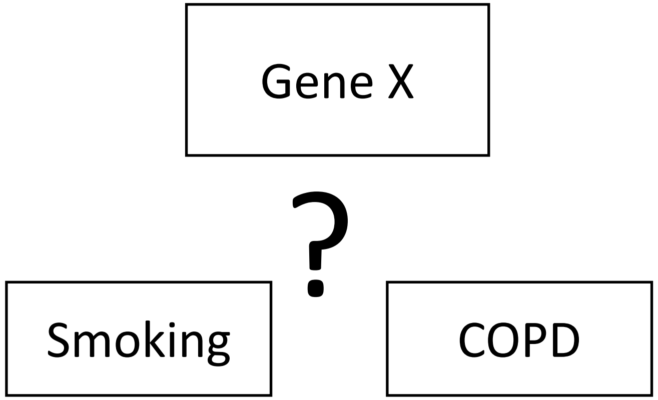
P(do(Smoking), Gene X, COPD) =
P(Gene X)*P(do(Smoking))*P(COPD|Gene X)

Gene X mutation	Smoking	COPD
No	Yes	No
Yes	No	Yes
No	No	No

If you know the causal graph, you can predict the effects of interventions

Learning Causality

Causal graph



Probability distribution

P(Smoking, Gene X, COPD) =
P(Gene X)*P(Smoking|Gene X)*P(COPD|Gene X)

Gene X mutation	Smoking	COPD
No	Yes	No
Yes	No	Yes
No	No	No

The field of causal learning (causal discovery and inference) is about learning causal structure and quantifying causal effects from limited or no interventions.

Causal learning from observational vs experimental data

Observational data
(e.g., health record data)

+ Large sample sizes

+ Sampled from the entire population

- Biased for causal effect estimation due to confounders unless you know the ground truth causal model.

Experimental data
(e.g., randomized control trial data)

-Limited sample sizes

-Sampled from selected subpopulations

+ Unbiased for causal effect estimation – high variance estimates due to low sample size

Causal learning from observational vs experimental data

Observational data (e.g., health record data)	**Experimental data (e.g., randomized control trial data)**
+ Large sample sizes	-Limited sample sizes
+ Sampled from the entire population	-Sampled from selected subpopulations
- Biased for causal effect estimation due to confound unless you know the groun causal model.	+ Unbiased for causal effect variance low sample size

Combining observational and experimental data with causal models can improve effect estimation and get the best of both worlds.

Summary

- Causal models allow us to formalize and reason with causal relationships, connect observations and experiments, and automate the scientific process.