Introduction to Causal Inference: Basics and Advances

Hang Wu

Dr. May D. Wang

Dept. Biomedical Engineering, Georgia Tech

Link: https://github.com/hang-wu/CI

Objective and Takeaways

- What is causal inference?
- How to identify causal effect?

Agenda

Break

Part 1

- Introduction of Causal Inference

3:00 – 3:30

Part 2

- Basic of Causal Effect Estimation Algorithm

3:30 – 4:15

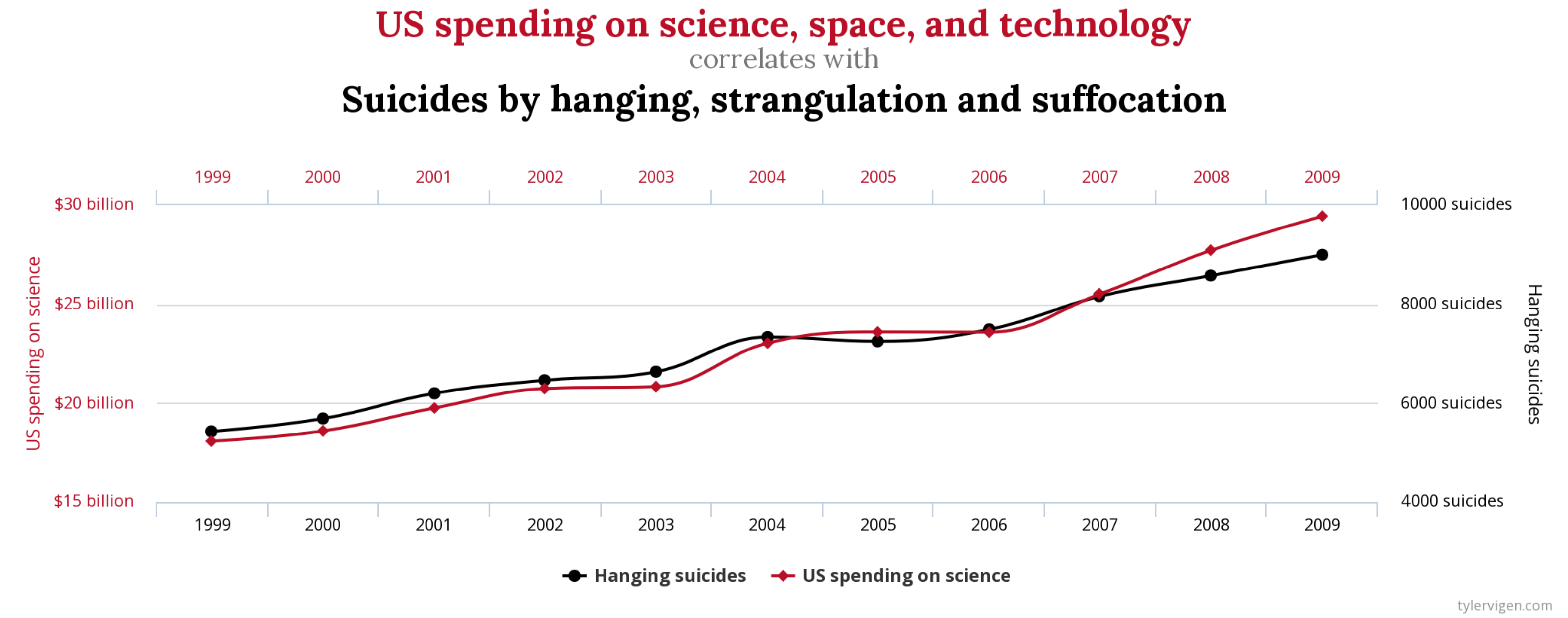
Part 3

- Recent Advances
- Challenges and Opportunities

4:35 – 5:15

I'll attend the NSF student award from 4:20 – 4:30 CNB-MAC

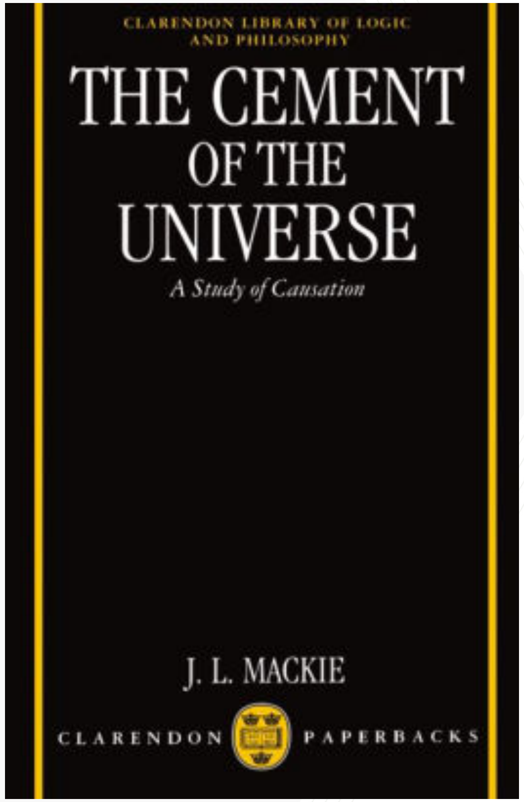
Correlation vs. causality

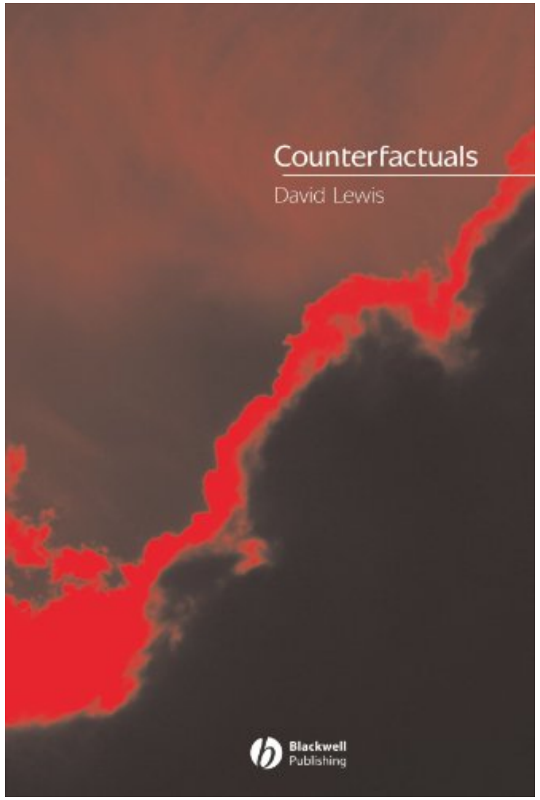


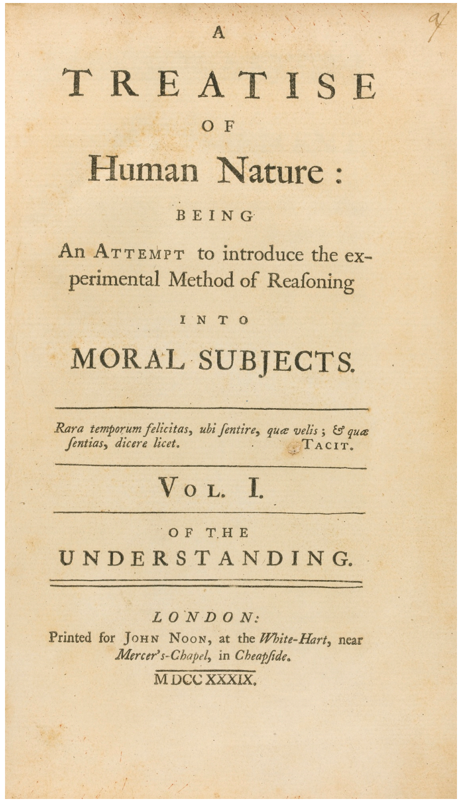
[Figure: https://www.tylervigen.com/spurious-correlations](https://www.tylervigen.com/spurious-correlations)

What is causality?

The study on causality






A
TREATISE
OF
Human Nature:
BEING
An ATTEMPT to introduce the experimental Method of Reaſoning
INTO
MORAL SUBJECTS.

Rara temporum felicitas, ubi ſentire, quæ velis; & quæ ſentias, dicere licet. TACIT.

VOL. I.
OF THE
UNDERSTANDING.

LONDON:
Printed for JOHN NOON, at the *White-Hart*, near *Mercer's-Chapel*, in *Cheapſide*.
MDCCXXXIX.

......

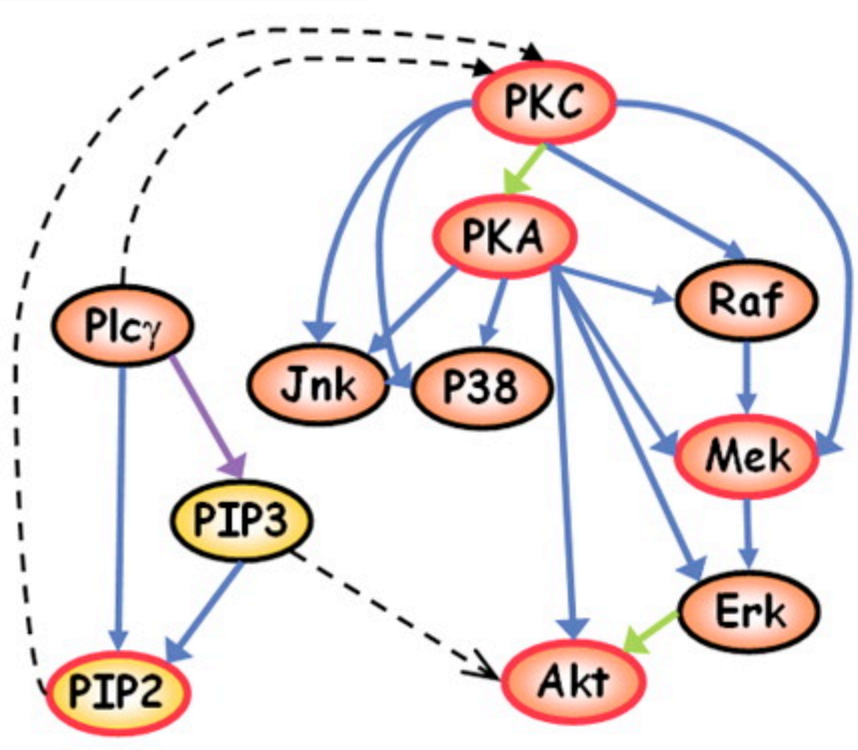
A philosophical debate since the start of human knowledge….

Our Definitions

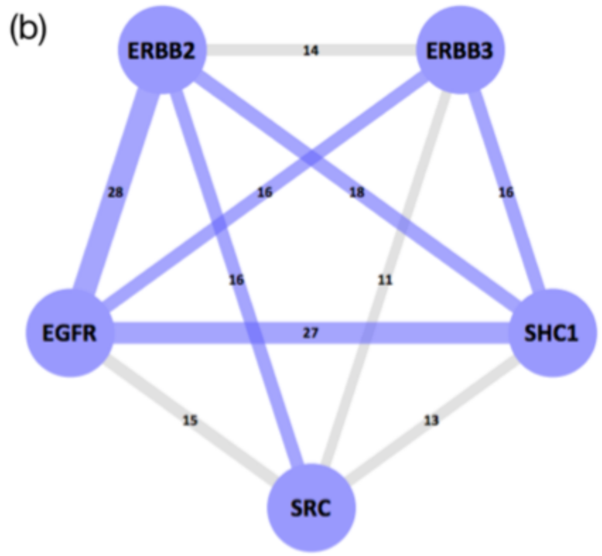
- X causes Y iff
- changing X leads to a change in Y,
- while keeping everything else constant.
- The magnitude of change in Y when changing X by 1 unit is the causal effect of X on Y.

Causality in Bio/medical Research

- Protein-Signaling Networks



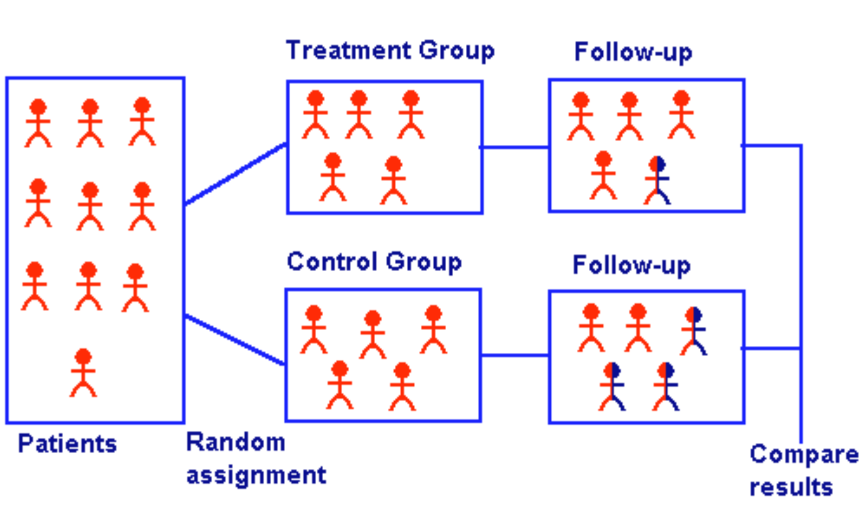
- Pan-cancer Pathways



1. Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A., and Nolan, G. P. (2005). Causal protein-signaling networks derived from multiparameter single-cell data. Science 308, 523–529. doi: 10.1126/science.1105809
2. Ha, M. J., Banerjee, S., Akbani, R., Liang, H., Mills, G. B., Do, K. A., & Baladandayuthapani, V. (2018). Personalized Integrated Network Modeling of the Cancer Proteome Atlas. *Scientific reports*, *8*(1), 14924.

Causality in Bio/medical Research

- Randomized Controlled Trials/ Epidemiology
- Dynamic Treatment Regimes



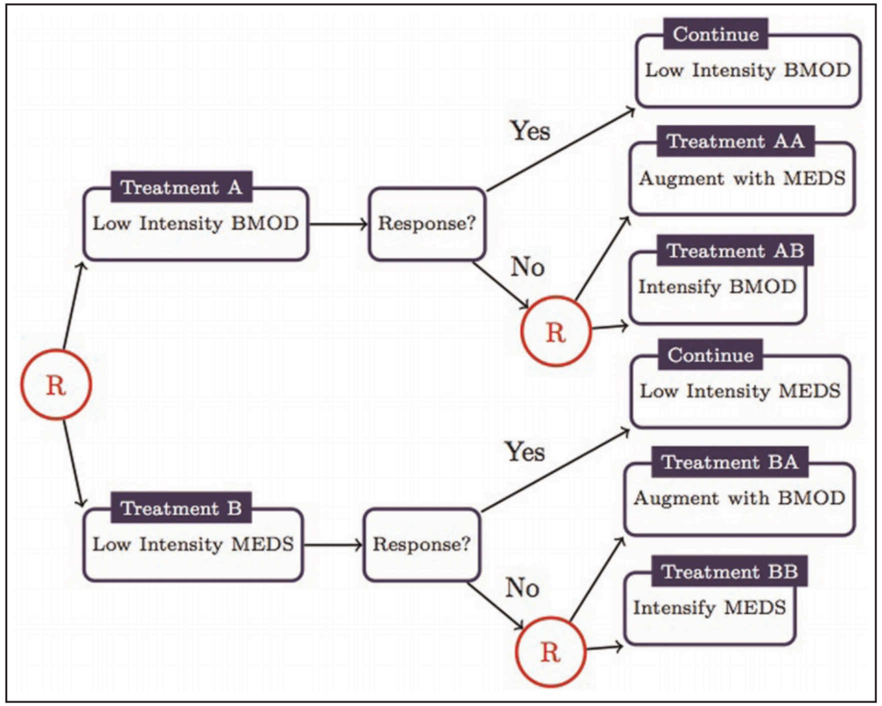
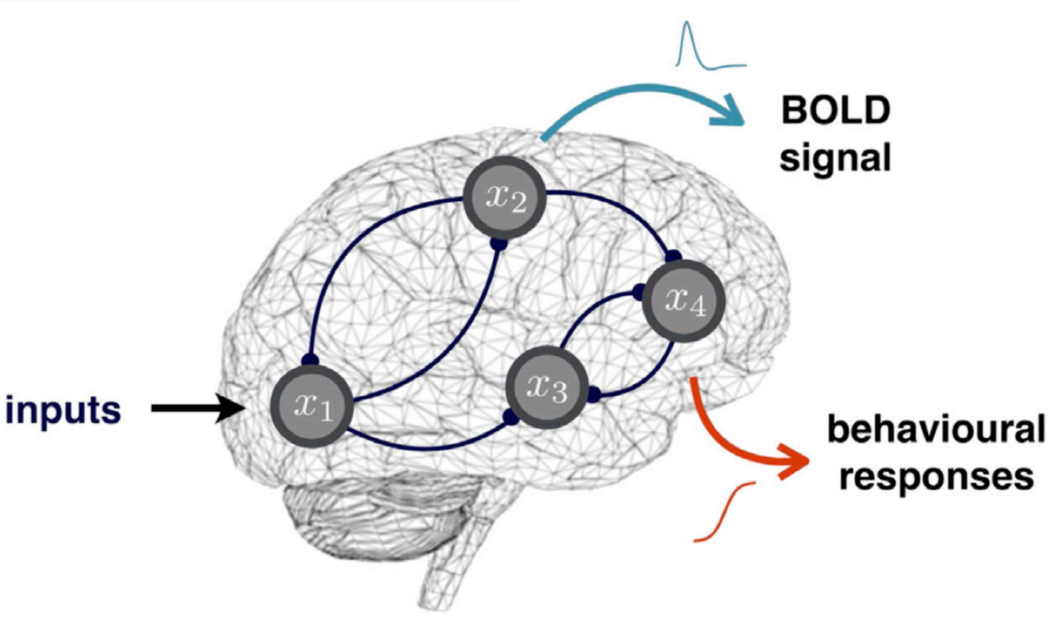
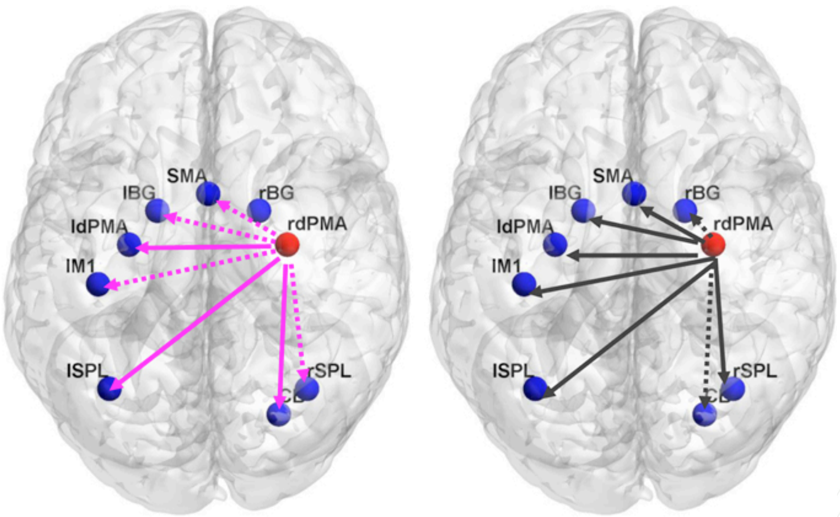


Figure: https://www.statisticshowto.datasciencecentral.com/experimental-design/randomized-clinical-trial-rcts/
Zhao, Ying-Qi, and Eric B. Laber. "Estimation of optimal dynamic treatment regimes." Clinical Trials 11.4 (2014): 400-407.

Dynamics of Causality

- Dynamic Causal Modeling[1]
- Granger Causality[2]

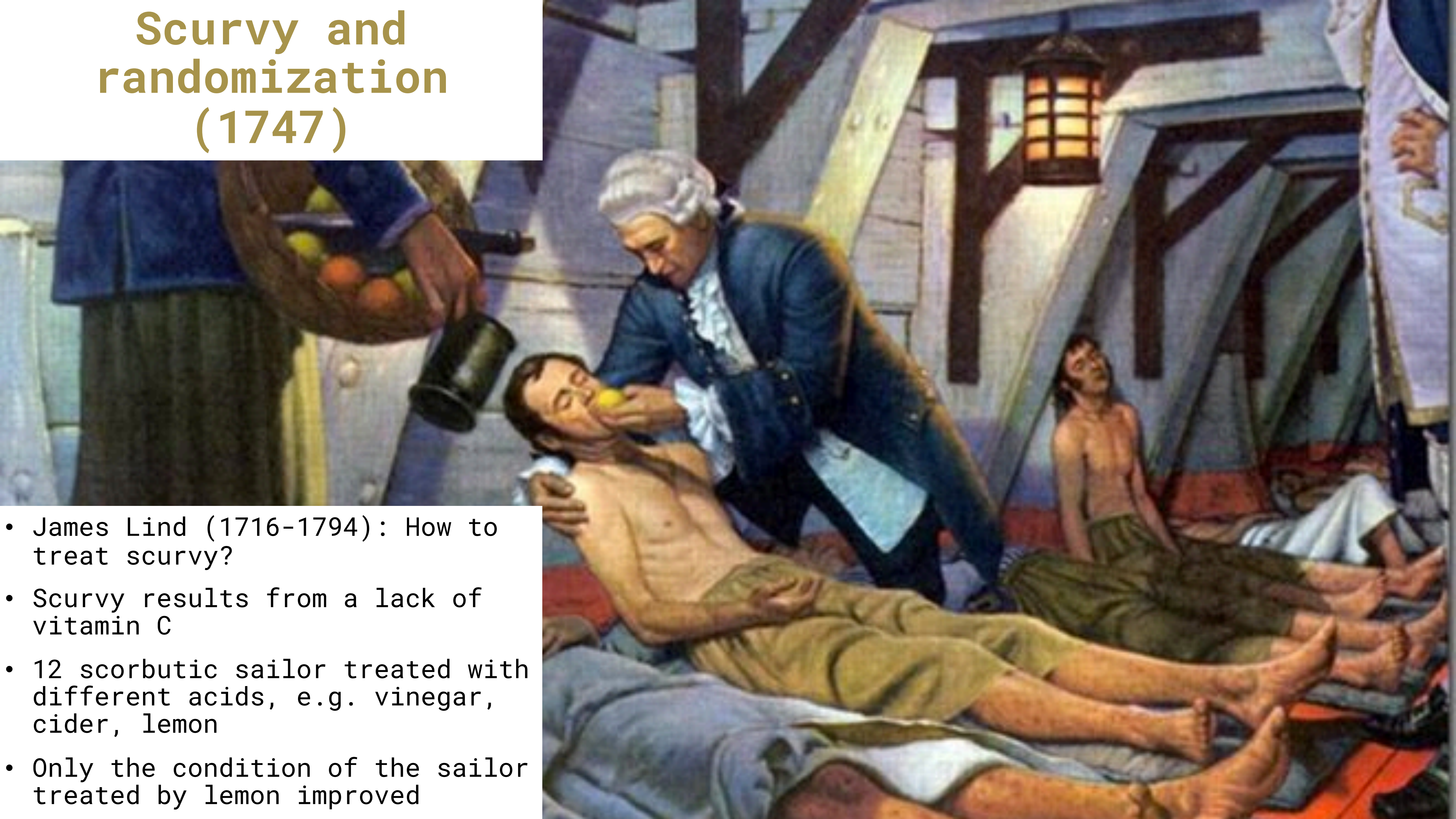




1. https://mbb-team.github.io/VBA-toolbox/wiki/behavioural-DCM/
2. Causal interaction following the alteration of target region activation during motor imagery training using real-time fMRI - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-difference-of-Granger-causality-index-between-TrainD-and-TrainA-in-the-experiment_fig3_259499871 [accessed 3 Sep, 2019]

A historic view of causal inference

Scurvy and randomization (1747)



- James Lind (1716-1794): How to treat scurvy?
- Scurvy results from a lack of vitamin C
- 12 scorbutic sailor treated with different acids, e.g. vinegar, cider, lemon
- Only the condition of the sailor treated by lemon improved

Fisher Experiment (1923) Potential Outcome by Rubin (1925)

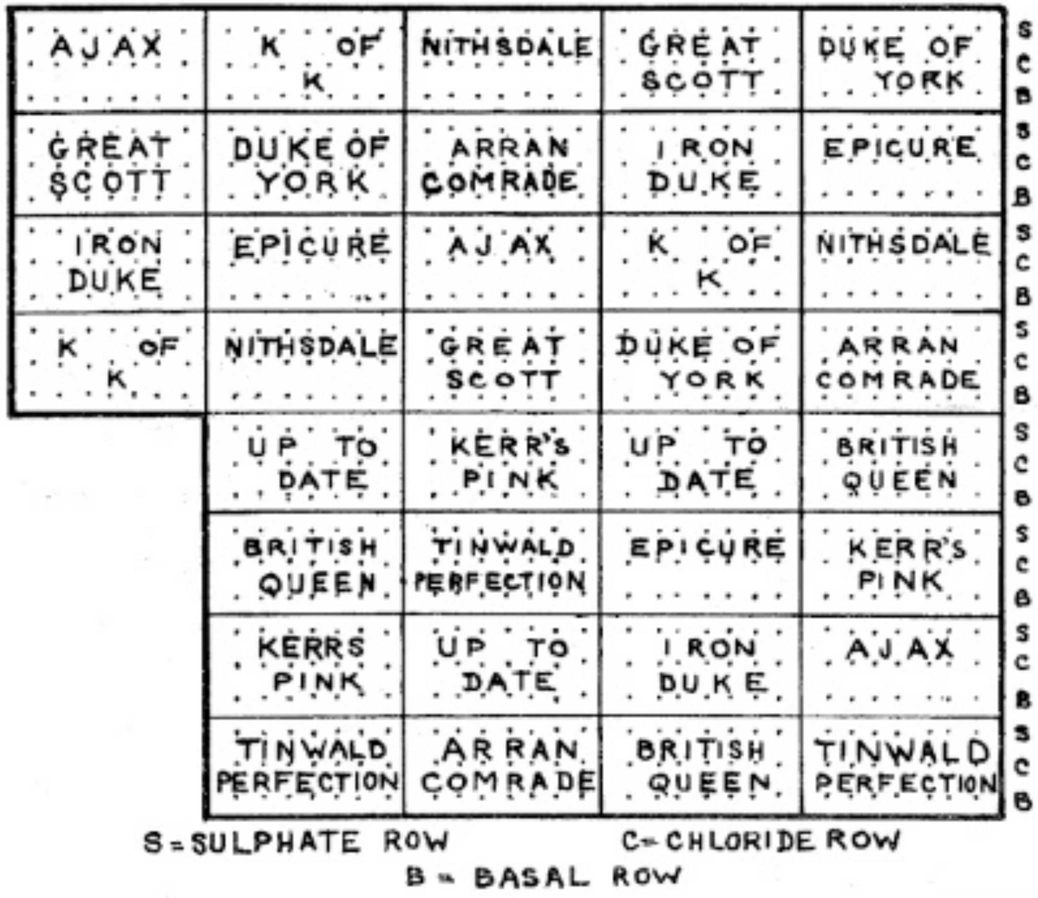
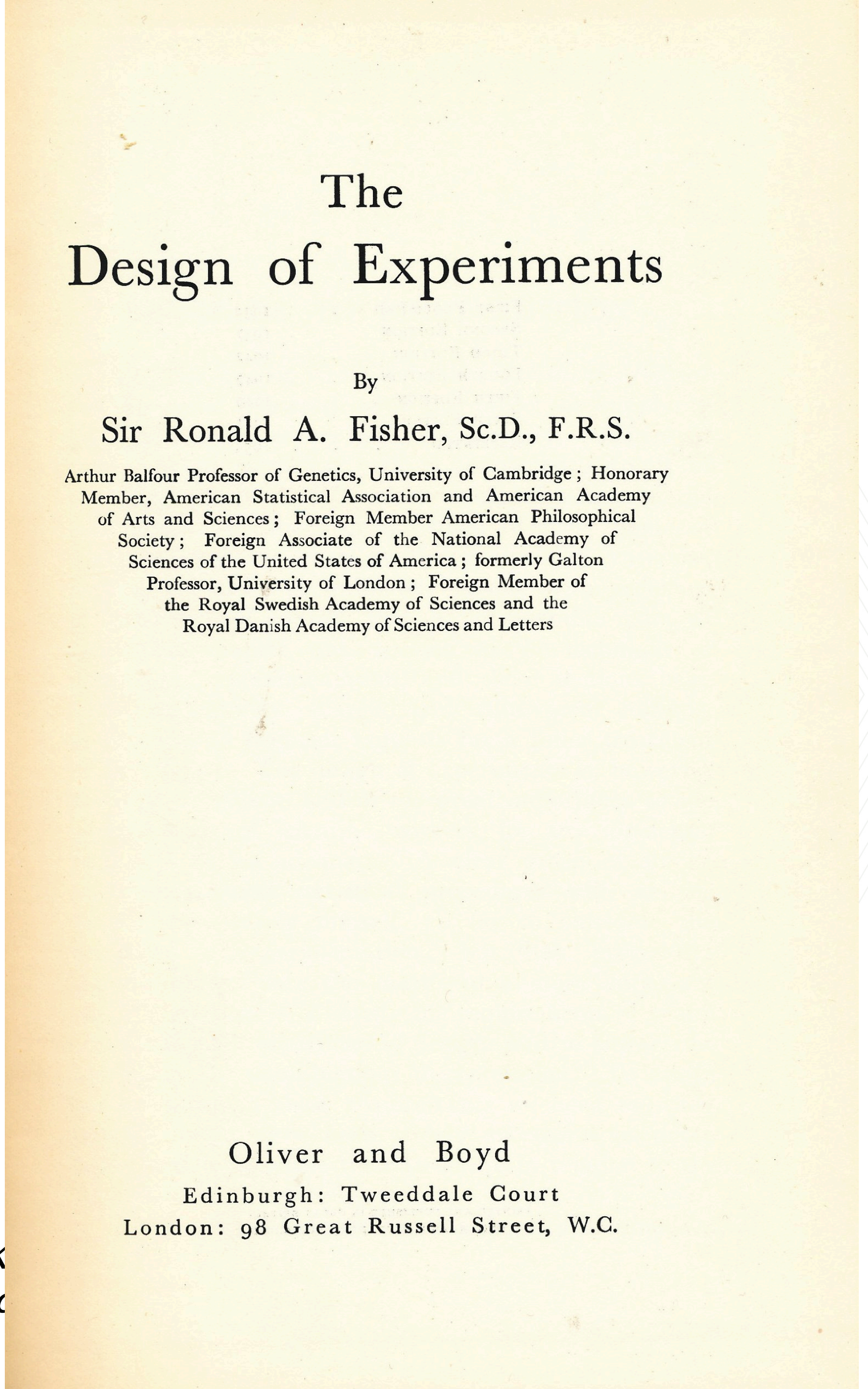


Diagram 1. Plan of experiment. Farmyard manure series.

A plan for a controlled agricultural experiment, from: R. A. Fisher and W. A. Mack Variation. II. The Manural Response of Different Potato Varieties," Journal of Agric 311–320.



The
Design of Experiments

By

Sir Ronald A. Fisher, Sc.D., F.R.S.

Arthur Balfour Professor of Genetics, University of Cambridge; Honorary Member, American Statistical Association and American Academy of Arts and Sciences; Foreign Member American Philosophical Society; Foreign Associate of the National Academy of Sciences of the United States of America; formerly Galton Professor, University of London; Foreign Member of the Royal Swedish Academy of Sciences and the Royal Danish Academy of Sciences and Letters

Oliver and Boyd
Edinburgh: Tweeddale Court
London: 98 Great Russell Street, W.C.

Modern Randomized Controlled Trial (1948)

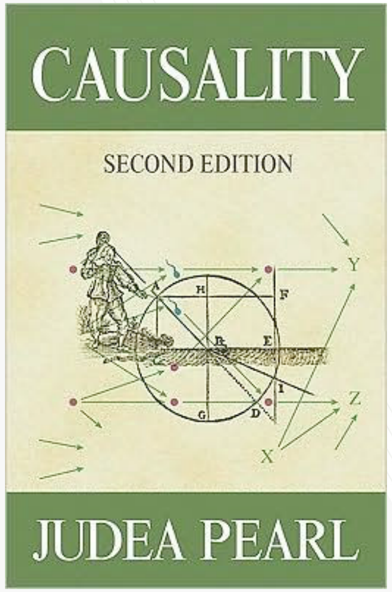
TABLE II.—*Assessment of Radiological Appearance at Six Months as Compared with Appearance on Admission*

Radiological Assessment	Streptomycin Group		Control Group	
Considerable improvement ..	28	*51%*	4	*8%*
Moderate or slight improvement	10	*18%*	13	*25%*
No material change	2	*4%*	3	*6%*
Moderate or slight deterioration	5	*9%*	12	*23%*
Considerable deterioration ..	6	*11%*	6	*11%*
Deaths	4	*7%*	14	*27%*
Total	55	*100%*	52	*100%*

- The first published RCT
  - "Streptomycin treatment of pulmonary tuberculosis”

Neyman–Rubin Causal Model (1974)

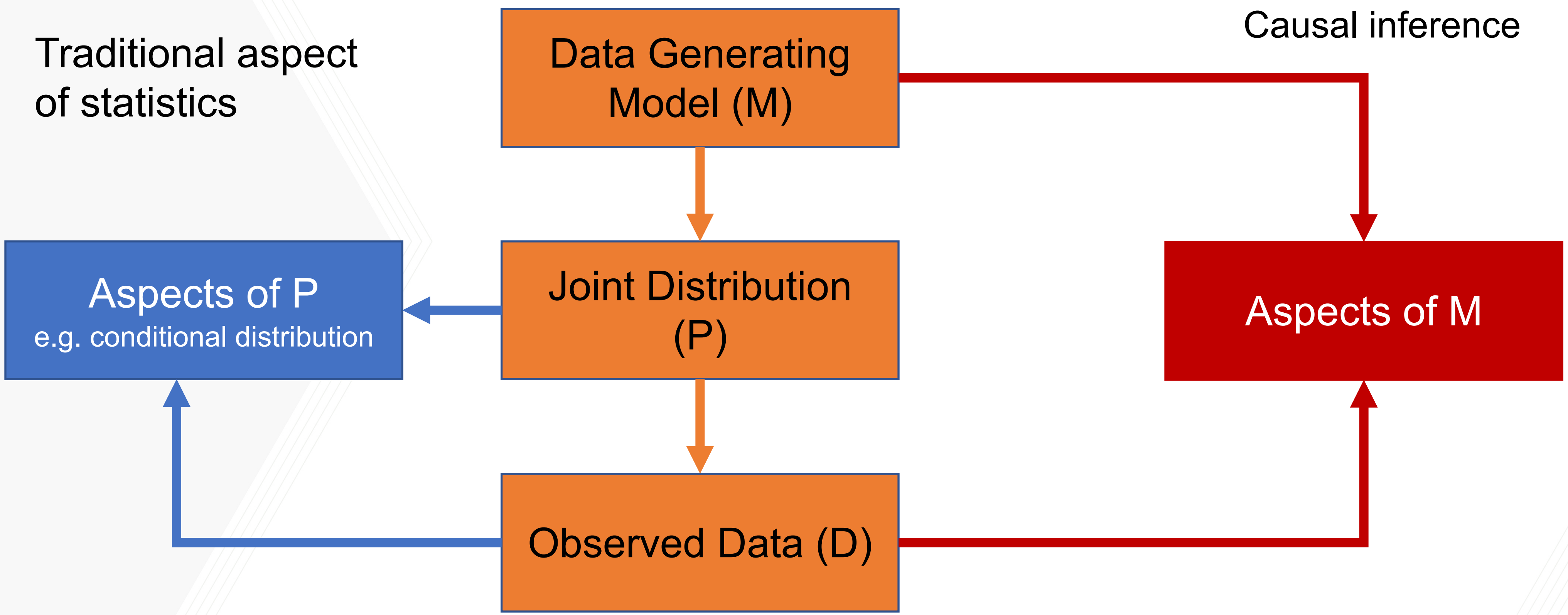
- Fundamental problem facing causal inference (Rubin, 1975)
- It is a missing data problem
- For an individual, only one out of the two outcome can be observed (assuming binary treatment).



Structural Causal Models (1995)

- ACM Turing Award 2011: “For fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning.”

Diagram View



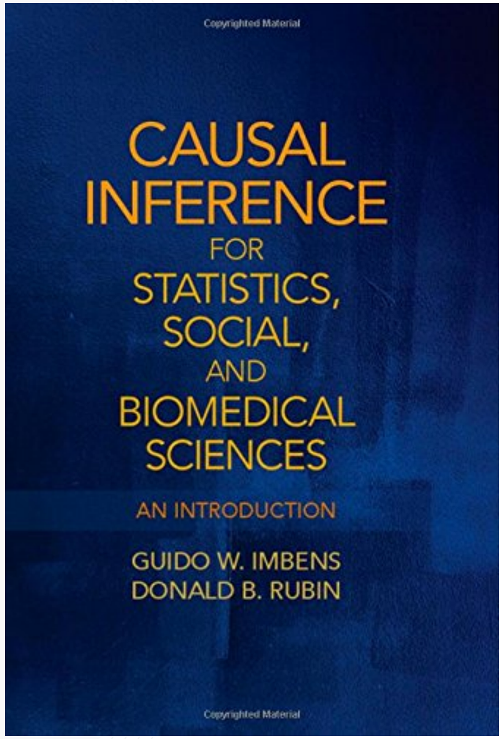
Two basic questions of CI

- Q1. Model Discovery
  - Does X cause Y?
  - Does X, Y have a common cause?

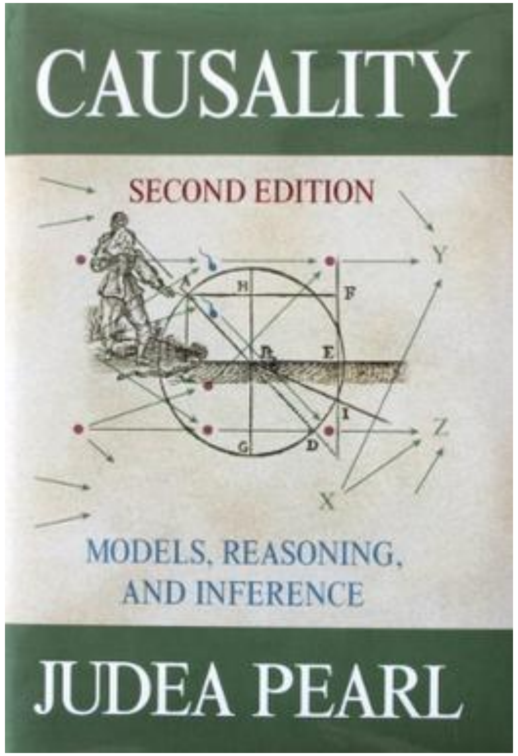
- Q2. Variable Inference
  - What is the causal effect of X on Y?
  - What if I take action T?
  - Had I taken action T=1 instead of T=0, what would the outcome Y had been?

Basic Causal Inference Frameworks

A taxonomy of CI frameworks



Potential Outcomes



Structural Causal Models

Potential Outcome

Drug A

Patient

Which one to take?

Drug B

In a factual world

Patient

Takes

Drug A

Cured

In a counter-factual world

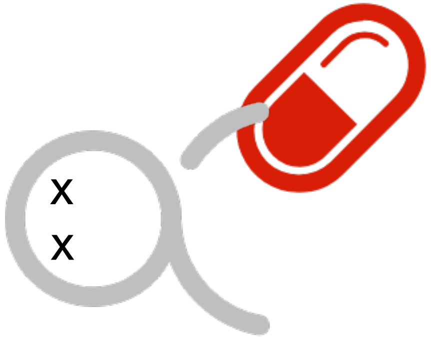
Patient

Takes

Drug B

Died

Causal Effect Under Potential Outcome



Outcome | Treatment=A — Outcome | Treatment=B

Patient	Drug	$Y_{T=A}$	$Y_{T=B}$
P1	A	0.1	
P2	B		0.6
P3	B	0.3	
P4	A		0.1
P5	B	0.5	
P6	A		0.5

- Causal Effect
  - $E[Y_{T=A} - Y_{T=B}]$

Patient	Drug	$Y_{T=A}$	$Y_{T=B}$
P1	A	0.1	?
P2	B	?	0.6
P3	B	0.3	?
P4	A	?	0.1
P5	B	0.5	?
P6	A	?	0.5

- Causal Effect
  - $E[Y_{T=A} - Y_{T=B}]$

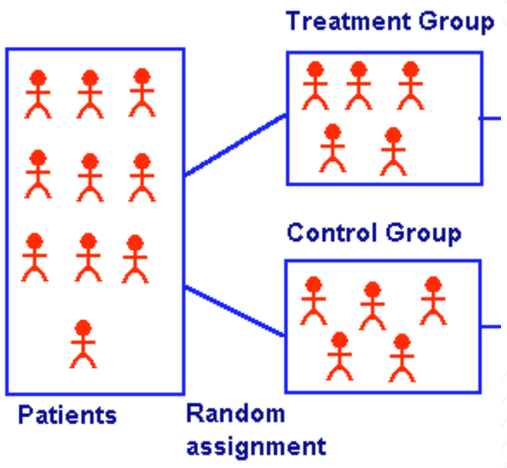
Causal inference is the problem of estimating the missing data

Patient	Drug	$Y_{T=A}$	$Y_{T=B}$
P1	A	0.1	?
P2	B	?	0.6
P3	B	0.3	?
P4	A	?	0.1
P5	B	0.5	?
P6	A	?	0.5

- Causal Effect
  - $E[Y_{T=A} - Y_{T=B}]$
- Fundamental challenge:
  - We can only observe one of the two potential outcomes
  - A missing data problem

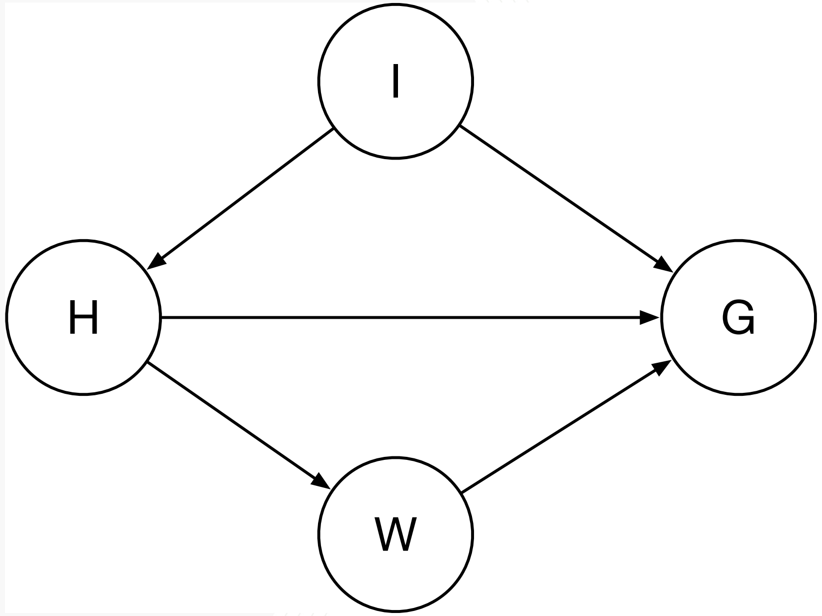
What can we do?

- Randomized Controlled Trials (RCTs) are the golden standard
- Cons:
  - Unethical
  - External validity
  - Infeasible or costly
- => need to learn from data where we don’t have actual control (i.e. observation data)



Structural Causal Models

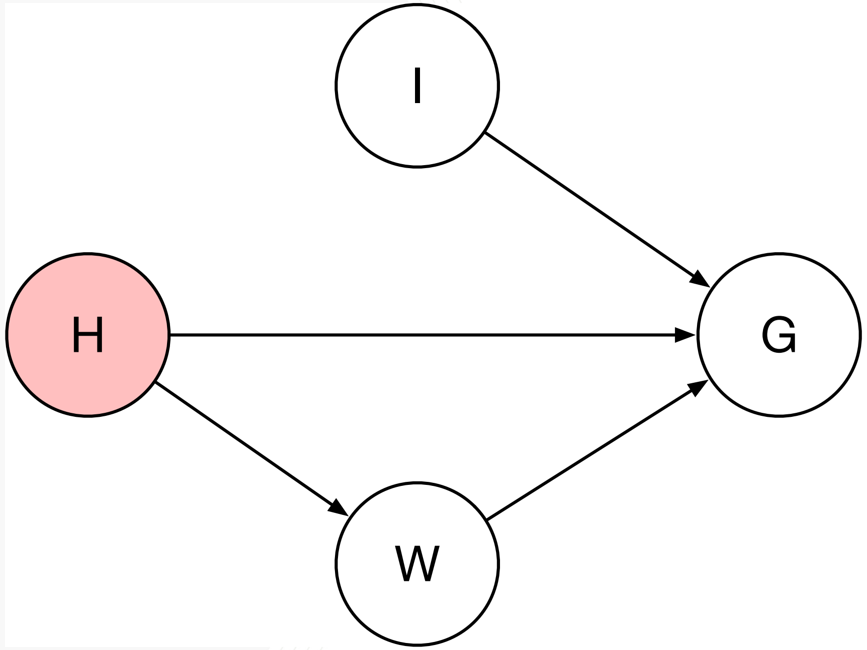
- A graph with edges for "causes" relationships



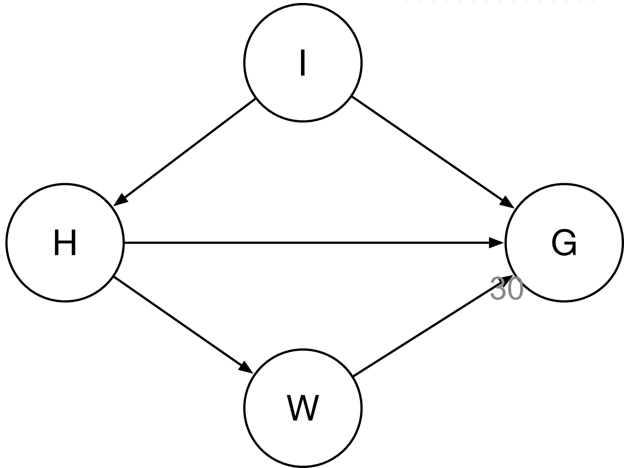
- A set of equations for quantitative relationships
  - Interest: $I = N_I \sim (0,1)$
  - Hour of study: $H = 2I + 1 + N_H$
  - Working method: $W = H^2 - 3 - N_W$
  - Exam Grade: $G = I + H + W + N_G$
- Here N are noise variables

Intervention: do(H=2)

- Intervention on the graph



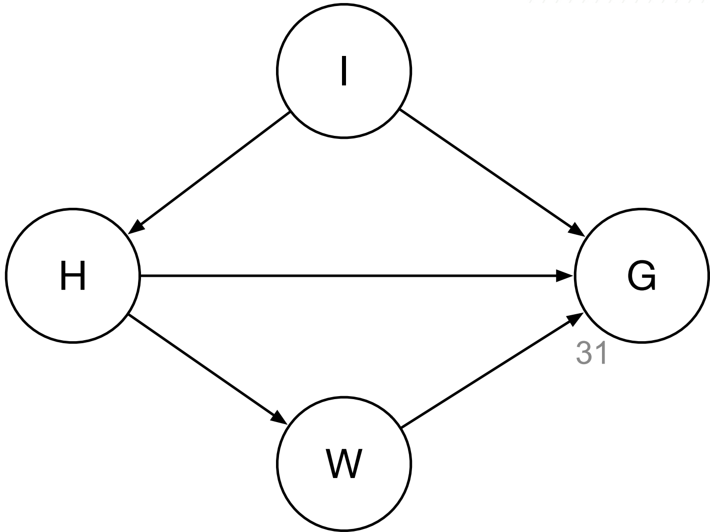
- Intervention on the equations
  - Interest: $I = N_I \sim (0,1)$
  - Hour of study: $H = 2$
  - Working method: $W = H^2 - 3 - N_W = 1 - N_W$
  - Exam Grade: $G = I + H + W + N_G = I + 3 - N_W + N_G$



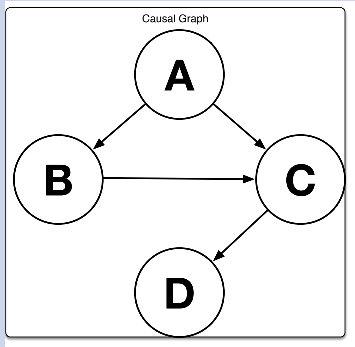
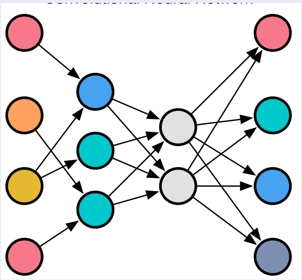
Counterfactual

- Factual: now we observe in the world
  - I=0.5
  - H=4
  - W=12
  - G=17
- => We can deduct all the noise variables
  - $N_W = 1$
  - $N_G = 0.5$

- Counterfactual: what would have happened, had I studied for 2 hours?
  - We know after do(H=2)
  - $G = I + H + W + N_G = I + 3 - N_W + N_G$ = 3



Taxonomy of Causal Inference Methods

	Description	Pros	Cons
SCM  	use/assume/learn a causal model (e.g. SCM), and perform inference based on it	- More sample efficient - Incorporate domain knowledge	- Hard to construct the model graph - Computationally expensive - Often assume all variables observed
Potential Outcome 	directly estimate causal quantities of interest using data	- Easy to model	- Needs a lot of data - Less of insight on the actual data generation process

Both frameworks have merits

Use structural causal model and do-calculus for

modeling the problem

making assumptions explicit

identifying the causal effect

Use potential outcomes-based methods for

estimating the causal effect

Next 2 units

	Causal language 1: structural causal model	Causal language 2: potential outcomes
Model discovery (discover the causal relationship between variables)		
Effect identification (predict what effect an intervention can have)		This is the key focus of this tutorial

Agenda

Link: https://github.com/hang-wu/CI

Break

Part 1

- Introduction of Causal Inference

3:00 – 3:30

Part 2

- Basic of Causal Effect Estimation Algorithm

3:30 – 4:15

Part 3

- Recent Advances
- Challenges and Opportunities

4:35 – 5:15

I'll attend the NSF student award from 4:20 – 4:30 CNB-MAC