WHAT IS CAUSAL INFERENCE AND WHERE IS DATA SCIENCE GOING?

Judea Pearl
UCLA
Computer Science and Statistics

IDRE Seminar, UCLA
May 27, 2022, 12-1pm

OUTLINE

1. Data Science: Successes, limitations, and tensions
2. The Causal Revolution: From fitting to understanding
3. The ladder of causation
4. The seven wisdoms of causal thinking
5. Future directions:
   a. Automated scientists
   b. From population to individual decisions

DATA SCIENCE –
A CLASH OF TWO PARADIGMS

1. The data-centric paradigm
   - How best to fit the data so as to maximize success on the training set.
2. The scientific paradigm
   - What should the world be like before I can answer my research question?
3. Extracting Knowledge from Data (IDRE)
4. Extracting Understanding from Data

WHAT CAPABILITIES DOES DEEP UNDERSTANDING ENTAIL?

A state of knowledge evoking a sensation of “being in control.”

1. Predict future events from past/present observations
2. Predict consequence of contemplated actions
3. Provide explanations of unanticipated events
4. Imagine alternative worlds or “Roads not Taken”
5. Design new experiments, seek new observations (attention, curiosity, and conjectures)

TYPICAL QUESTIONS NEEDING UNDERSTANDING

1. How effective is a given treatment in preventing a disease?
2. Was it the new tax break that caused our sales to go up? Or our marketing campaign?
3. What is the annual health-care costs attributed to obesity?
4. Can hiring records prove an employer guilty of sex discrimination?
5. I am about to quit my job, will I regret it?
   - Unarticulatable in the standard grammar of science.

$$Y = aX \quad \text{vs.} \quad Y \leftarrow aX$$

SEWALL WRIGHT – CAUSALITY'S FIRST FORMAL VOICE

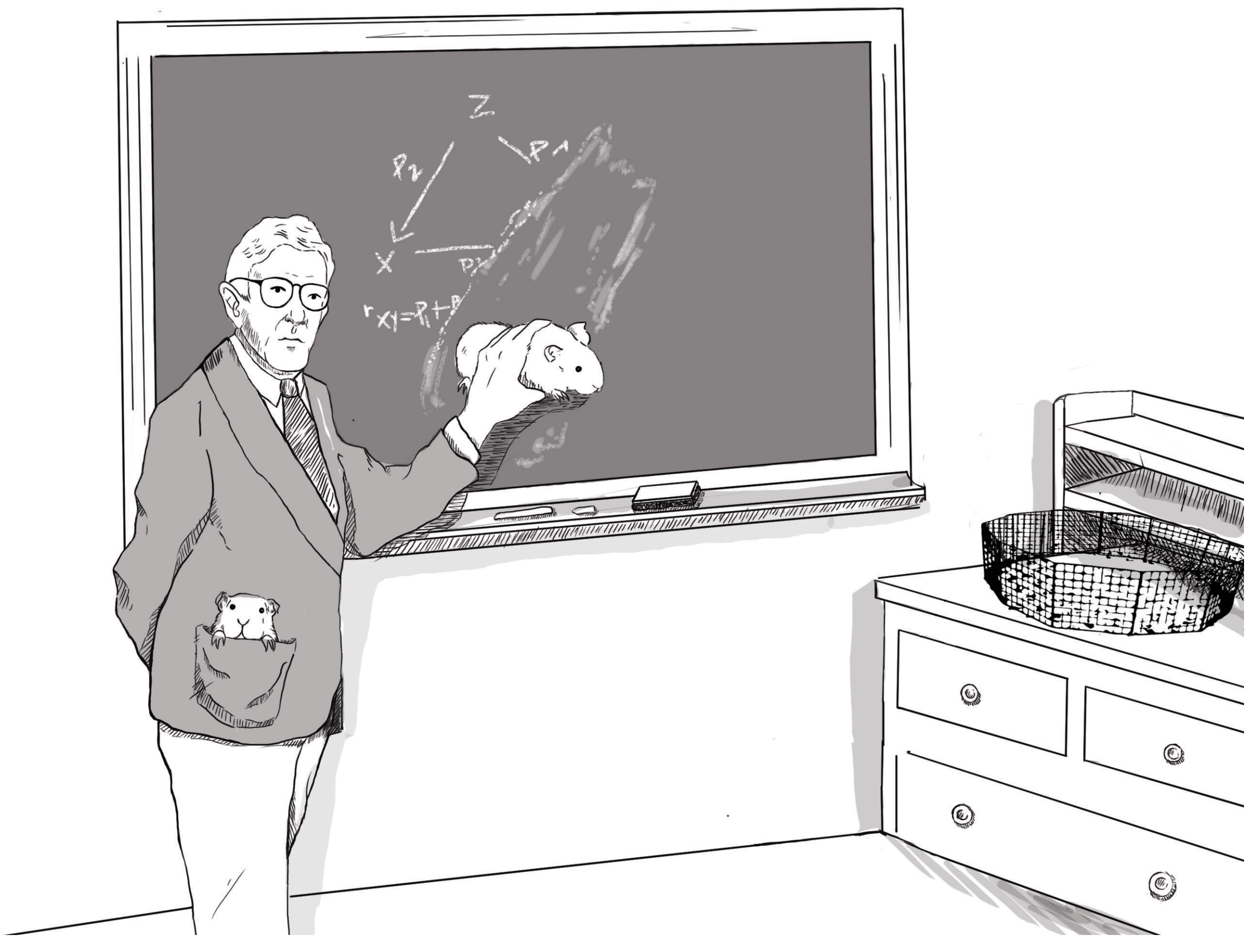


Figure 2.6. Sewall Wright was the first person to develop a mathematical method for answering causal questions from data, known as path diagrams. His love of mathematics surrendered only to his passion for guinea pigs.

SEWALL WRIGHT – CAUSALITY'S FIRST FORMAL VOICE

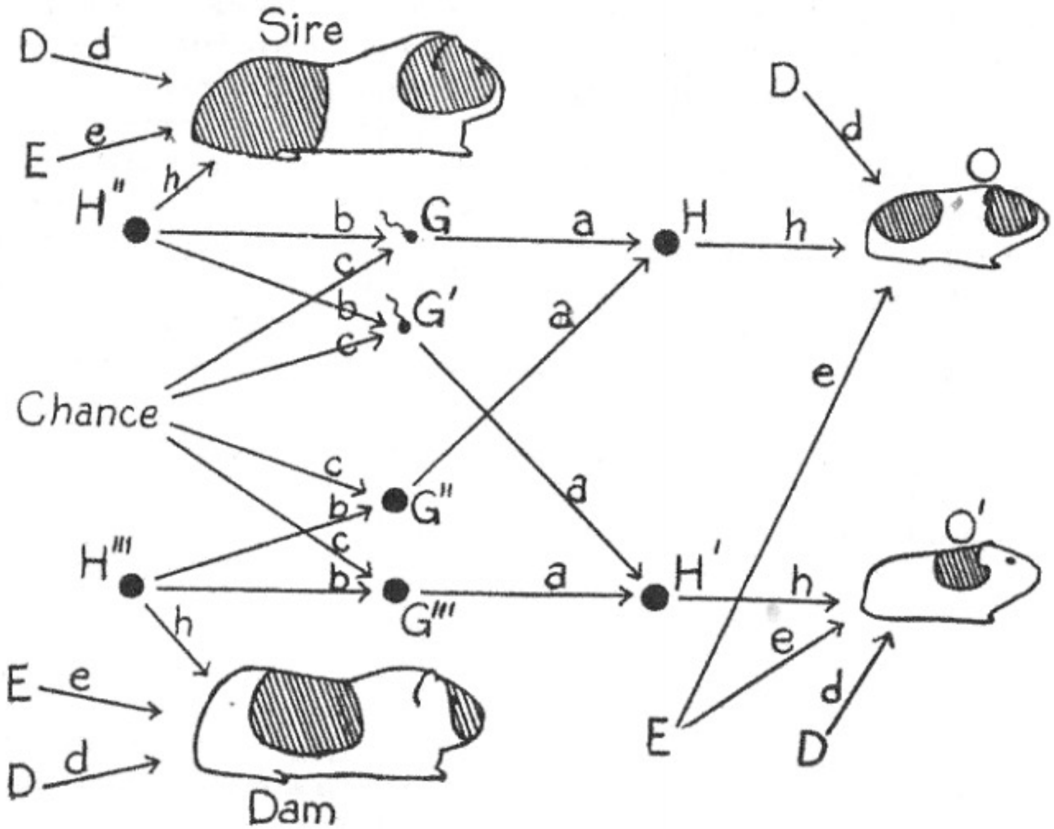
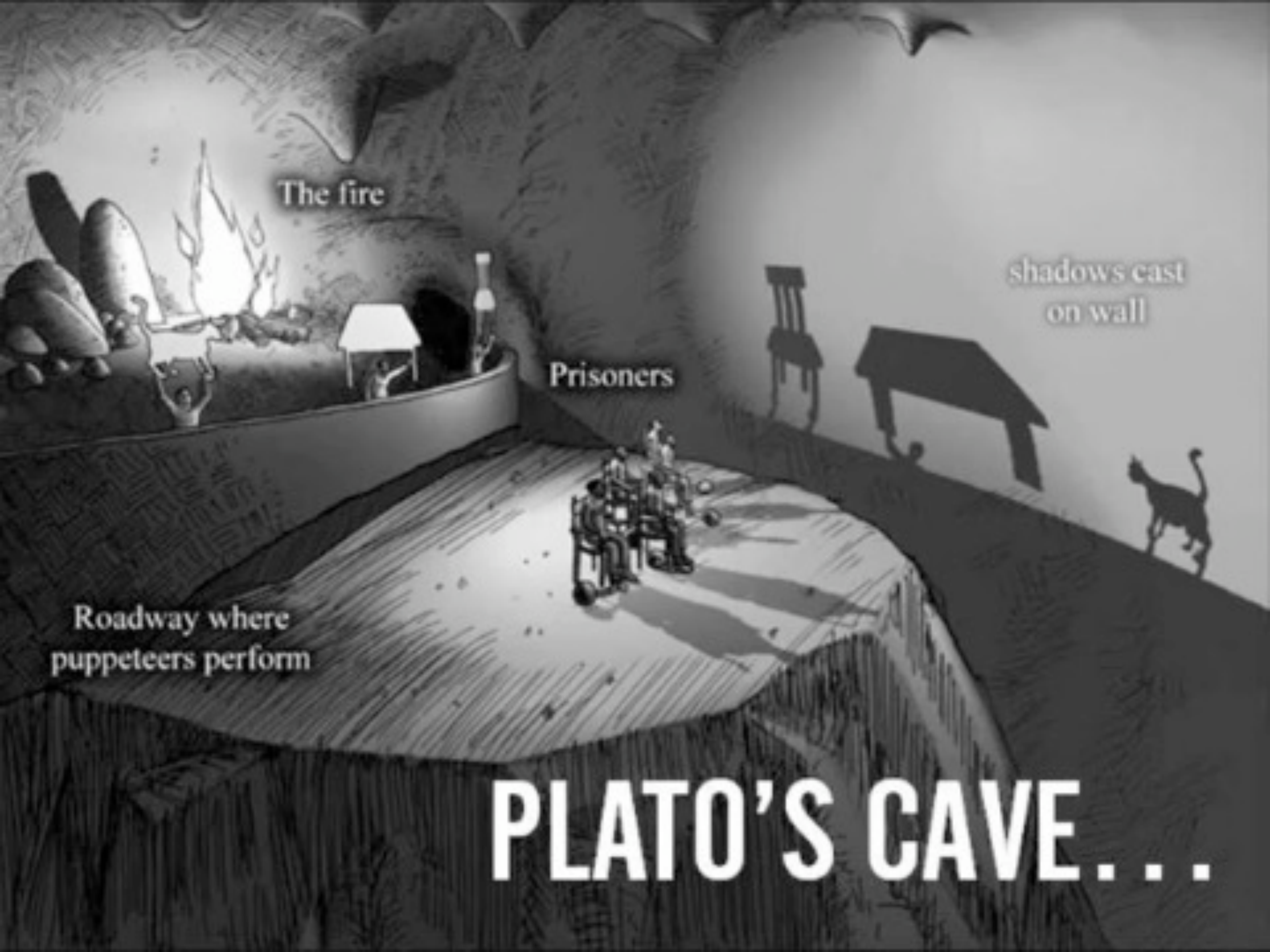


FIG. 5.

Figure 2.7. Sewall Wright's first path diagram, illustrating the factors leading to coat color in guinea pigs. D = developmental factors (after conception, before birth), E = environmental factors (after birth), G = genetic factors from each individual parent, H = combined hereditary factors from both parents, O, O¢ = offspring. The objective of analysis was to estimate the strength of the effects of D, E, H (written as d, e, h in the diagram). (Source: Sewall Wright, Proceedings of the National Academy of Sciences [1920], 320–332.)

WHY WAS WRIGHT ATTACKED? THE STATISTICS PARADIGM 1834–2022

- “The object of statistical methods is the reduction of data” (Fisher 1922).
- Statistical concepts are those expressible in terms of joint distribution of observed variables.
- All others are: “substantive matter,” “domain dependent,” “metaphysical,” “ad hockery,” i.e., outside the province of statistics, ruling out all interesting questions.
- Slow awakening since Neyman (1923) and Rubin (1974).
- Traditional Statistics Education = Causalophobia


The fire
shadows cast
on wall
Prisoners
Roadway where
puppeteers perform
PLATO'S CAVE...

FROM STATISTICAL TO CAUSAL ANALYSIS: 2. THE SHARP BOUNDARY

1. Causal and associational concepts do not mix.

CAUSAL	ASSOCIATIONAL
Spurious correlation	Regression
Randomization / Intervention	Association / Independence
"Holding constant" / "Fixing"	"Controlling for" / Conditioning
Confounding / Effect	Odds and risk ratios
Instrumental variable	Collapsibility / Granger causality
Ignorability / Exogeneity	Propensity score

2.

3.

4.

FROM STATISTICAL TO CAUSAL ANALYSIS: 3. THE MENTAL BARRIERS

1. Causal and associational concepts do not mix.

CAUSAL	ASSOCIATIONAL
Spurious correlation	Regression
Randomization / Intervention	Association / Independence
"Holding constant" / "Fixing"	"Controlling for" / Conditioning
Confounding / Effect	Odds and risk ratios
Instrumental variable	Collapsibility / Granger causality
Ignorability / Exogeneity	Propensity score

2. No causes in – no causes out (Cartwright, 1989)

$$\left.\begin{array}{r}\text{data}\\ \text{causal assumptions (or experiments)}\end{array}\right\} \Rightarrow \text{causal conclusions}$$

3. Causal assumptions cannot be expressed in the mathematical language of standard statistics.

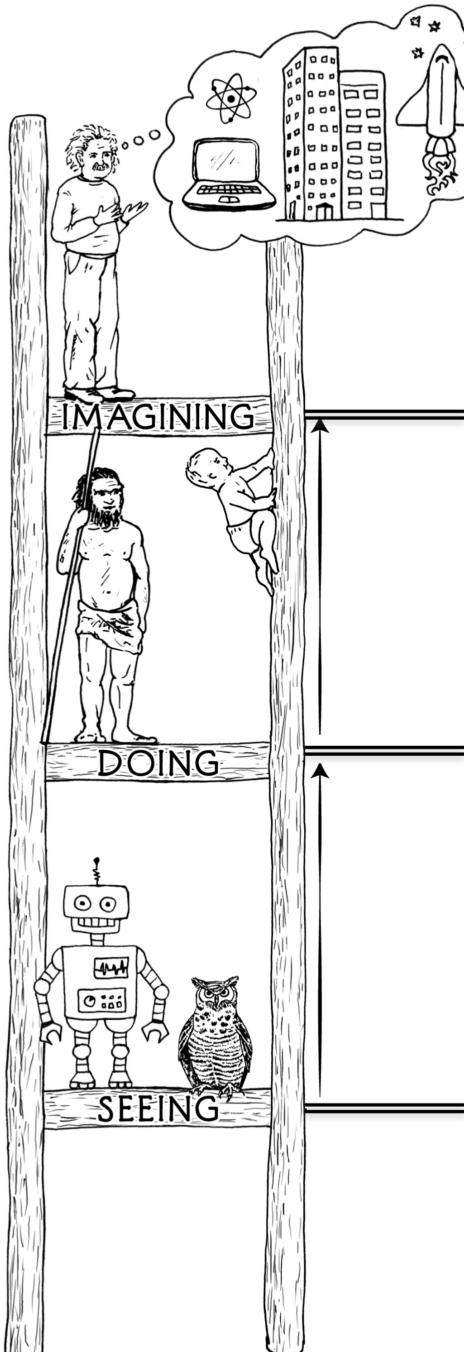
4. Non-standard mathematics:
   a) Structural equation models (Wright, 1920; $X \rightarrow Y$)
   b) Counterfactuals (Neyman-Rubin (Y_x), Lewis $(x \mathbin{\Box\!\!\rightarrow} Y)$)

THE CAUSAL REVOLUTION

1. "More has been learned about causal inference in the last few decades than the sum total of everything that had been learned about it in all prior recorded history."
   (Gary King, Harvard, 2014)
2. From liability to respectability
   - JSM 2003 – 13 papers
   - JSM 2013 – 130 papers
   - 2022 – dozens of causality-specific workshops and conferences

3. Fun, profit, education, and Mind

THE LADDER OF CAUSATION



3. COUNTERFACTUALS

ACTIVITY: Imagining, Retrospection, Understanding

QUESTIONS: *What if I had done . . . ? Why?*
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)

EXAMPLES: Was it the aspirin that stopped my headache?
Would Kennedy be alive if Oswald had not killed him? What if I had not smoked the last 2 years?

2. INTERVENTION

ACTIVITY: Doing, Intervening

QUESTIONS: *What if I do . . . ? How?*
(What would Y be if I do X?)

EXAMPLES: If I take aspirin, will my headache be cured?
What if we ban cigarettes?

1. ASSOCIATION

ACTIVITY: Seeing, Observing

QUESTIONS: *What if I see . . . ?*
(How would seeing X change my belief in Y?)

EXAMPLES: What does a symptom tell me about a disease?
What does a survey tell us about the election results?

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way

Tool 2: Predicting the effects of actions and policies

Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)

Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)

Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)

Tool 6: Recovering from missing Data

Tool 7: Causal Discovery

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way

Tool 2: Predicting the effects of actions and policies

Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)

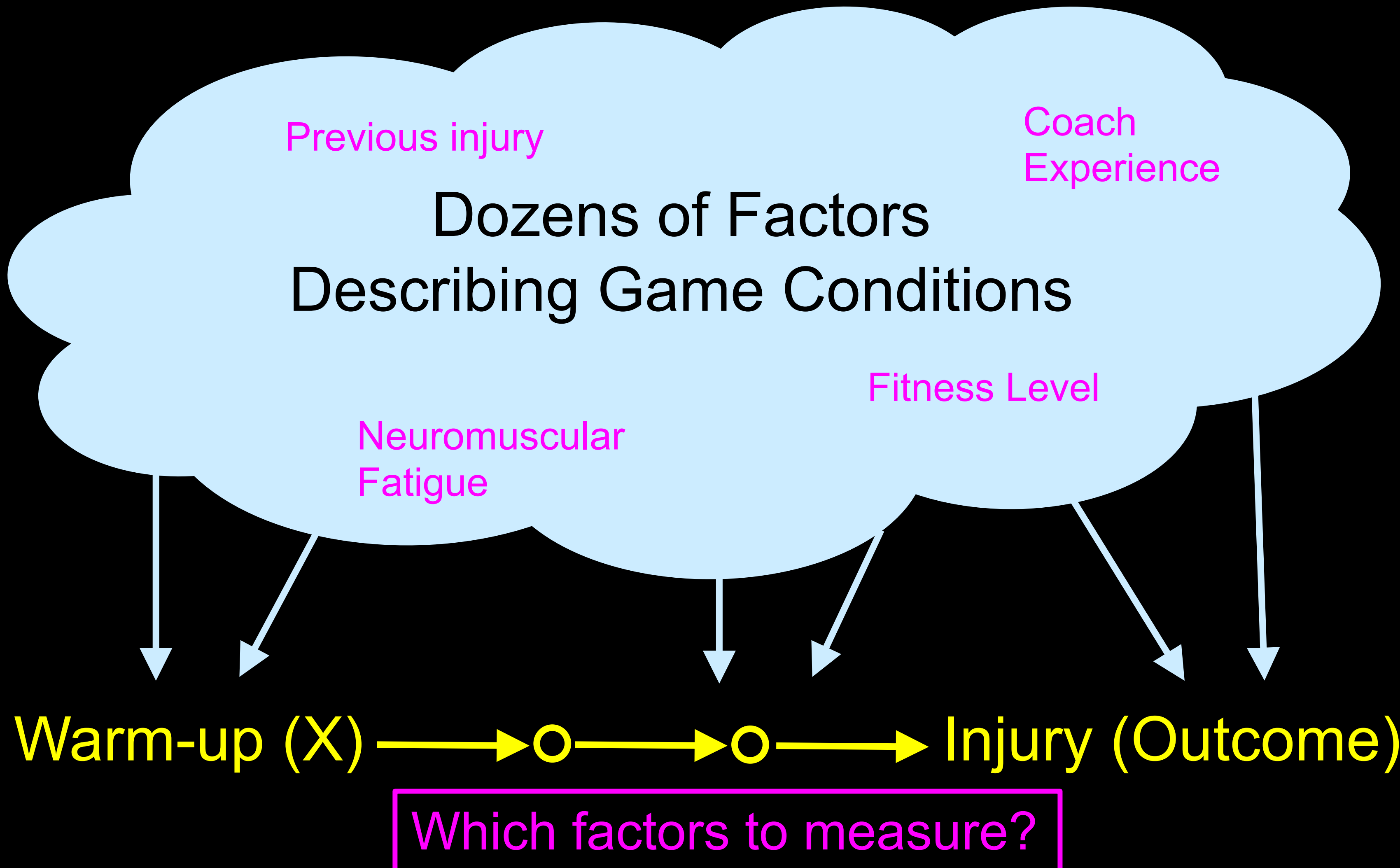
Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)

Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)

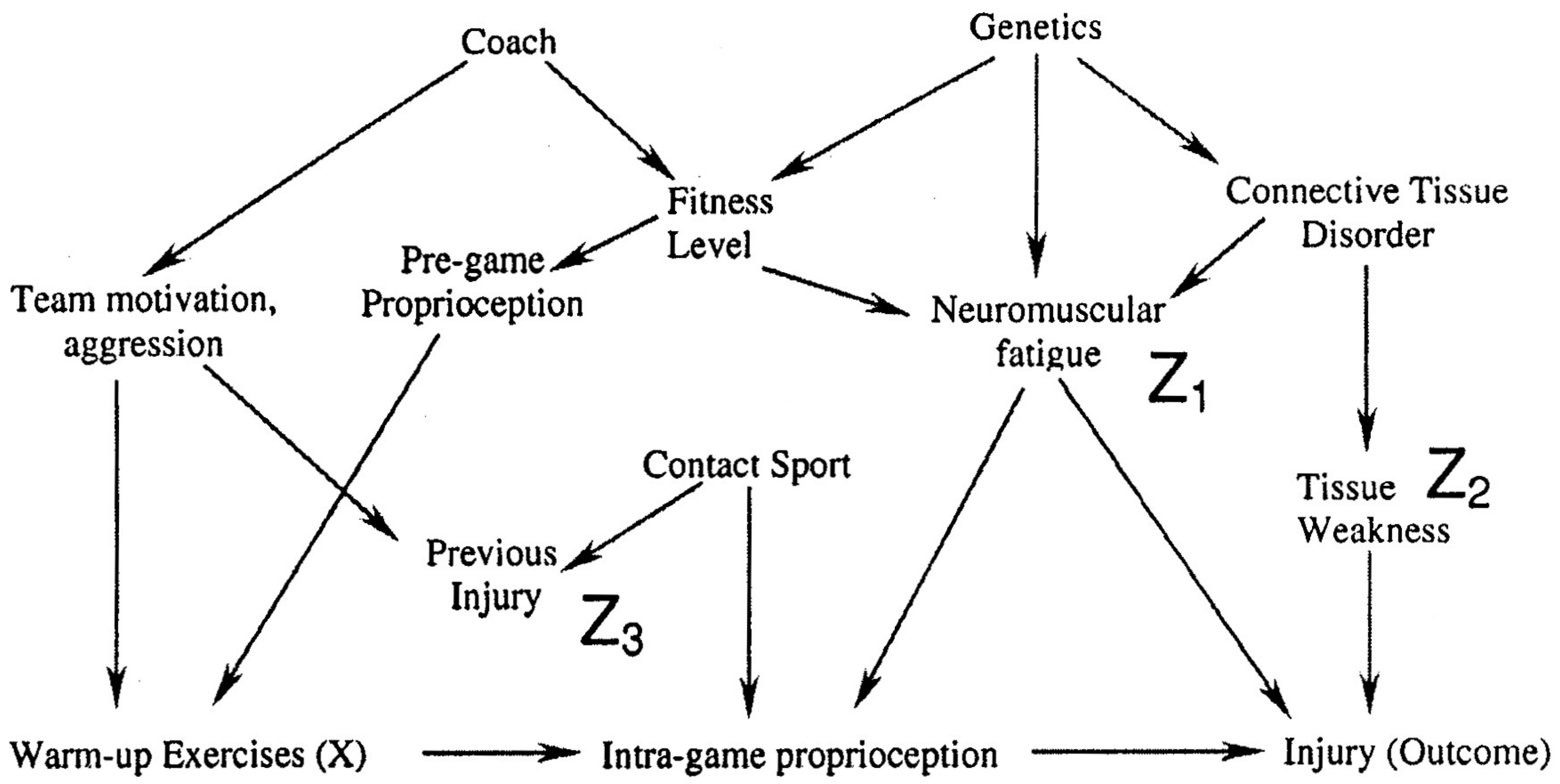
Tool 6: Recovering from missing Data

Tool 7: Causal Discovery

EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)

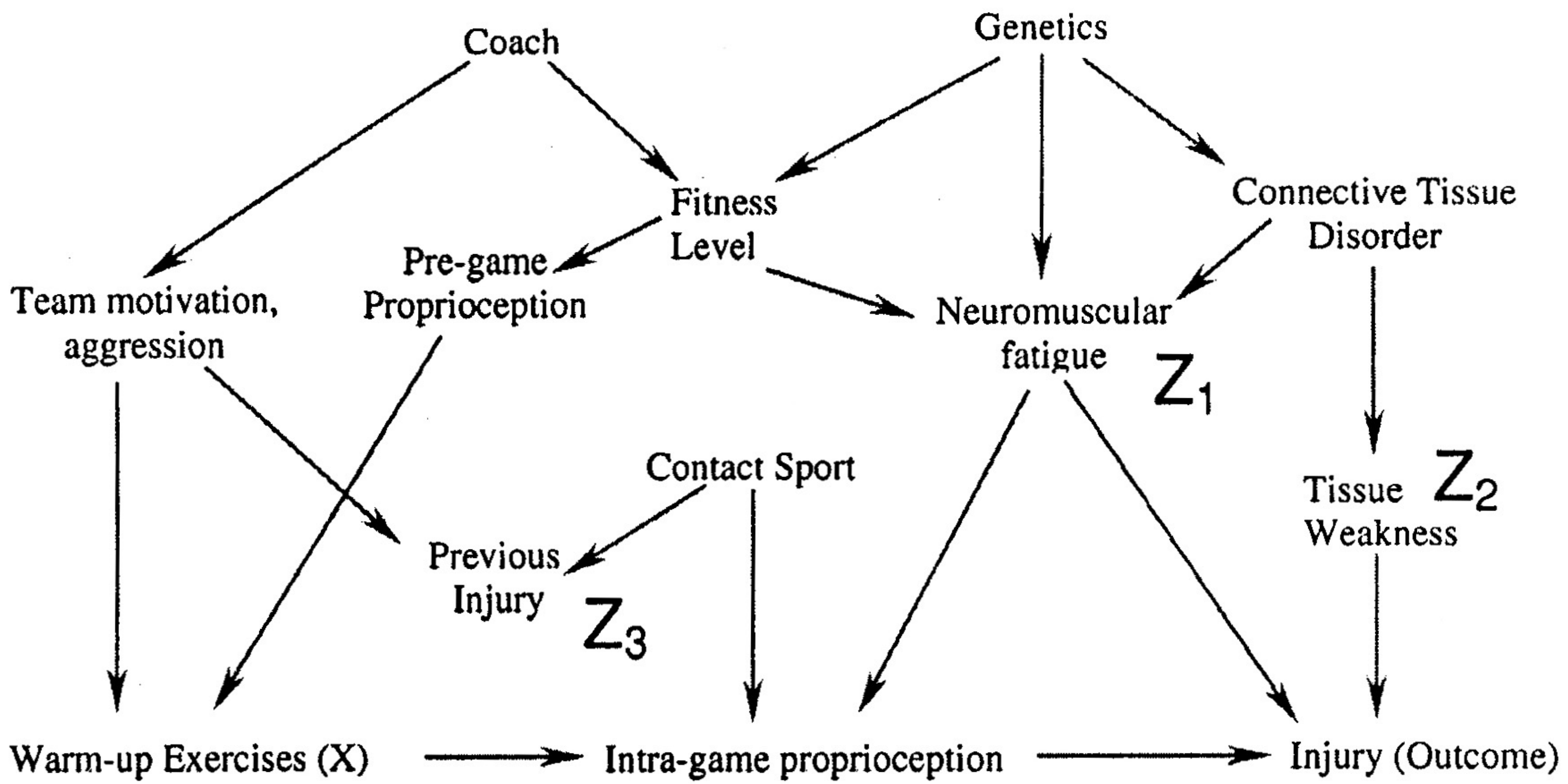


EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)



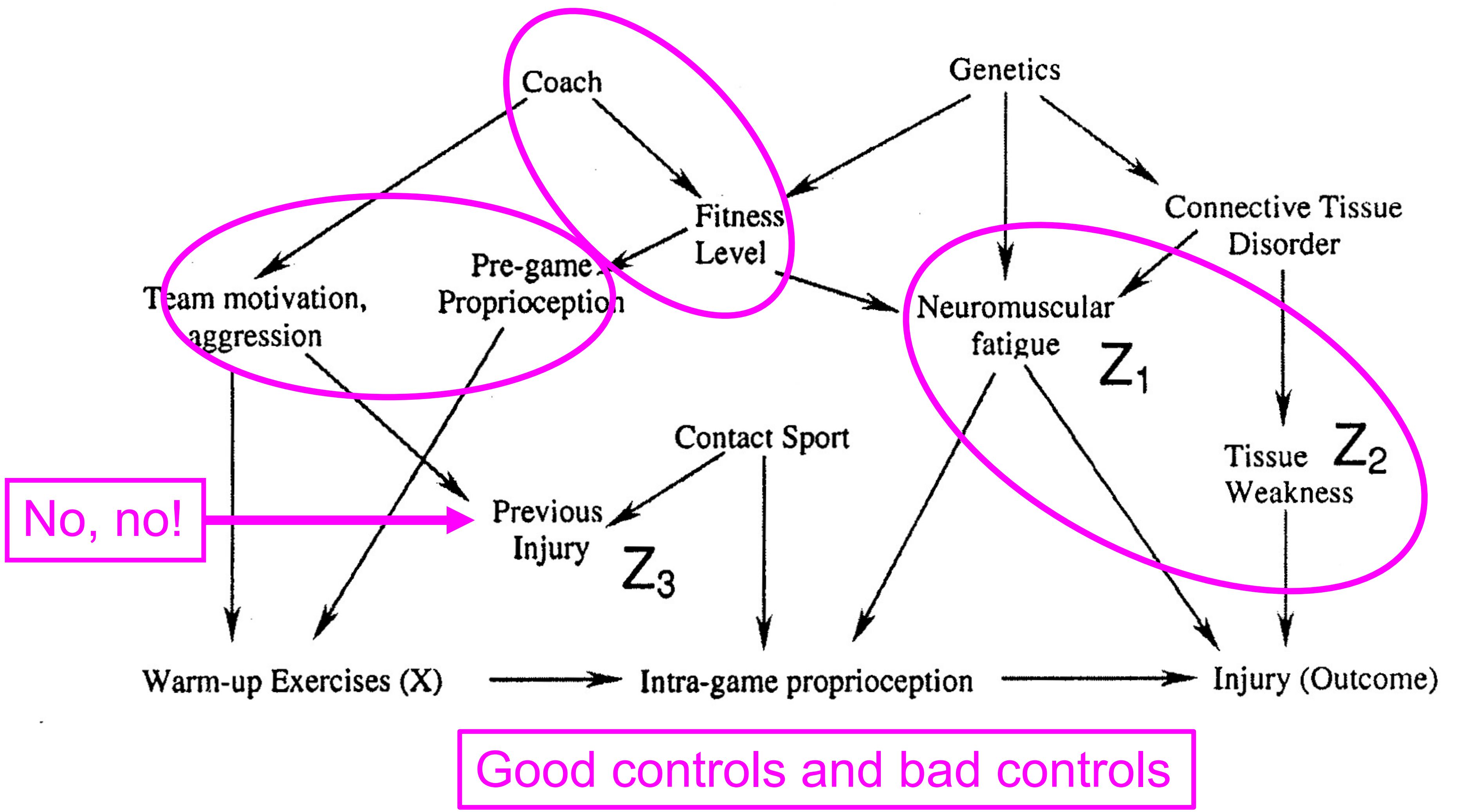
Which factors to adjust?

EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)

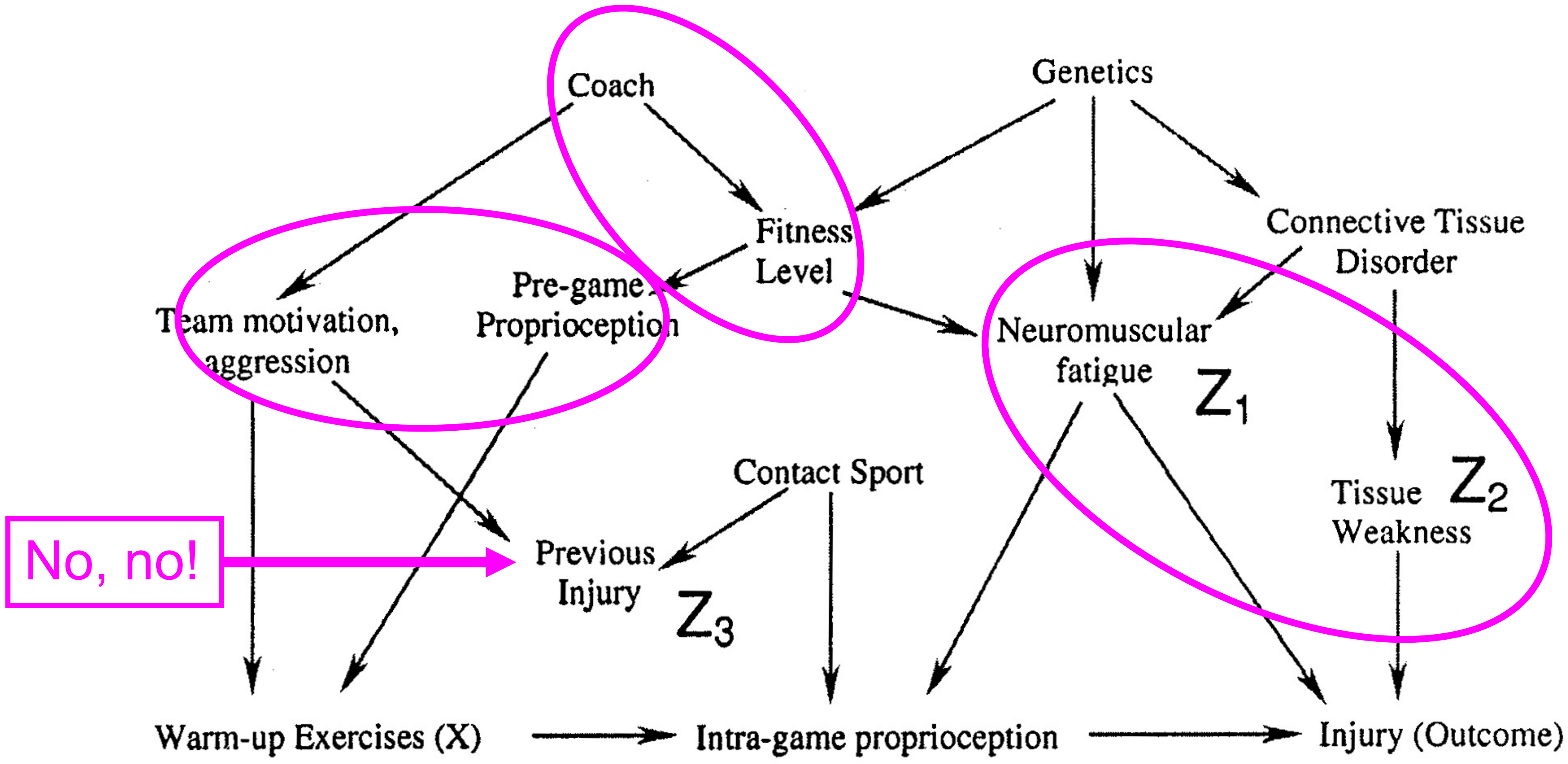


To adjust or not to adjust?

EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)

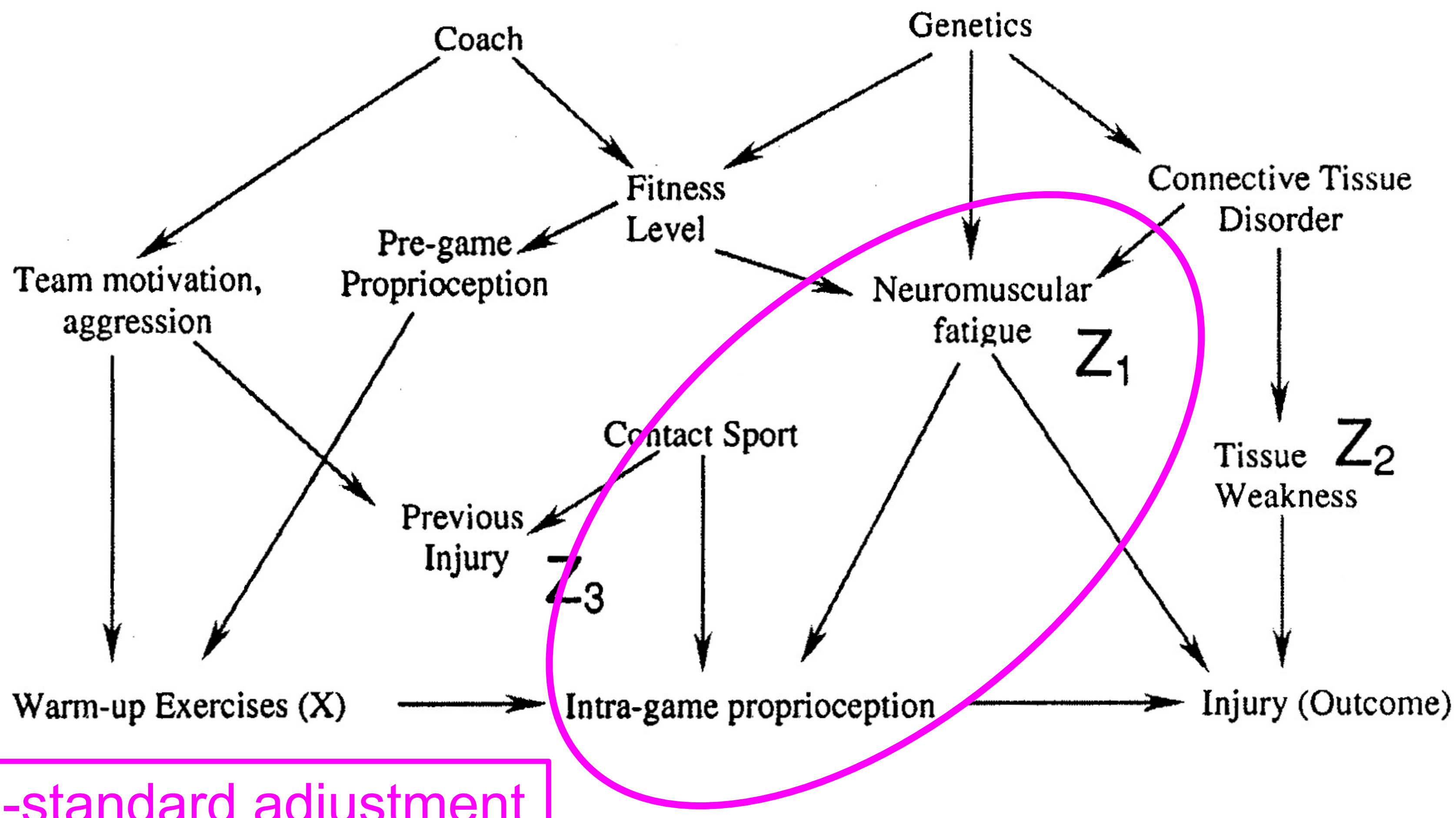


EFFECT OF WARM-UP ON INJURY
(After Shrier & Platt, 2008)



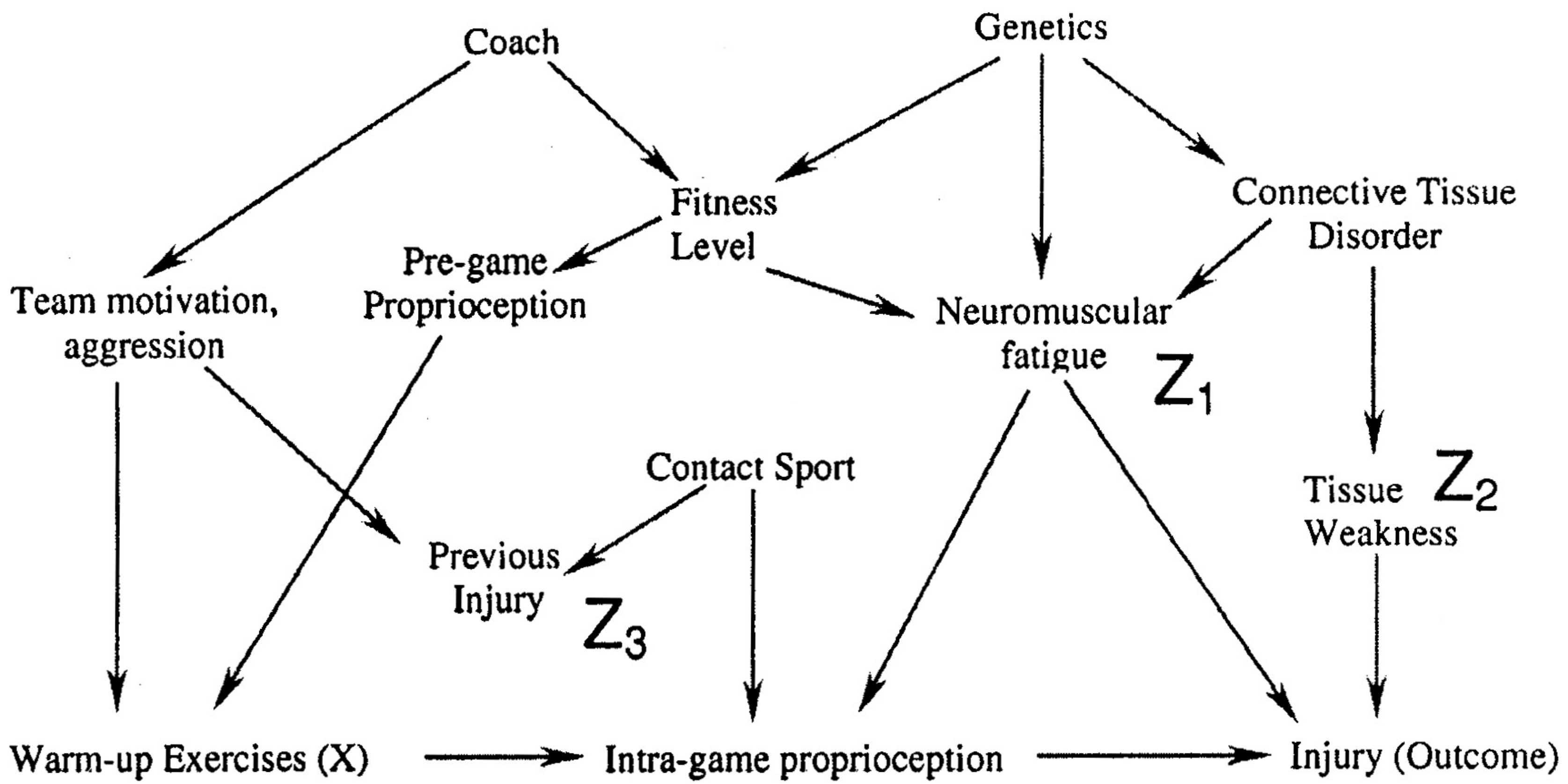
Back-door Victory! Confounding deconfounded!

EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)



Non-standard adjustment

EFFECT OF WARM-UP ON INJURY (After Shrier & Platt, 2008)



Wisdom: Effect-identifiability, completeness, testability

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way

Tool 2: Predicting the effects of actions and policies

Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)

Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)

Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)

Tool 6: Recovering from missing Data

Tool 7: Causal Discovery

ATTRIBUTION (Rung-3)

- Your Honor! My client (Mr. A) died BECAUSE he used this drug.

ATTRIBUTION (Rung-3)

- Your Honor! My client (Mr. A) died BECAUSE he used this drug.

- Court to decide if it is MORE PROBABLE THAN NOT that Mr. A would be alive BUT FOR the drug!
- $PN = P(alive_{\{no\ drugs\}} \mid dead, drug) \geq 0.50$

CAN WE COMPUTE COUNTERFACTUALS?

- Yes! If we know the functions behind the arrows, every counterfactual gets a truth value.

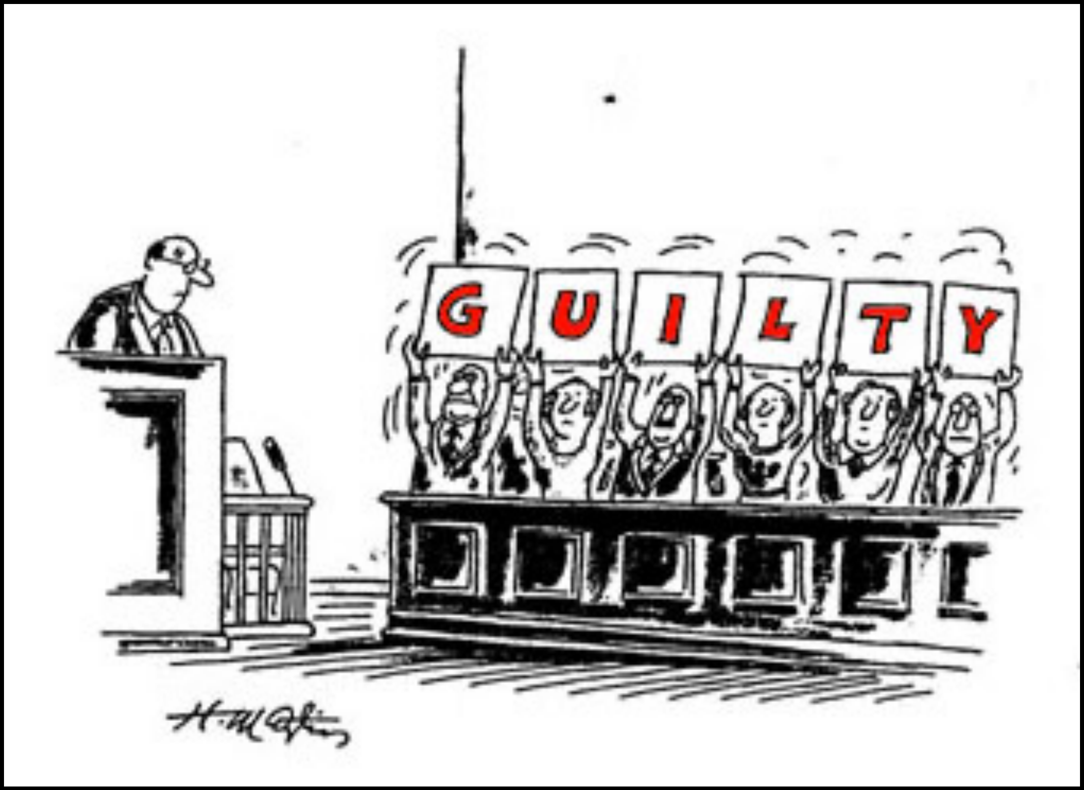
- If we don't, we can bound them using the logic of counterfactuals (Halpern & Pearl).

- The bounds improve when combined data are available and may narrow down sufficiently to reveal individual responsibility.

CAN FREQUENCY DATA DETERMINE LIABILITY?

Sometimes:

When PN is bounded above 0.50.



- WITH PROBABILITY ONE $1 \leq PN \leq 1$
- Combined data reveals individual behavior

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

PNS = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$PNS = P(Y(1) = 1 \;\&\; Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on PNS.

In general, going from group data to individual behavior requires counterfactual logic.

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

PNS = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$PNS = P(Y(1) = 1 \;\&\; Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on PNS.

In general, going from group data to individual behavior requires counterfactual logic.

- Situation-specific decisions

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

PNS = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$PNS = P(Y(1) = 1 \,\&\, Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on PNS.

In general, going from group data to individual behavior requires counterfactual logic.

- Personalized medicine

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

PNS = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$PNS = P(Y(1) = 1 \ \& \ Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on PNS.

In general, going from group data to individual behavior requires counterfactual logic.

- Identify customers worthy of offer/recommendation

IDENTIFYING "PATIENTS IN NEED"

Counterfactual: Patients susceptible to treatment.

PNS = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$PNS = P(Y(1) = 1 \;\&\; Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on PNS.

In general, going from group data to individual behavior requires counterfactual logic.

- Characterize voters swayable by a slogan

IDENTIFYING “PATIENTS IN NEED”

Counterfactual: Patients susceptible to treatment.

PNS = Probability that a patient with characteristics c will improve IF AND ONLY IF treated.

$$PNS = P(Y(1) = 1 \; \& \; Y(0) = 0 \mid C = c)$$

Experimental and observational studies provide informative bounds on PNS.

In general, going from group data to individual behavior requires counterfactual logic.

- Unit Selection: Li, Mueller and Pearl (2021)

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way

Tool 2: Predicting the effects of actions and policies

Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)

Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)

Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)

Tool 6: Recovering from missing Data

Tool 7: Causal Discovery

TOOL 4:
MEDIATION ANALYSIS – DIRECT AND INDIRECT EFFECTS

Task: Given {Data + Model}, unveil and quantify the mechanisms that transmit changes from a cause to its effects.

Wisdom: Counterfactual analysis tells us when direct and indirect effects are estimable from data, and, if so, how necessary (or sufficient) mediation is for the effect.

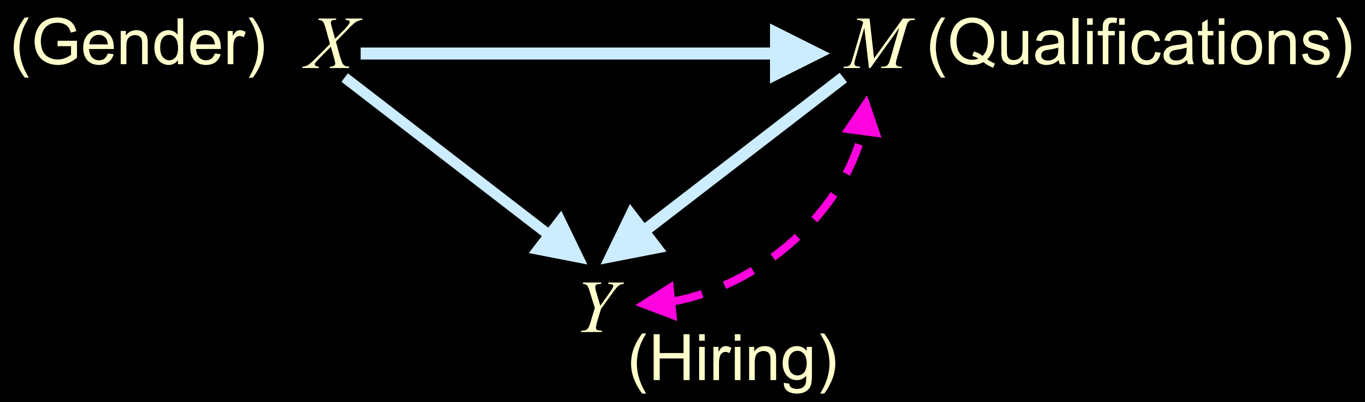
COUNTERFACTUAL DEFINITION OF DESCRIMINATION

"The central question in any employment-discrimination case is whether the employer would have taken the same action had the employee been of a different race (age, sex, religion, national origin, etc.) and everything else had been the same."

(In Carson vs Bethlehem Steel Corp., 70 FEP Cases 921, 7th Cir. (1996).)

LEGAL IMPLICATIONS OF DIRECT EFFECT

Can data prove an employer guilty of hiring discrimination?



What is the direct effect of X on Y?

- $NDE(X,Y) =$ The expected change in Y had X changed and had M been constant at whatever value it attained before the change.
- Meditation formulas, identification, standardization

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way

Tool 2: Predicting the effects of actions and policies

Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)

Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)

Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)

Tool 6: Recovering from missing Data

Tool 7: Causal Discovery

THE DATA FUSION PROBLEM

The general problem

- How to combine results of several experimental and observational studies, each conducted on a different population and under a different set of conditions,
- so as to construct a valid estimate of effect size in yet a new population, unmatched by any of those studied.
- Subproblems: External validity, selection bias

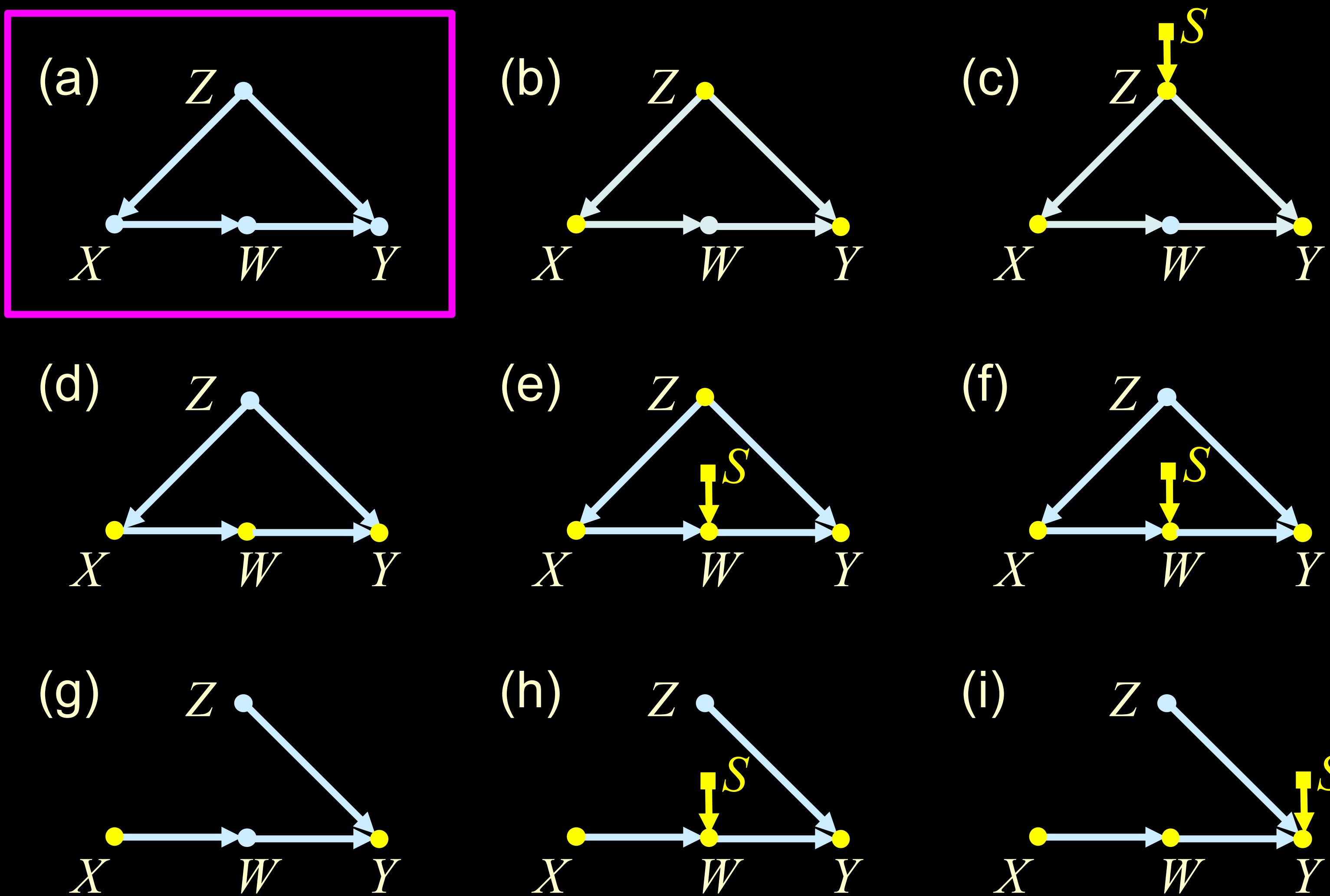
THE PROBLEM IN REAL LIFE

Target population Π^* Query of interest: $Q = P^*(y \mid do(x))$

(a) Arkansas Survey data available	(b) New York Survey data Resembling target	(c) Los Angeles Survey data Younger population
(d) Boston Age not recorded Mostly successful lawyers	(e) San Francisco High post-treatment blood pressure	(f) Texas Mostly Spanish subjects High attrition
(g) Toronto Randomized trial College students	(h) Utah RCT, paid volunteers, unemployed	(i) Wyoming RCT, young athletes

THE PROBLEM IN MATHEMATICS

Target population Π^* Query of interest: $Q = P^*(y \mid do(x))$



WHAT CAN BE FUSED AND HOW?

- Experimental results from multiple sources can be fused provided that commonalities and differences are encoded in selection diagrams.

- When estimation is feasible, a fusion formula can be derived in polynomial time.

- The algorithm is complete.

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way

Tool 2: Predicting the effects of actions and policies

Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)

Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)

Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)

Tool 6: Recovering from missing Data

Tool 7: Causal Discovery

THE SEVEN WISDOMS (TOOLS) OF CAUSAL INFERENCE

Tool 1: Encoding causal information in transparent and testable way
Tool 2: Predicting the effects of actions and policies
Tool 3: Computing counterfactuals and finding causes of effects (attribution, explanation, susceptibility)
Tool 4: Computing direct and indirect effects (Mediation) (discrimination, inequities, fairness)
Tool 5: Integrating data from diverse sources (fusion, transportability, transfer-learning)
Tool 6: Recovering from missing Data
Tool 7: Causal Discovery

CONCLUSIONS

"More has been learned about causal inference in the last few decades than the sum total of everything that had been learned about it in all prior recorded history."

(Gary King, Harvard, 2014)

"The next revolution will be even more impactful upon realizing that data science is the science of interpreting reality, not of summarizing data."

(The Author, UCLA, 2022)

Paper available: http://ftp.cs.ucla.edu/pub/stat_ser/r475.pdf
Refs: http://bayes.cs.ucla.edu/jp_home.html

Every science that has thriven has thriven upon its own symbols
~Augustus de Morgan (1864)

THANK YOU

Joint work with:
Elias Bareinboim
Karthika Mohan
Ilya Shpitser
Jin Tian
Many more . . .

For a trailer, click WHY on my home page.