Statistical Learning in Mediation Analysis

Chapter 1: Introduction to causal inference and mediation analysis

David Benkeser
Emory University

Iván Díaz
New York University

Marco Carone
University of Washington

MODULE 13

Summer Institute in Statistics for Clinical and Epidemiological Research

July 2023

Contents of this chapter

1. What is causal inference?
2. Why are DAGs and SWIGs useful for causal inference?
3. What is causal mediation analysis?
4. What types of effects and estimators will we see in this course?

Causal inference

Statistical inference asks questions about "what is"?

- What is the risk of COVID-19 amongst those who are recently vaccinated?
- What is the average length of life following a cancer diagnosis among patients using an experimental treatment?

Causal inference asks questions about "what would be"?

- What would the risk of COVID-19 be for a recently vaccinated individual be if 90% of the population were vaccinated?
- What would the average length of life following a cancer diagnosis if the experimental treatment became standard of care?

Causal inference

We are very often interested in causal inference in biomedical science, but...

Data alone only tell us what is and not what would be.

The field of causal inference provides a framework for:

- codifying our knowledge about relationships between variables in the system that we are studying;
- conceptualizing a hypothetical experiment that would generate the data capable of answering our causal question of interest;
- assessing whether the data collected in the real world can be used to such an end;
- teaching us how to describe "what is" in such a way as to infer "what would be".

Encoding knowledge into DAGs

It is critical to map out the relationship between all key variables involved in studying a particular scientific question. Why? For example, such information can be used:

- to determine what variables must be adjusted for;
- to determine what variables should not be adjusted for;
- to determine what variables could be used to gain precision in analyses.

Causal diagrams can be used for this purpose. They also allow investigators to encode prior knowledge about the relationship under study.

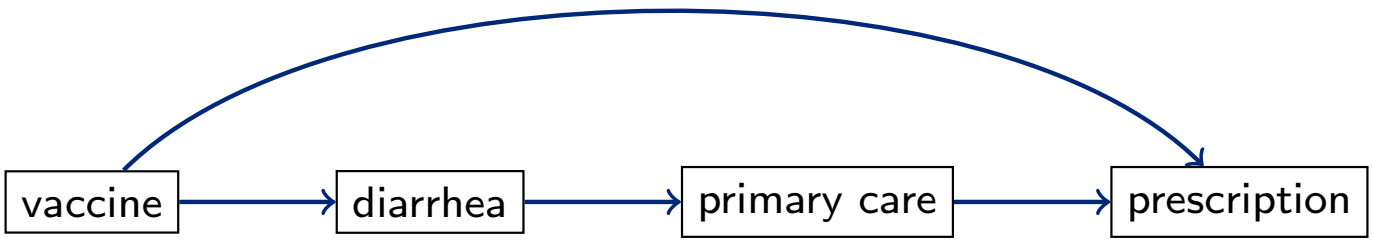
A causal diagram is made up of:

- nodes = variables (observed or not);
- arrows = edges + direction = directed causal relationships.

A graph is a **directed acyclic graph** (DAG) if it is impossible to find a nontrivial path from a node onto itself following only the direction of arrows.

Example DAG

Example: The relationship between rotavirus vaccination and antibiotic prescribing.



Encoding knowledge into DAGs

Guidelines for constructing or interpreting a causal DAG:

- absence of arrow = assumption of no causal relationship;
- any common cause of two nodes in a causal DAG must be included in the graph, even if it is not actually measured;
- all arrows should be unidirectional (else, incorporate time-dependent nodes);

Using DAGs to infer (conditional) independence:

- two nodes can be related yet neither is the cause of the other:
  - (A $\longleftarrow$ B $\longrightarrow$ C) B is a confounder;
  - (A $\longrightarrow$ B $\longleftarrow$ C) B is a collider;
- types of path between variables A and B:
  - directed path: sequences of edges from A to B following direction of arrows;
  - backdoor path: sequences of edges from A to B with at least one edge pointing into A;
- a path is blocked if it includes a collider;
- a path can be blocked by controlling (i.e., fixing) one of its nodes;
- controlling a collider opens a (undirected) path between its parents.

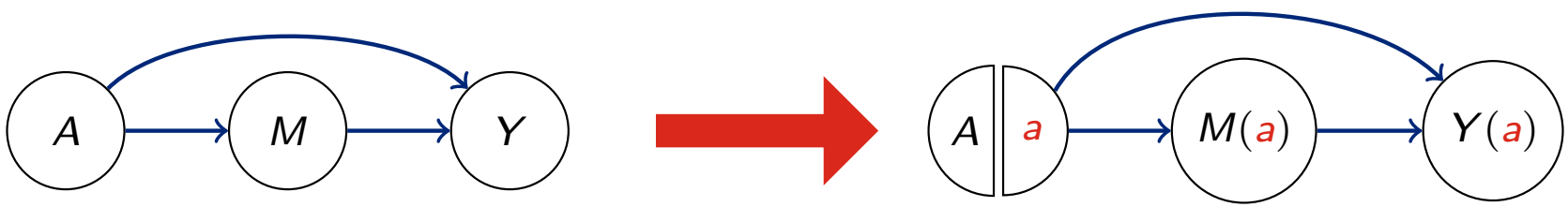
Conceptualizing ideal experiments

The DAG describes what is, a SWIG describes what would be.

- Richardson and Robins introduced the idea of a single world intervention graph.
- "Split" nodes in a DAG to describe your experiment.
- SWIGs unify causal graphs and counterfactual variables.

Example of a DAG → SWIG:

A = exposure M = mediator Y = outcome



Counterfactual variables

Counterfactual variables describe data that would be seen under intervention.

- Also known as **potential outcomes**.

Example: Suppose A, M and Y are all binary (0/1) variables. Let

- $M(0)$ = mediator that would be observed under intervention $A = 0$.
- $M(1)$ = mediator that would be observed under intervention $A = 1$.
- $Y(0)$ = outcome that would be observed under intervention $A = 0$.
- $Y(1)$ = outcome that would be observed under intervention $A = 1$.

Patient	M(1)	M(0)	Y(1)	Y(0)
1	1	0	1	1
2	0	1	0	0
3	1	0	0	1
4	0	0	1	0
...	...	...	...	...

Counterfactual variables

There are two **implicit assumptions** in our notation for counterfactuals.

No interference

- $Y(1)$ is Your outcome if You are given $A = 1$.
- $Y(1)$ only depends on the treatment You are given.
- Counterexamples: indirect vaccine effects, social network interference

Consistency

- There are not multiple forms of treatment.
- In the real world, we observe $A = 1, Y$.
- In the counterfactual world, we set $A = 1$ and observe $Y(1)$.
- Consistency states that $Y = Y(1)$.
- Counterexamples: interventions on "clinical score", BMI

Causal effects

Counterfactual variables are random due to sampling from a population. In other words, they have some distribution.

- Example: $P[Y(1) = 1]$ = proportion with outcome value 1 if given treatment.
- Example: $P[Y(1) = 1, M(1) = 1]$ = proportion with mediator value 1 and outcome value 1 if given treatment.
- These are examples of causal parameters.

A causal effect contrasts causal parameters under different interventions.

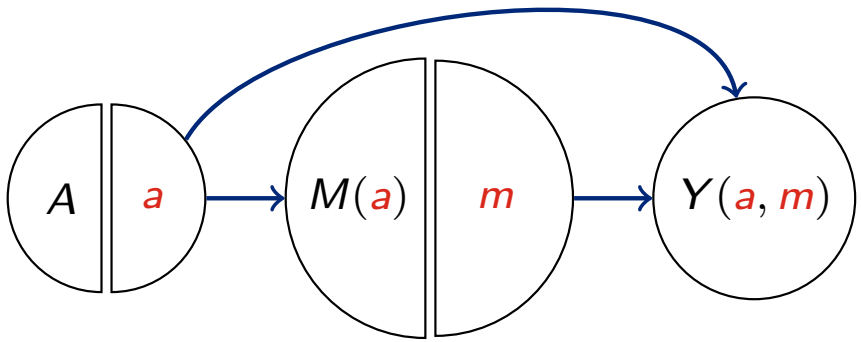
Examples:

- Average treatment effect: $E[Y(1) - Y(0)]$
- Causal risk ratio: $P[Y(1) = 1]/P[Y(0) = 1]$
- Causal odds ratio: $\{P[Y(1) = 1]/P[Y(1) = 0]\}/\{P[Y(0) = 1]/P[Y(0) = 0]\}$

Causal effects

Questions of mediation will generally involve two interventions:

1. one on the exposure/treatment (A);
2. the second on a mediator (M).



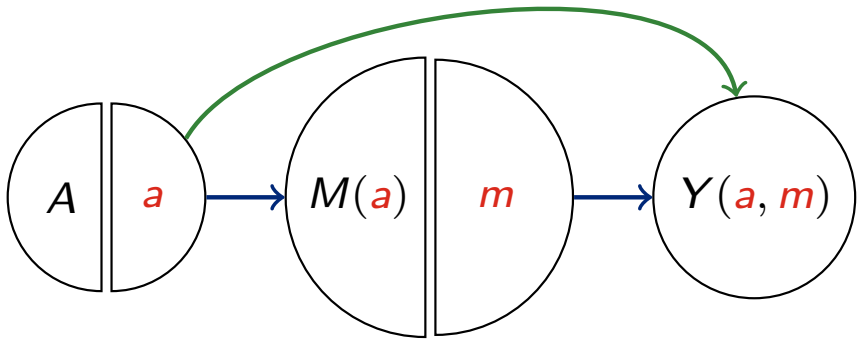
$Y(a, m)$ = the outcome that would be observed if A were set to a and M to m.

Causal mediation effects will be defined based on different choices of a and m.

Direct and indirect effects

We are often interested in understanding

- the **indirect effect** of A on Y, and/or
- the **direct effect** of A on Y.



An **indirect effect** describes impact of A on Y operating through M.

- **Example**: Does a COVID-19 vaccine (A) prevent disease (Y) because it impacts neutralizing antibodies (M)?
- Comparing distribution of $Y(a, m)$ vs. $Y(a, m^*)$

A **direct effect** describes impact of A on Y through pathways that **do not** involve M.

- **Example**: If we prevent COVID-19 vaccines from inducing any neutralizing antibodies, would they still prevent disease?
- Comparing distribution of $Y(a, m)$ vs. $Y(a^*, m)$.

Causal mediation effects

Example: Is the impact of type 2 diabetes (A) on all-cause mortality (Y) attributable to its impact on heart failure (M)?

The relevant counterfactuals are:

- $Y(0,0)$ = mortality if no diabetes and no heart failure;
- $Y(1,0)$ = mortality if diabetes and no heart failure;
- $Y(0,1)$ = mortality if no diabetes and heart failure;
- $Y(1,1)$ = mortality if diabetes and heart failure.

Question: What is the effect of diabetes on mortality "fixing" heart failure status?

- Compare $P[Y(1,0) = 1]$ vs. $P[Y(0,0) = 1]$?
- This is a controlled direct effect.

Causal mediation effects

Example: Is the impact of type 2 diabetes (A) on all-cause mortality (Y) attributable to its impact on heart failure (M)?

The relevant counterfactuals are:

- $M(1)$ = "natural" heart failure status if diabetes;
- $M(0)$ = "natural" heart failure status if no diabetes;
- $Y(1, M(1))$ = mortality if diabetes and natural heart failure if diabetes;
- $Y(1, M(0))$ = mortality if diabetes and natural heart failure if no diabetes;

Question: What is was the impact of diabetes-induced heart failure on mortality?

- Compare $P[Y(1, M(1)) = 1]$ vs. $P[Y(1, M(0)) = 1]$?
- This is a natural indirect effect.

Roadmap for causal inference

Fundamental problem of causal inference:

For each patient, we get to see only one of the counterfactuals.

The data alone cannot tell us how outcomes would change under different interventions on the exposure and mediator.

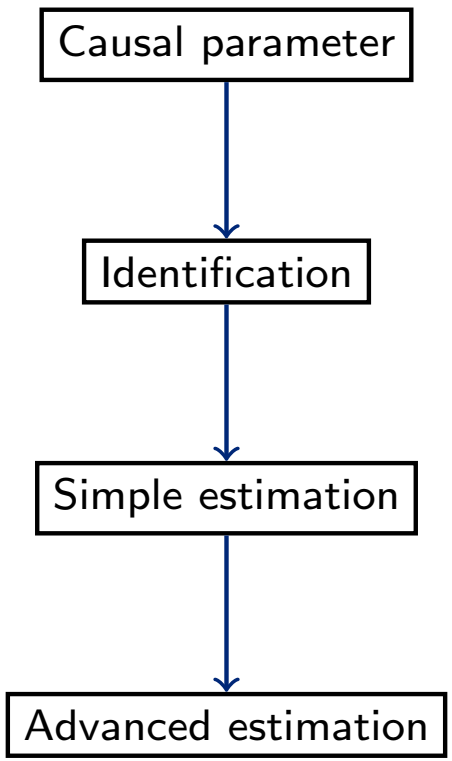
Consequently, causal analysis drawn from all observational studies and many clinical trials will rely on untestable assumptions about confounding.

We will hear about so-called randomization assumptions.

- Do we have enough data captured to disentangle confounding between the exposure, the mediator, and the outcome?
- DAGs/SWIGs can be useful to this end!
- Unforunately, randomized designs are infeasible for many mediation problems.

Roadmap for causal inference

General roadmap for each chapter:



References and additional reading

References:

Richardson TS and Robins JR (2013). Single world intervention graphs: a primer. *Second UAI workshop on causal structure learning, Bellevue, Washington*. [link].

Additional reading:

Nguyen TQ, Schmid I, Stuart EA (2021). Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. *Psychological Methods*. doi: 10.1037/met0000299

Breskin A, Cole SR, Hudgens MG (2018). A Practical Example Demonstrating the Utility of Single-world Intervention Graphs. *Epidemiology*. doi: 10.1097/EDE.0000000000000797.