Causal diagrams (DAGs)

ADVANCED EPIDEMIOLOGY

22 JANUARY 2019

REGINA PRIGGE: S1570624@SMS.ED.AC.UK

COLIN FISCHBACHER: COLIN.FISCHBACHER@NHS.NET

Why do we need causal diagrams?

- Interest: The relationship between depression and subsequent cardiovascular diseases
- Literature
  - 1. To define hypotheses
  - 2. To identify other variables that have a relationship with depression as well as with CVD:
    - Physical activity, alcohol intake, menopausal status, age, sex, ethnicity, diet, weight gain, different medications, biological dysregulations, and many more...
- Need to communicate to others what my theory is
  - Can be difficult to explain in words
  - Causal diagram as a way of making clear what our theories and assumptions are

Usefulness of causal diagrams

- Identifies variables relevant to research
- Summarize knowledge
- Visualize assumptions
- Graphic representation of causal network
- Enhance communication among researchers

Directed acyclic graphs (DAGs)

- Directed: Edges (arrowheads) imply a direction

$$\underset{\text{(treatment)}}{A} \longrightarrow \underset{\text{(outcome)}}{Y}$$

- Acyclic: A variable cannot cause itself

$$A \rightleftarrows Y$$

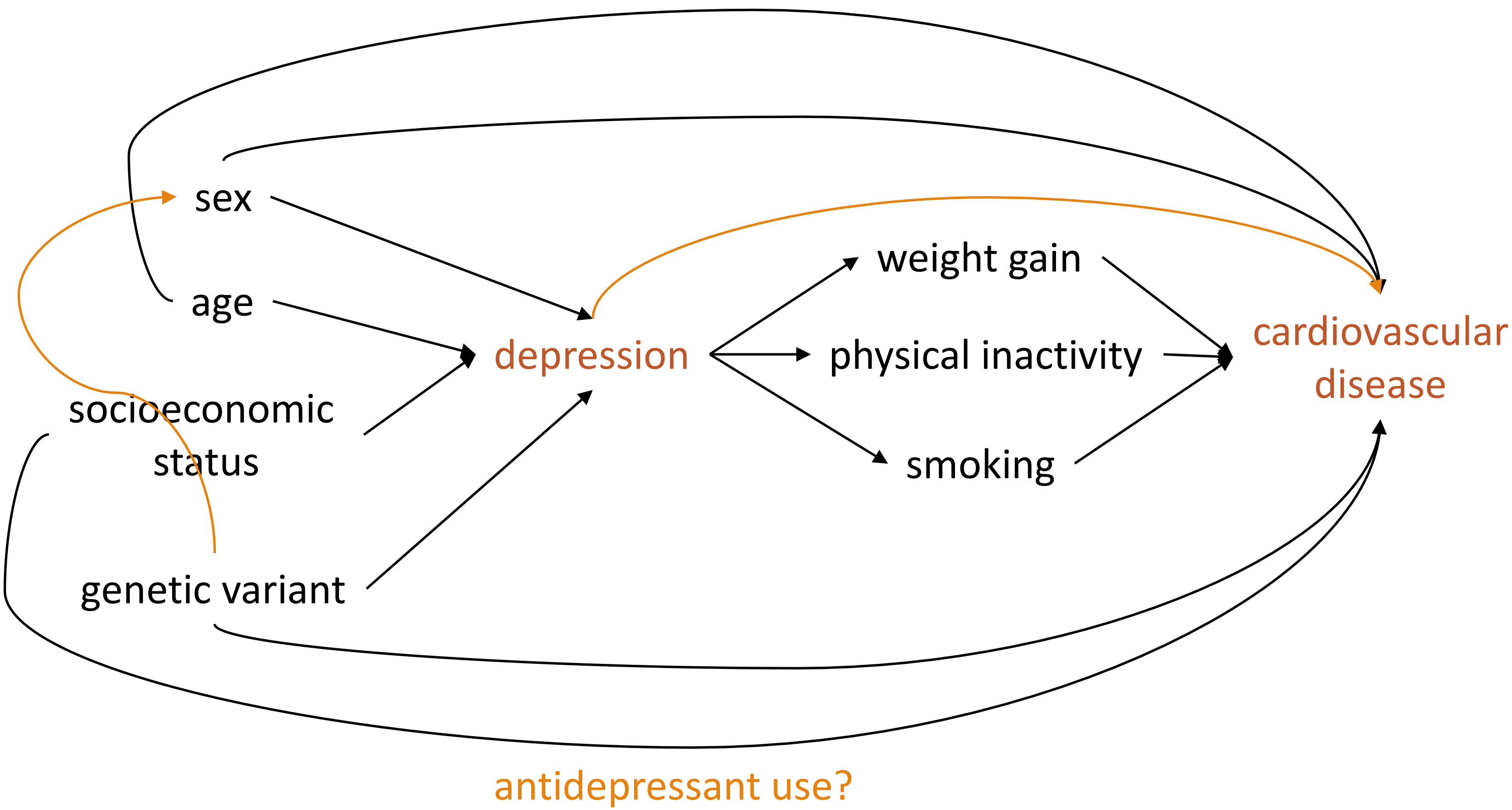
Visualization of DAGs

- Presence of arrow:
  - We assume direct causal effect
  - We are not willing to assume that causal effect does not exist
- Absence of arrow = strong assumption:
  - We are willing to assume that causal effect does not exist
- Direction of arrow:
  - Assumed direction of effect
- Time flows from left to right
- We do not distinguish between harmful and protective effects

$$A \xrightarrow{\text{causes}} Y$$

earlier in life → later in life

What does this DAG tell us?



Ideal RCTs

- What do I mean by “ideal” RCT?
  - Treatment groups are exchangeable
  - No loss to follow-up
  - Double-blinding
  - Perfect adherence to treatment strategies
- Unconditional exchangeability

$$A \longrightarrow Y$$

A	Y
exposure, eg a drug	outcome, eg heart attack

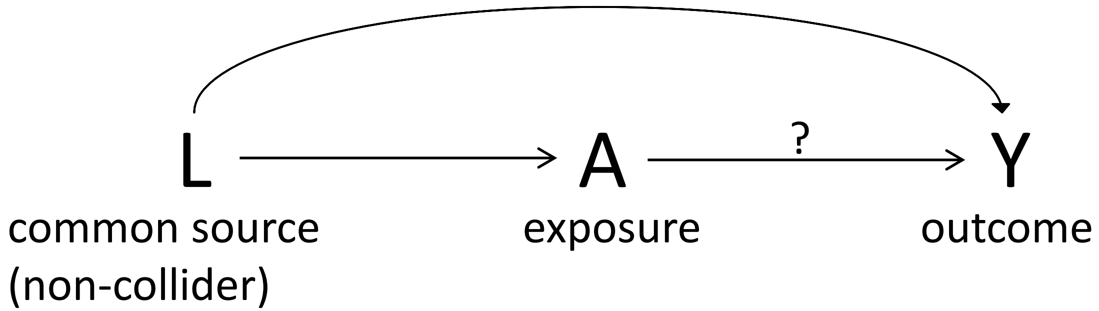
Non-ideal RCTs and observational studies

- Conditional exchangeability (more than two variables)

- What type of other variables could there be?
  - Common sources
  - Common effects
  - Mediators

Common sources – non-collider

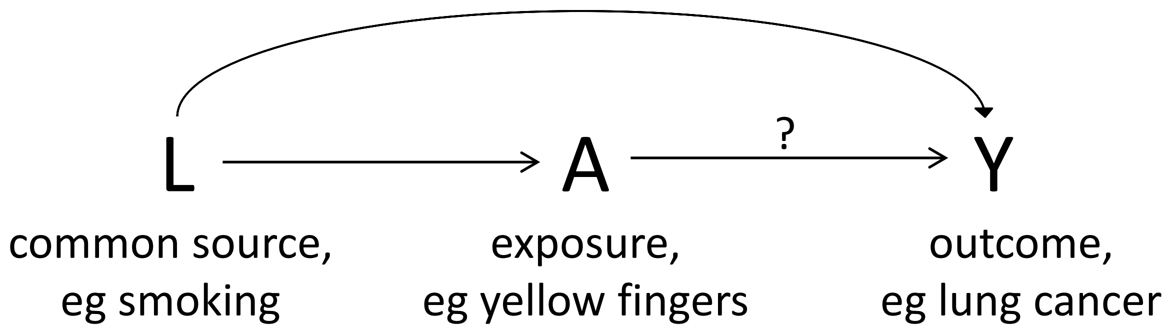
- H0: There is no causal relationship between exposure A and outcome Y, given variable L



- Association flows between variables regardless of the direction of the causal arrows (association from A→L→Y)
  - "back-door path" is open

Common sources – non-collider

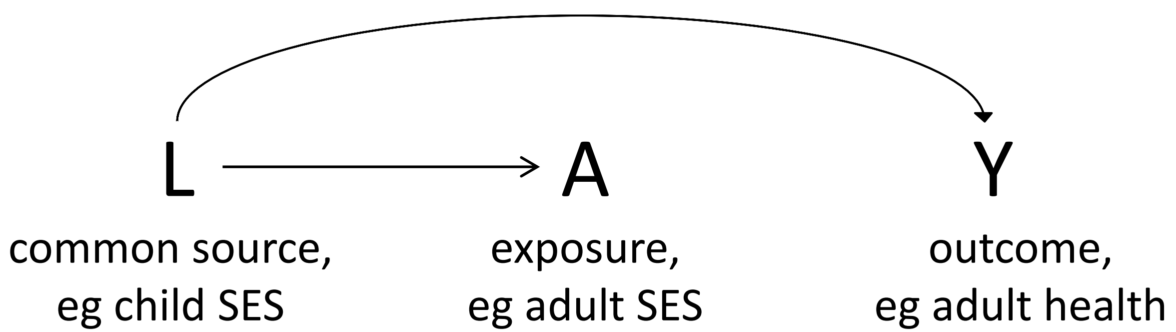
- H0: There is no causal relationship between exposure A and outcome Y, given variable L



- “Back-door path” is open
- We can only estimate causal effects if there is no open back-door path

Common sources – non-collider

- H0: There is no causal relationship between exposure A and outcome Y, given variable L

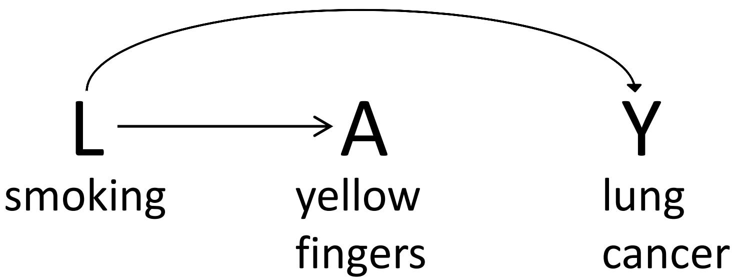


- “Back-door path” is open
- We can interpret our estimates causally if there is no open backdoor-path

Open and blocked paths – non-collider

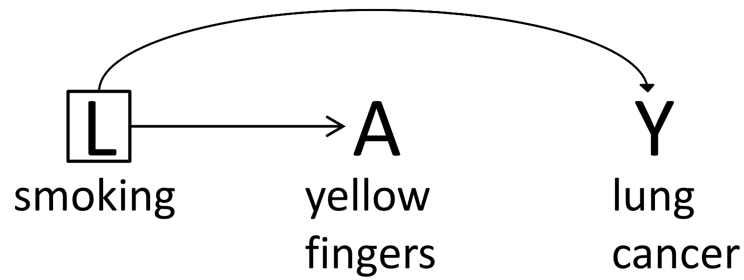
OPEN PATH

- No conditioning on non-collider (common source)



BLOCKED PATH

- Conditioning on non-collider (common source)

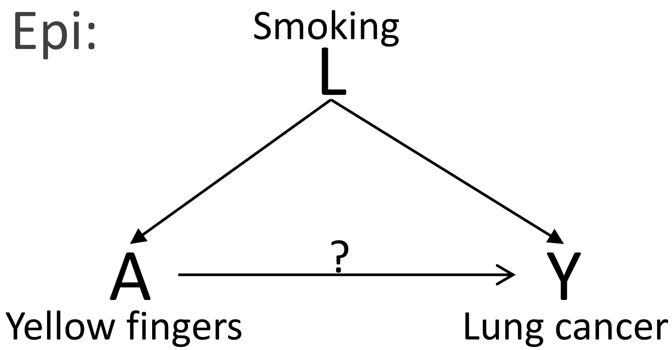


Conditioning

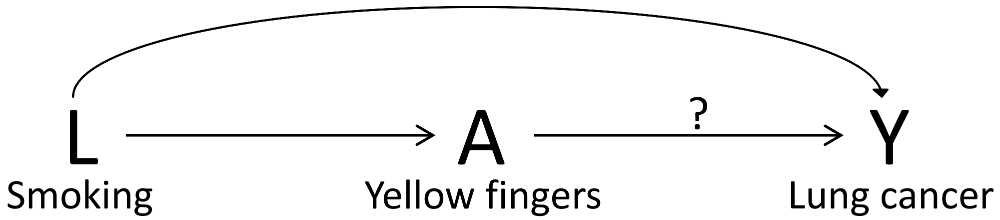
ADJUSTMENT	STRATIFICATION	RESTRICTION
"In our statistical analysis we adjusted our results for smoking"	"We report our results separately for smoker and non-smoker"	"We only selected participants that were smokers"
Effect estimates:	Effect estimates:	Effect estimates:
Unadjusted (crude): ◦ OR 1.22 (1.14-1.56)	Smoker: ◦ OR: 1.48 (1.36-1.57)	Smoker: ◦ OR: 1.48 (1.36-1.57)
Adjusted for smoking: ◦ OR 1.15 (1.05-1.25)	Non-smoker: ◦ OR: 1.21 (1.15-1.34)	

Confounding

- Confounding diagram in Intro to Epi:



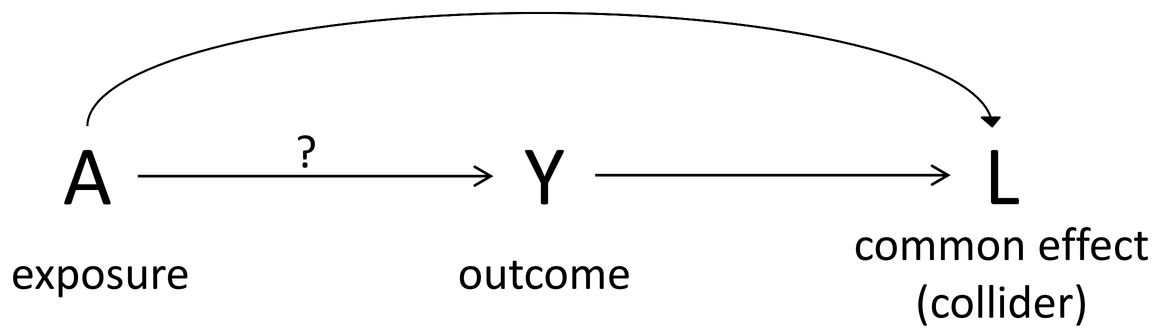
- This is what our diagram looked like when we did not condition on the common source (non-collider):



- We did not close an open “back-door path“

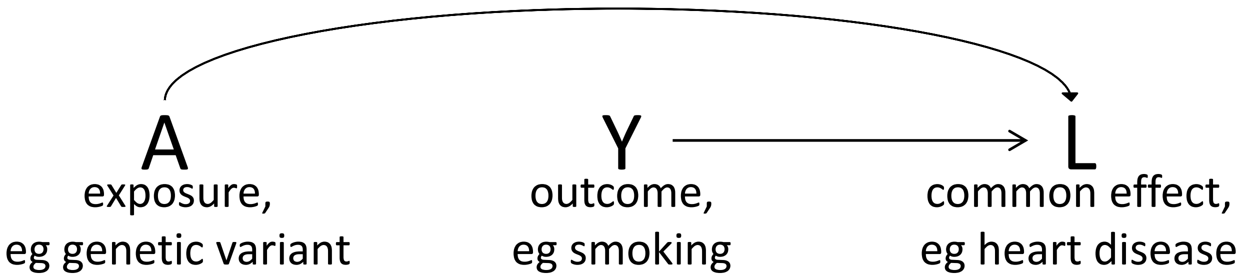
Common effect - collider

- H0: There is no causal relationship between exposure A and outcome Y, given variable L



Common effect - collider

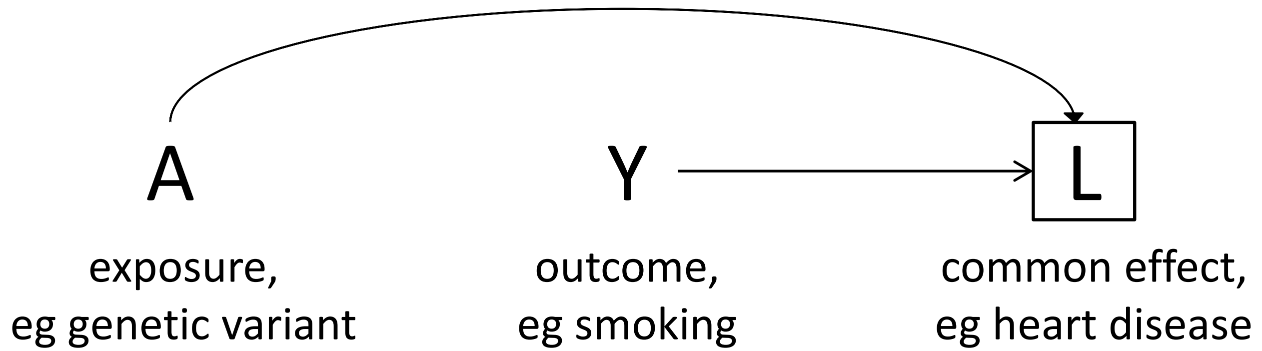
- H0: There is no causal relationship between exposure A and outcome Y, given variable L



- Association does not flow between variables when two arrowheads point towards a variable (no association from A→L←Y)
  - “Back-door path” is blocked
  - We can interpret our estimates causally if there is no open backdoor-path

Common effect - collider

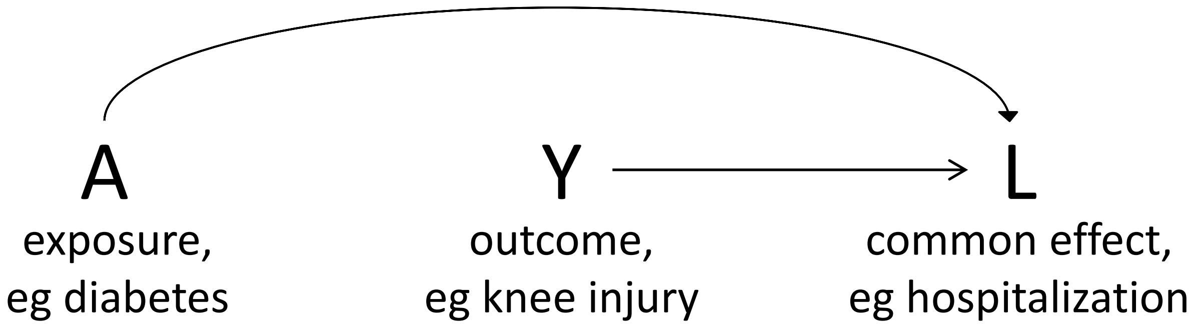
- If we restrict our sample to participants with heart disease…



- Association does flow between variables when two arrowheads point towards a variable, and this variable was conditioned on (association from A→$\boxed{\text{L}}$←Y)
  - “back-door path” is open

Common effect - collider

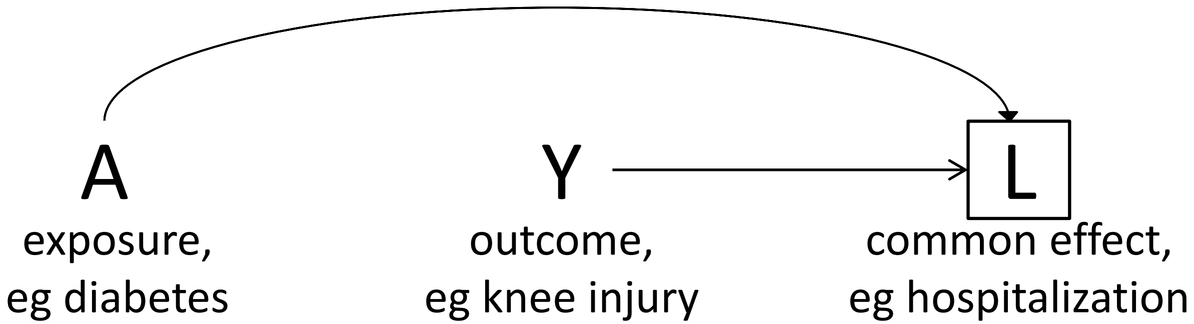
- H0: There is no causal relationship between exposure A and outcome Y, given variable L



- Association does not flow between variables when two arrowheads point towards a variable (no association from A→L←Y)
  - “back-door path” is closed

Common effect - collider

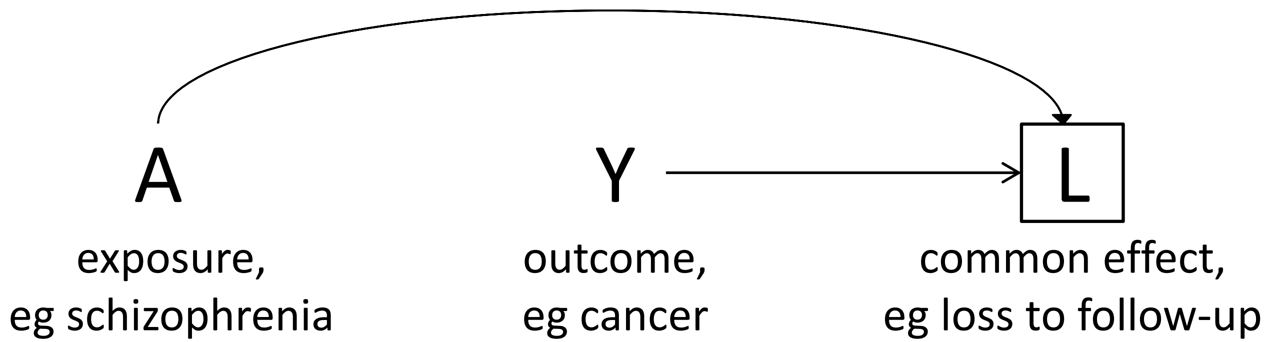
- H0: There is no causal relationship between exposure A and outcome Y, given variable L



- "Back-door path" is open
- Estimates cannot be interpreted causally

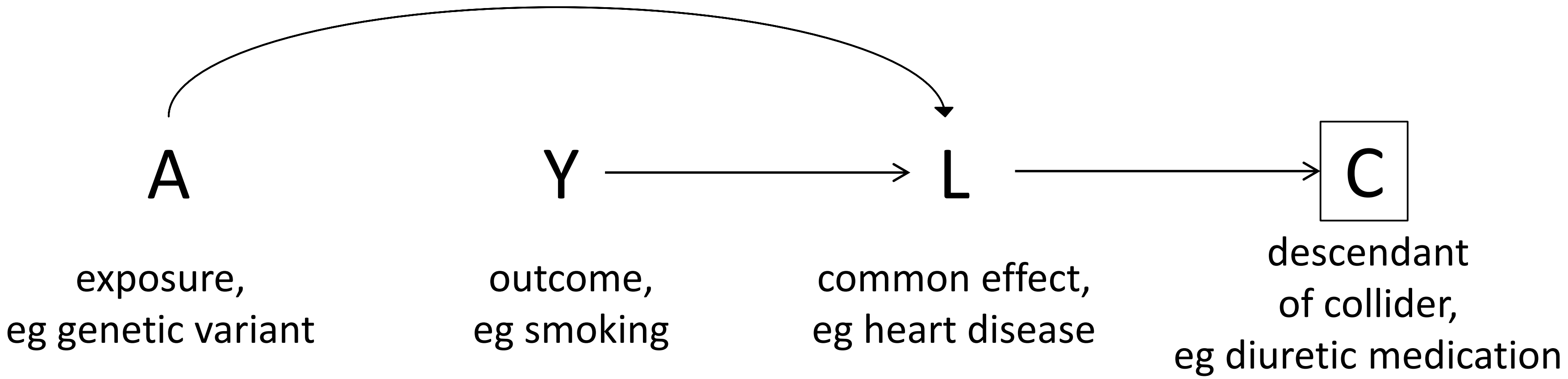
Common effect - collider

- We might unwillingly condition on a collider because we lose information of those lost to follow-up



- Association does flow between variables when two arrowheads point towards a variable, and this variable was conditioned on (association from A→$\boxed{L}$←Y)
  - “back-door path” is open

Conditioning on descendent of a collider

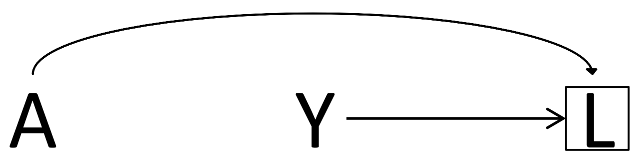


- Association flows between variables when it was conditioned on the descendent of a collider
  - “back-door path” is open
- Think of descendent of collider as common effect of A and Y through L

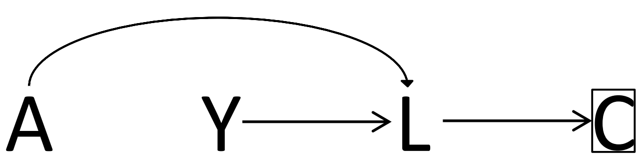
Open and blocked paths – collider

OPEN PATH

- Conditioning on collider (common effect)

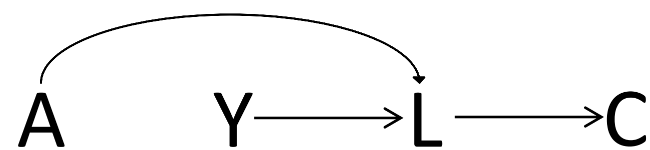


- Conditioning on effect of collider (common effect)



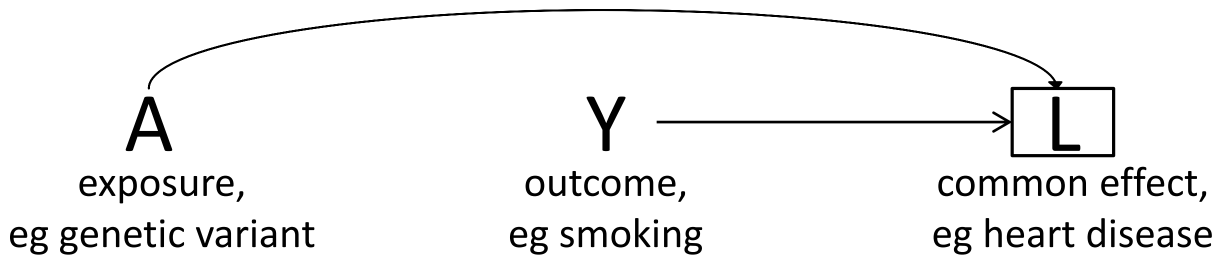
BLOCKED PATH

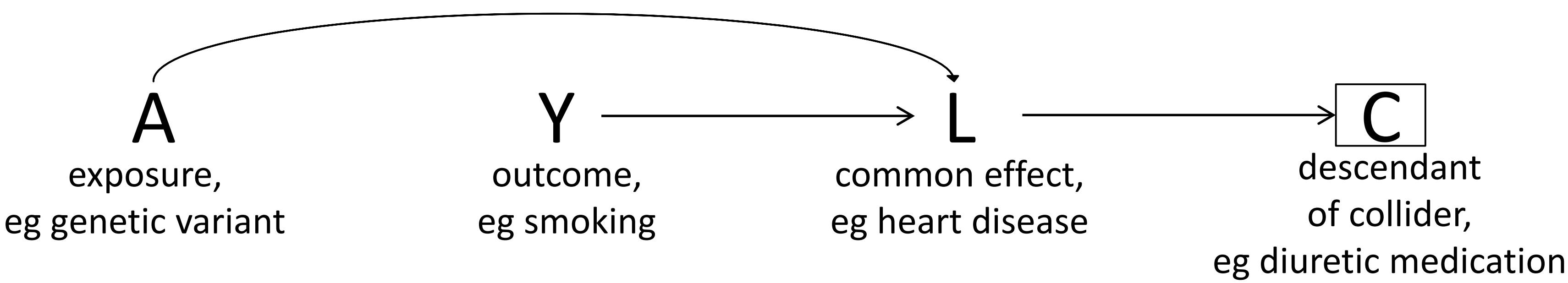
- No conditioning on collider AND no conditioning on effect of collider



Selection bias

- Conditioning on common effect (collider) or conditioning on descendent of collider





- We opened a closed “back-door path”

Confounding vs selection bias

Confounding:

- Not conditioning on common source (non-collider) → we did not close an open “back-door path“

$$\begin{tikzcd} L \arrow[r] \arrow[rr, bend left] & A & Y \end{tikzcd}$$

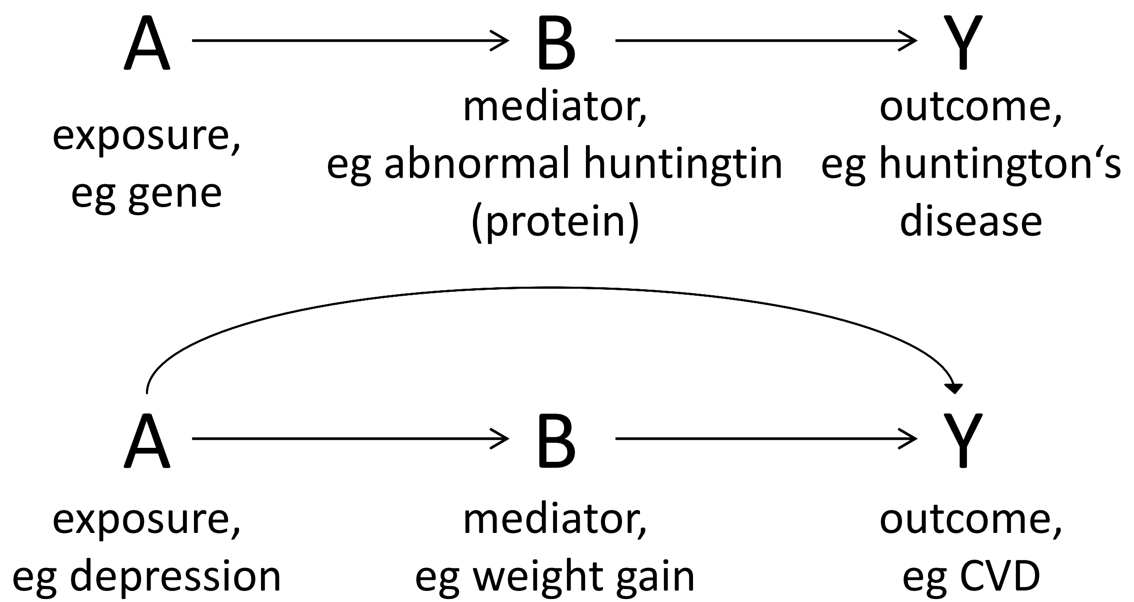
Selection bias:

- Conditioning on common effect (collider) or conditioning on effect of collider → we opened a closed “back-door path”

$$\begin{tikzcd} A \arrow[rr, bend left] & Y \arrow[r] & \boxed{L} \end{tikzcd} \qquad \begin{tikzcd} A \arrow[rr, bend left] & Y \arrow[r] & L \arrow[r] & \boxed{C} \end{tikzcd}$$

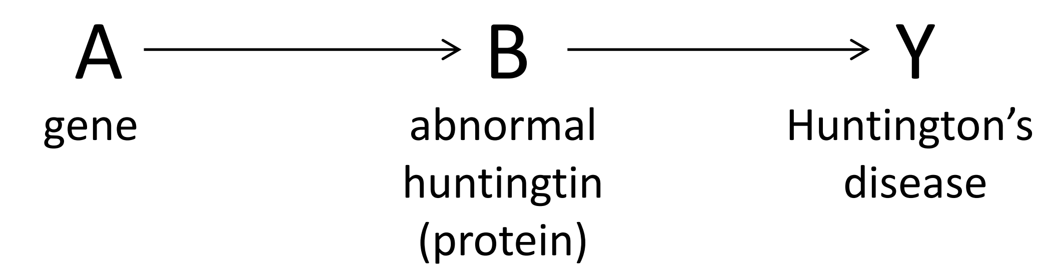
Mediation

- Factor is on the causal pathway

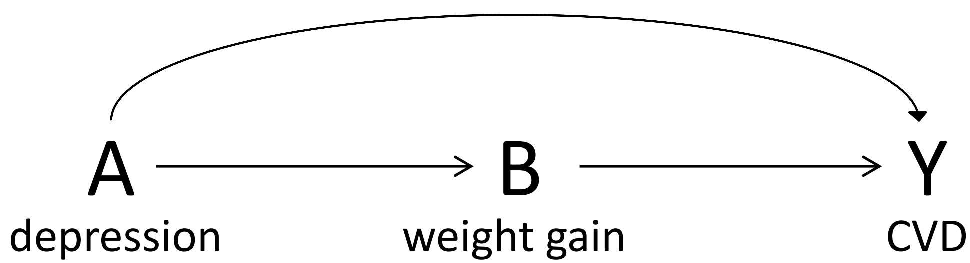


- What do we want to do here?

Mediation – Estimation of total effect

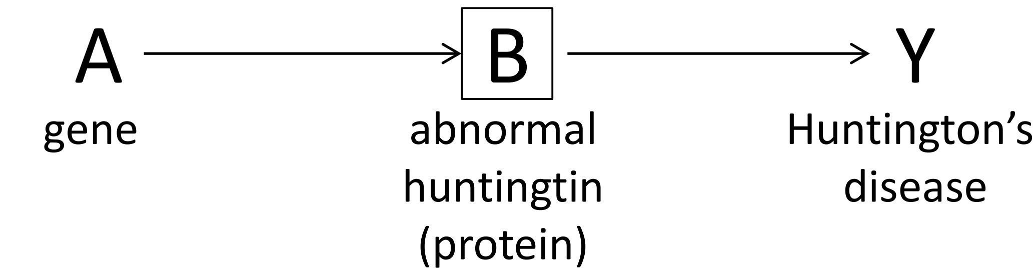


complete mediation

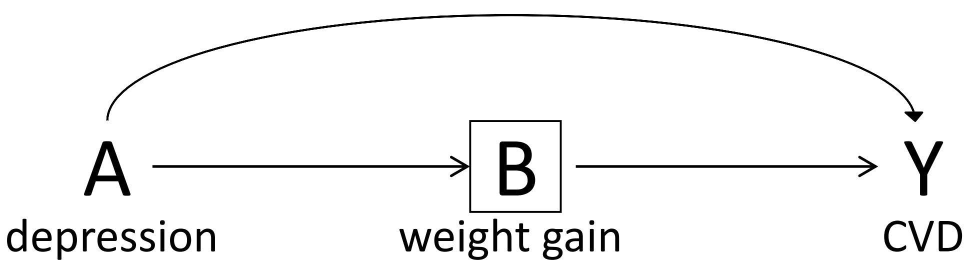


incomplete mediation

Mediation – Estimation of indirect and direct effect

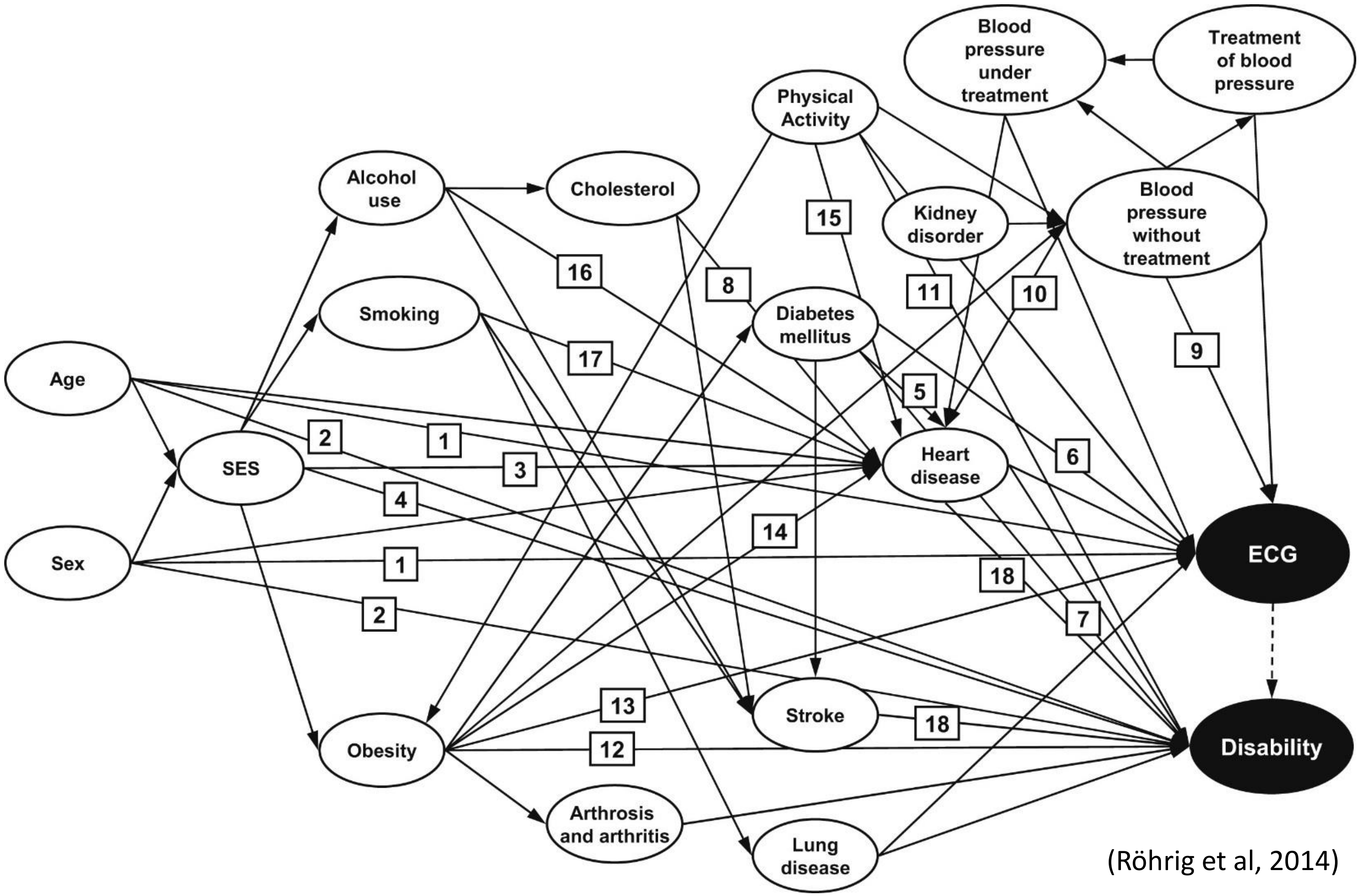


complete mediation



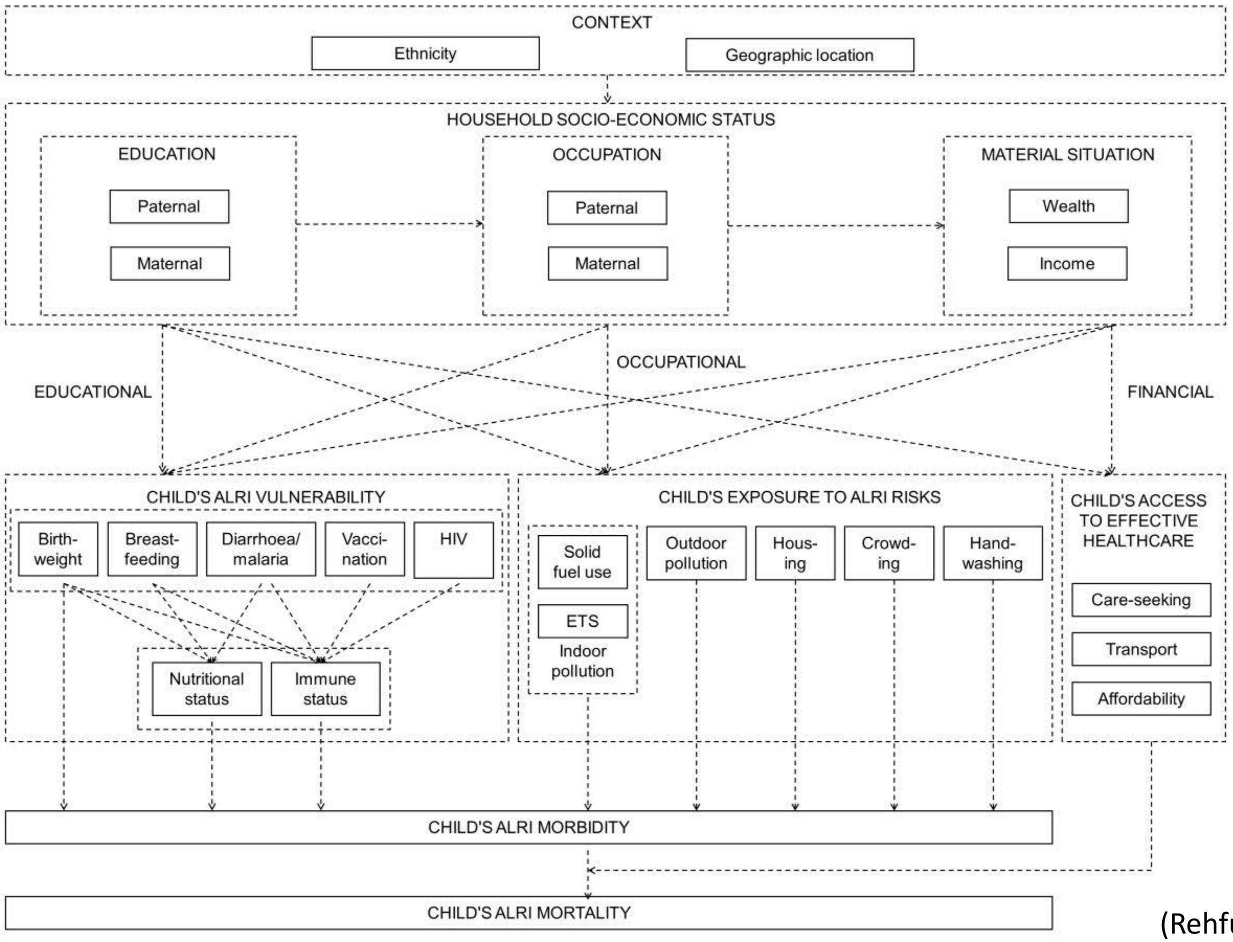
incomplete mediation

The world is more complicated than that...



(Röhrig et al, 2014)

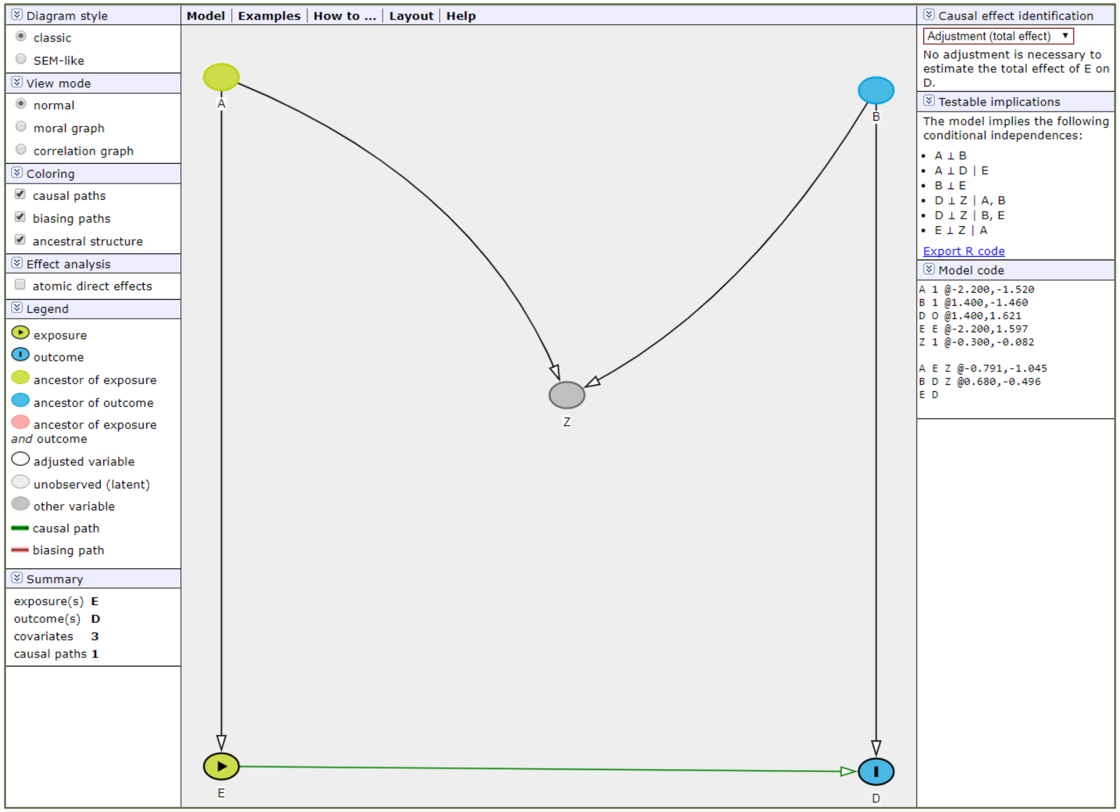
A conceptual causal diagram



(Rehfuess et al, 2013)

DAGitty

- Free software for drawing causal diagrams
- Helps you identify the variable you have to condition on in order to interpret your estimate causally
- http://www.dagitty.net/dags.html#



References

Greenland, S., Pearl, J. and Robins, J.M., 1999. Causal diagrams for epidemiologic research. *Epidemiology*, pp.37-48.

Hernán, M.A., Hernández-Díaz, S. and Robins, J.M., 2004. A structural approach to selection bias. *Epidemiology*, *15*(5), pp.615-625.

Hernán, M.A. and Robins, J.M., (forthcoming). Graphical representation of causal effects. In: *Causal inference*. Boca Raton, FL: CRC, p.69-82.

Hernán, M.A. and Robins, J.M., (forthcoming). Confounding. In: *Causal inference*. Boca Raton, FL: CRC, p.83-94.

Hernán, M.A. and Robins, J.M., (forthcoming). Selection bias. In: *Causal inference*. Boca Raton, FL: CRC, p.95-108.

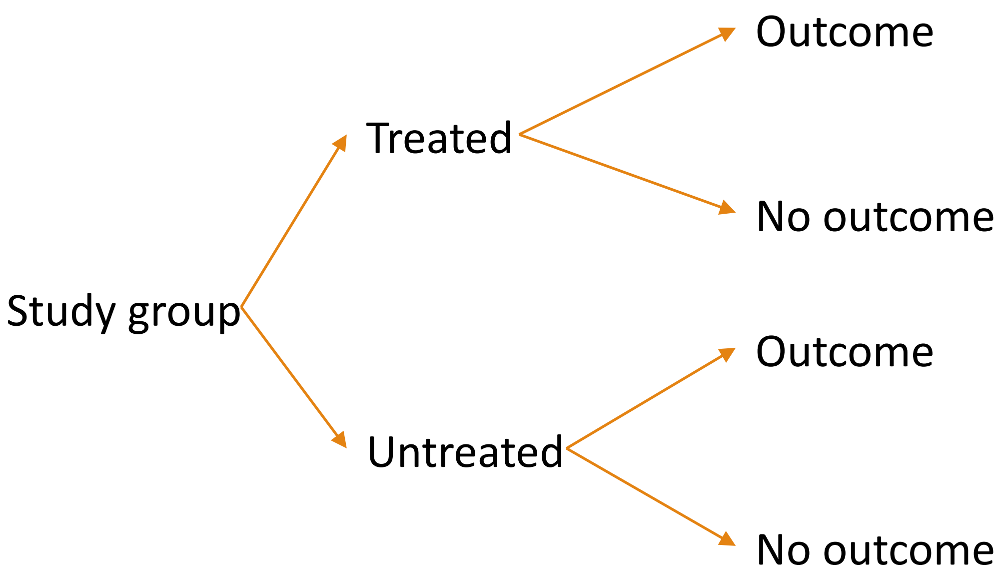
Textor, J., Hardt, J. and Knüppel, S., 2011. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology*, *22*(5), p.745.

Rehfuess, E., Best, N., Briggs, D. and Joffe, M. (2013). Diagram-based Analysis of Causal Systems (DACS): elucidating inter-relationships between determinants of acute lower respiratory infections among children in sub-Saharan Africa. *Emerging Themes in Epidemiology*, 10(1), p.13.

Röhrig, N., Strobl, R., Müller, M., Perz, S., Kääb, S., Martens, E., Peters, A., Linkohr, B. and Grill, E. (2014). Directed acyclic graphs helped to identify confounding in the association of disability and electrocardiographic findings: results from the KORA-Age study. Journal of Clinical Epidemiology, 67(2), pp.199-206.

Recap - Unconditional exchangeability

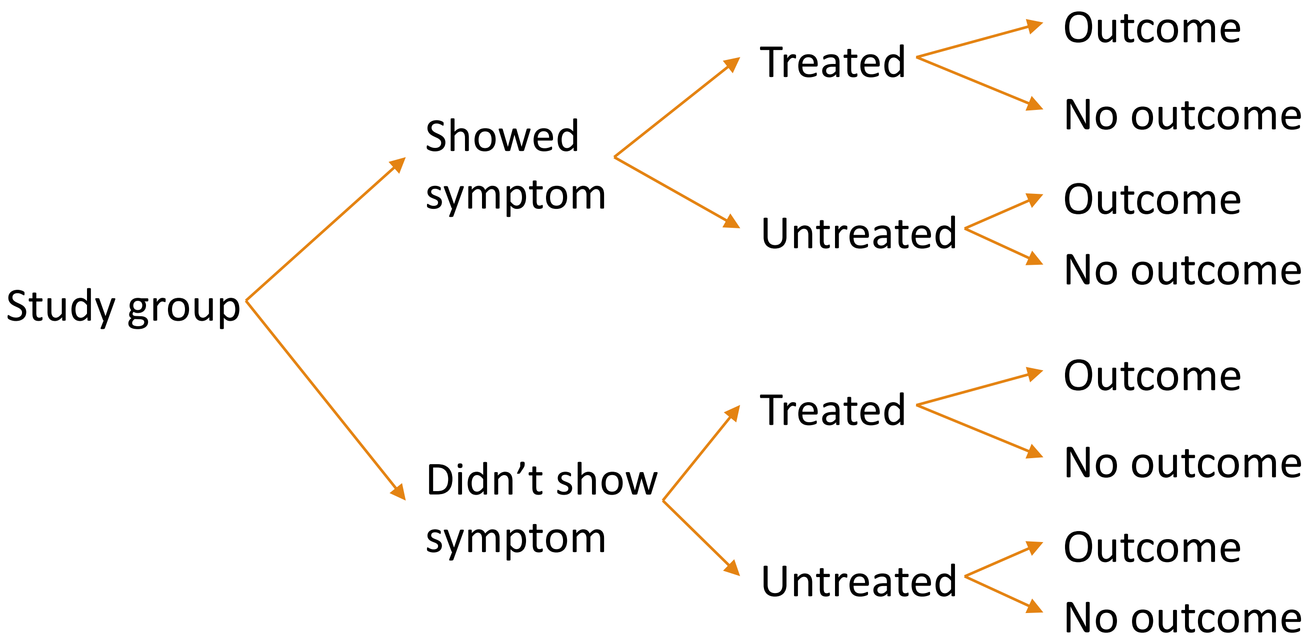
- Achieved through randomization
- Groups are <u>equal</u> in all aspects other than their exposure status



- Treatment effect would have been the same among the untreated if they had been treated

Recap - Conditional exchangeability

- Groups are <u>different</u> in aspects other than their exposure status



- Within each strata of symptom status, treatment effect would have been the same among the untreated if they had been treated