

# Counterfactuals and Directed acyclic graphs in epidemiology

JA Labrecque

Erasmus Medical Center, Rotterdam, The Netherlands

23 mei, 2018

# Causal inference

- ▶ The goal of causal inference: to know how outcome  $Y$  would change if we changed exposure  $A$  from  $a = 0$  to  $a = 1$
- ▶  $Y^a$ : the value outcome  $Y$  would take if we set  $A$  to  $a$
- ▶ So  $Y^{a=1}$  is the value  $Y$  would take if we set  $a = 1$  (N.B.: this is not  $Y|A=1$ )
- ▶ In a dataset we observe *at most* half of the vector  $Y^a$  because we can only observe each person as  $Y^{a=1}$  or  $Y^{a=0}$
- ▶ We write the causal effect we're interested in as  $Y^{a=1} - Y^{a=0}$   
or  $\frac{Pr[Y^{a=1}=1]}{Pr[Y^{a=0}=1]}$

# Counterfactuals

- ▶ Counterfactual notation has been used to achieve insight into estimation of longitudinal causal effects, instrumental variables and a number of other topics
- ▶ Some of the insight provided by counterfactuals can also be depicted in graphical form

# Directed Acyclic Graphs (DAGs)

 Questions Developer Jobs Tags Users Search...

 1    

Can someone explain in simple terms to me what a Directed acyclic graph is?

Ask Question



70



30

Can someone explain in simple terms to me what a Directed acyclic graph is? I have looked on Wikipedia but it doesn't really make me see its use in programming.

directed-acyclic-graphs

share edit

edited Mar 2 '10 at 2:15



polygenelubricants  
257k • 92 • 491 • 583

asked Feb 17 '10 at 19:26



Zubair  
20.2k • 46 • 190 • 327

17 Wikipedia frequently contains overwhelming technical content that would take beginners a great deal of studying to comprehend. Many of the math help sites are superior in this regard, but they tend not to get into computation related subjects, unfortunately. – Jonathon Faust Feb 17 '10 at 19:41

add a comment

asked 7 years, 11 months ago

viewed 37,741 times

active 7 months ago



dots with lines pointing to other dots

25



share edit

answered Dec 14 '12 at 11:46



smartcaveman  
24.1k • 17 • 94 • 184

1 +1 FOR A FUNNY ANSWER! :) – Zubair Dec 14 '12 at 12:03

11 This is one of the best answers because it is a simple way of describing what is a simple concept buried in complex terminology (if we're asking this question, we might not know graph theory... or even need to know). My variant would be something like "bar-hopping where you can never go to the same bar twice". Although the family-tree example from another answer is probably conceptually simpler, especially for those of us who aren't college students or alcoholics. – Tom Harrison Jr Jul 23 '16 at 16:48

4 ... in one direction – Mark Robson Jul 4 '17 at 12:34

## FEATURED ON META

-  Affiliate Ads Are Coming To The Network
-  Ask a question template v1 experiment results

## HOT META POSTS

-  customer location
-  Possibility to grow your project into a start-up
-  Global new hire training in Mallorca or Switzerland
-  Macbook Pro or high-end Dell laptop (OS of choice)  
+ 2 more benefits

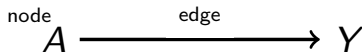
Learn more

Linked

16 Compute gradient norm of each part of

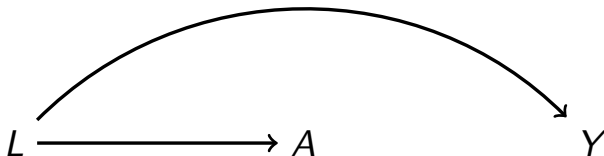
# Directed Acyclic Graphs (DAGs)

- ▶ DAGs are diagrams with simple rules that allow us to clearly think about our causal questions to determine what sources of bias we might have and how we can avoid them
- ▶ DAGs are composed of two elements: nodes and edges
- ▶ Edges are directed (i.e. one causes the other)
- ▶ DAGs must be acyclic, no feedback loops are allowed
- ▶ The edges are not deterministic



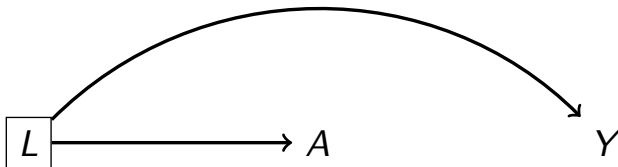
# Directed Acyclic Graphs (DAGs)

- ▶  $Y \not\perp\!\!\!\perp A$  and  $Y^a \perp\!\!\!\perp A$



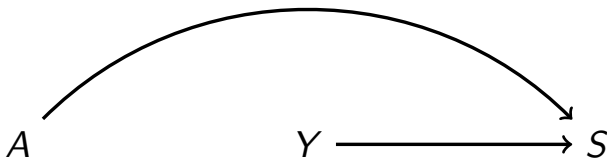
# Directed Acyclic Graphs (DAGs)

- ▶  $Y \perp\!\!\!\perp A|L$  and  $Y^a \perp\!\!\!\perp A|L$



# Directed Acyclic Graphs (DAGs)

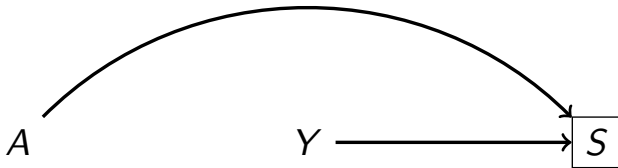
- ▶  $Y \perp\!\!\!\perp A$  and  $Y^a \perp\!\!\!\perp A$





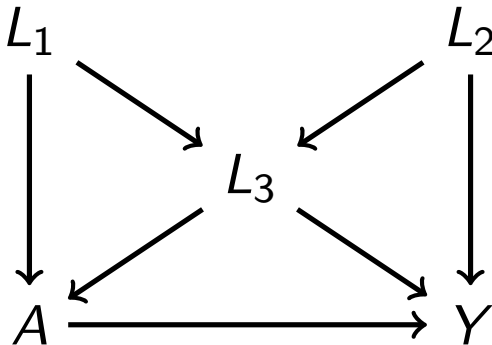
# Directed Acyclic Graphs (DAGs)

- ▶  $Y \not\perp\!\!\!\perp A|S$  and  $Y^a \perp\!\!\!\perp A|S$



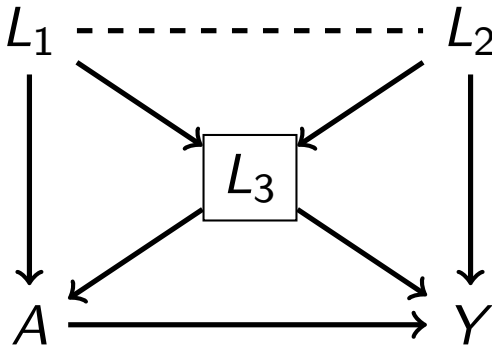
# Directed Acyclic Graphs (DAGs)

- ▶ With a DAG you can identify whether it's possible to achieve conditional exchangeability
- ▶ Here adjusting for  $L_1, L_3$  or  $L_2, L_3$  or  $L_1, L_2, L_3$  are all sufficient adjustment sets



# Directed Acyclic Graphs (DAGs)

- ▶ With a DAG you can identify whether it's possible to achieve conditional exchangeability
- ▶ Here adjusting for  $L_1, L_3$  or  $L_2, L_3$  or  $L_1, L_2, L_3$  are all sufficient adjustment sets



# Directed acyclic graphs (DAGs)



george davey smith

@mendel\_random

Following



Epidemiologists rejoice! "the task of selecting an appropriate set of covariates to control for confounding has been reduced to a simple "roadblocks" puzzle manageable by a simple algorithm" [arxiv.org/abs/1801.04016](https://arxiv.org/abs/1801.04016) and



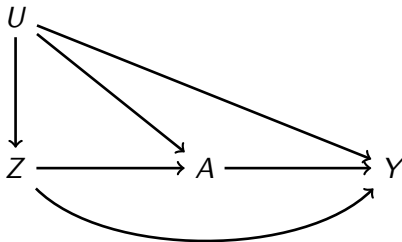
Comments on: The tale wagged by the DAG | International ...

I am grateful to the editors for the opportunity to comment on Nancy Krieger and George Davey Smith's article, 'The tale wagged by the DAG', which appeared in [academic.oup.com](https://academic.oup.com)

This is true! But the DAG has to be correctly specified which is the hard (impossible) part.

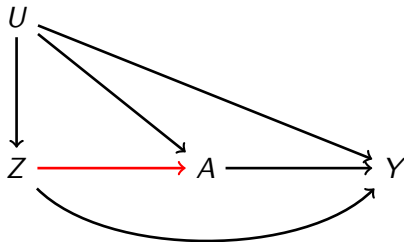
# Showing the IV assumptions in a DAG

- This DAG makes the least amount of assumptions about the observed data. Every node is connected to every other



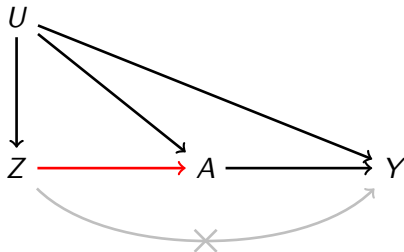
# Showing the IV assumptions in a DAG

- The relevance assumption is represented by the red arrow, showing that there is a relationship between  $Z$  and  $A$



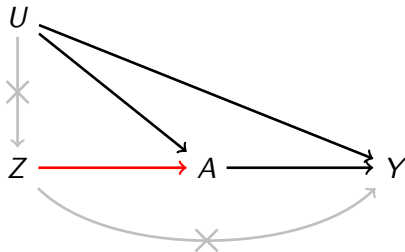
# Showing the IV assumptions in a DAG

- The exclusion restriction (ER) assumption assumes that the direct edge between  $Z$  and  $A$  is not present



# Showing the IV assumptions in a DAG

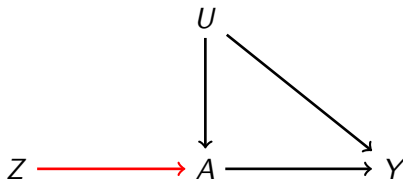
- The exchangeability assumption assumes that  $Z$  is not confounded with  $Y$





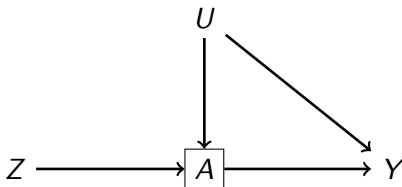
# Showing the IV assumptions in a DAG

- ▶ With our three assumptions, we end up with our classic IV DAG
- ▶ This DAG implies that if  $Z \not\perp\!\!\!\perp Y$  then it must also be that  $A \not\perp\!\!\!\perp Y^a$
- ▶ IV estimation requires monotonicity and homogeneity cannot be drawn on a DAG because they are parametric concepts



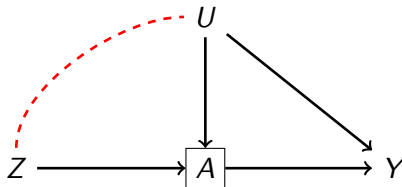
# Conditioning on exposure in IV analyses

- ▶ Sometimes people like to stratify or restrict on exposure
- ▶ But if we do, now  $Z \not\perp U$  and therefore  $Z \not\perp Y$ , AKA, the exchangeability assumption is violated
- ▶ Swanson et al, Am J Epi, 2015



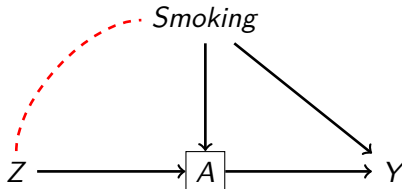
# Conditioning on exposure in IV analyses

- We often draw this on DAGs using a dashed line to remind us that these variables are associated now



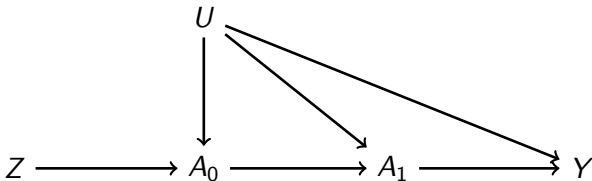
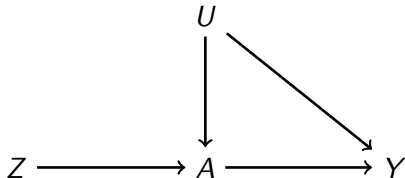
## An example where a simple DAG could have helped

- ▶ “Strong evidence of collider bias was observed for smoking behaviour” (Cho et al, Sci Rep, 2015)
- ▶ “To minimize the effect of risk factors susceptible to collider bias, associations between the SNP and cardiovascular outcomes were then assessed with adjustments for smoking”
- ▶ By saying smoking is what drives collider bias, they are implying smoking is the only confounder alcohol and cardiovascular outcomes. Better off simply adjusting then?



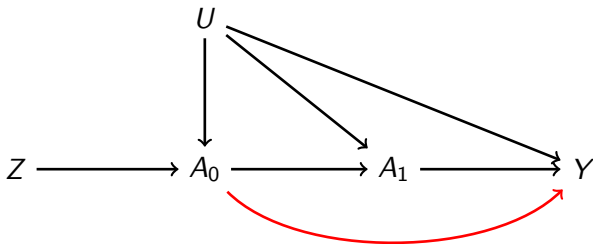
## Using DAGs to think about IV longitudinally

- ▶ We can depict a time-varying exposure by splitting the  $A$  node into separate nodes for each time point
- ▶ We are still assuming the IVs assumptions but splitting  $A$  into two nodes adds assumptions to the DAG



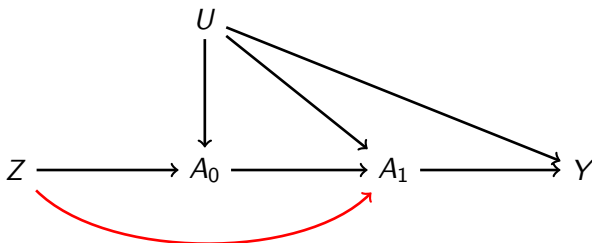
## Using DAGs to think about IV longitudinally

- ▶ We have assumed that  $A_0$  does not have a direct effect on  $Y$
- ▶ If  $A_0$  does affect  $Y$  then we can no longer estimate the effect of  $A_1$  on  $Y$
- ▶ This is because, although the ER holds for  $A$  as a whole, it does not hold for  $A_1$
- ▶ So if we measure  $A_0$  we're alright?



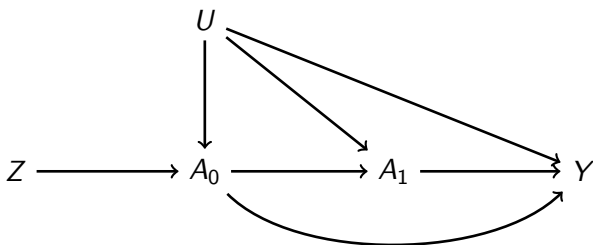
## Using DAGs to think about IV longitudinally

- ▶ We have also assumed that  $Z$  does not have a direct effect on  $A_1$
- ▶ If  $Z$  does affect  $A_1$  then we can no longer estimate the effect of  $A_0$  on  $Y$
- ▶ This is because, although the ER still holds for  $A$  as a whole, it does not hold for  $A_0$
- ▶ All this to illustrate that, in contexts such as Mendelian randomization, important thought must be given to how exposures change longitudinally



## Using DAGs to think about IV longitudinally

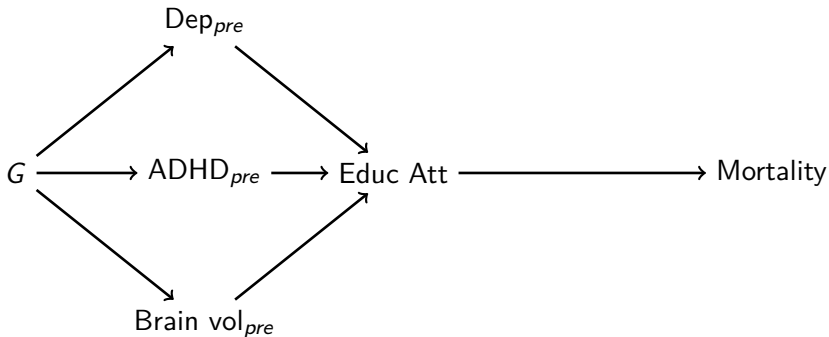
- ▶ What if we think about joint effects?
- ▶ First, though “lifetime effects” is often used in Mendelian randomization, it has no precise definition
- ▶ We propose  $E[Y_k^{\bar{A}+1} - Y_k^{\bar{A}}]$
- ▶ Only identifiable with IV if the effect of the  $Z$  on  $A$  does not change with time (N.B.:  $A$  can change as long as the effect of  $G$  doesn't)
- ▶ This also has implications for which null hypotheses are testable (Swanson et al, Eur J Epi, 2018)





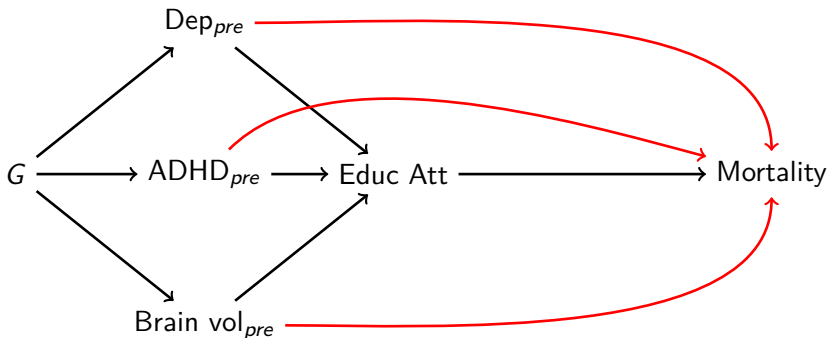
# Using DAGs to think about Mendelian randomization

- DAG for a MR study of the effect of educational attainment on mortality



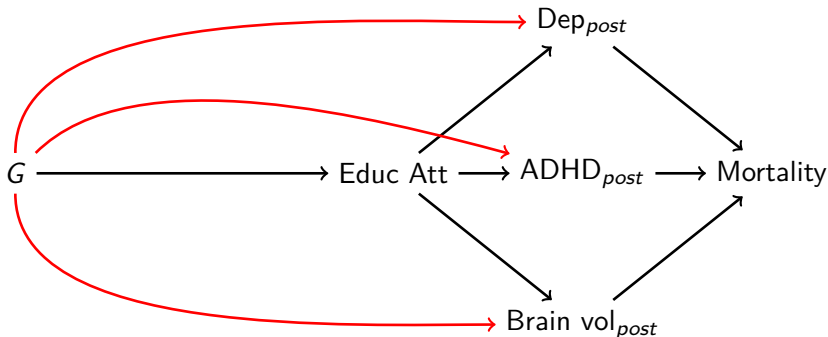
# Using DAGs to think about Mendelian randomization

- ▶ But what if the intermediates between  $G$  and education attainment have a direct effect on mortality?



# Using DAGs to think about Mendelian randomization

- Or what if  $G$  has a direct affect on the intermediates after educational attainment



# Using DAGs to think about Mendelian randomization

- ▶ We have also assumed that  $Z$  does not have a direct effect on  $A_1$

