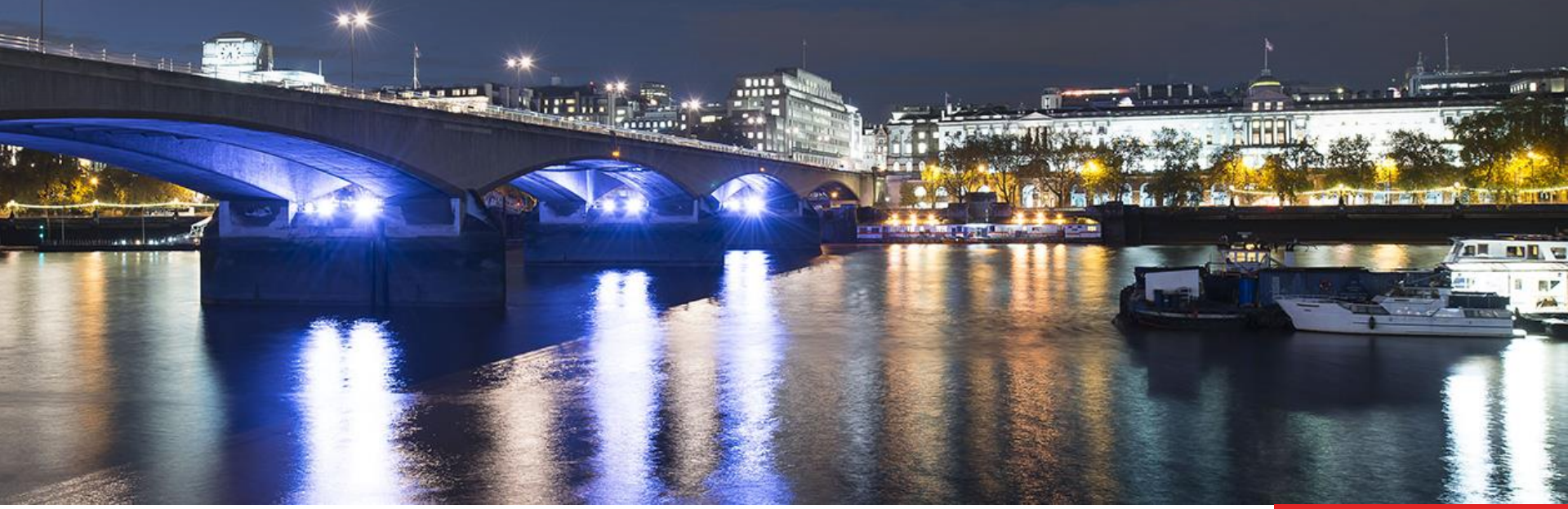
Careful with Causal Inference

Richard Emsley
Professor of Medical Statistics and Trials Methodology
Department of Biostatistics & Health Informatics
Institute of Psychiatry, Psychology & Neuroscience

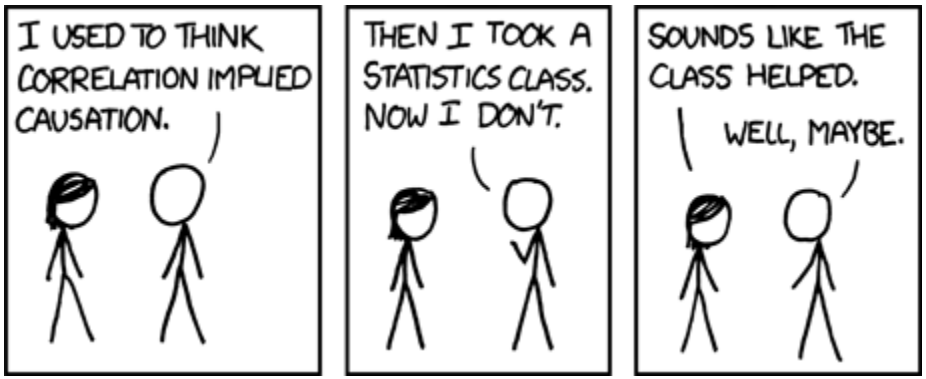


KING'S HEALTH PARTNERS

NHS National Institute for Health Research
Biomedical Research Centre for Mental Health

KING'S College LONDON

Correlation and causation



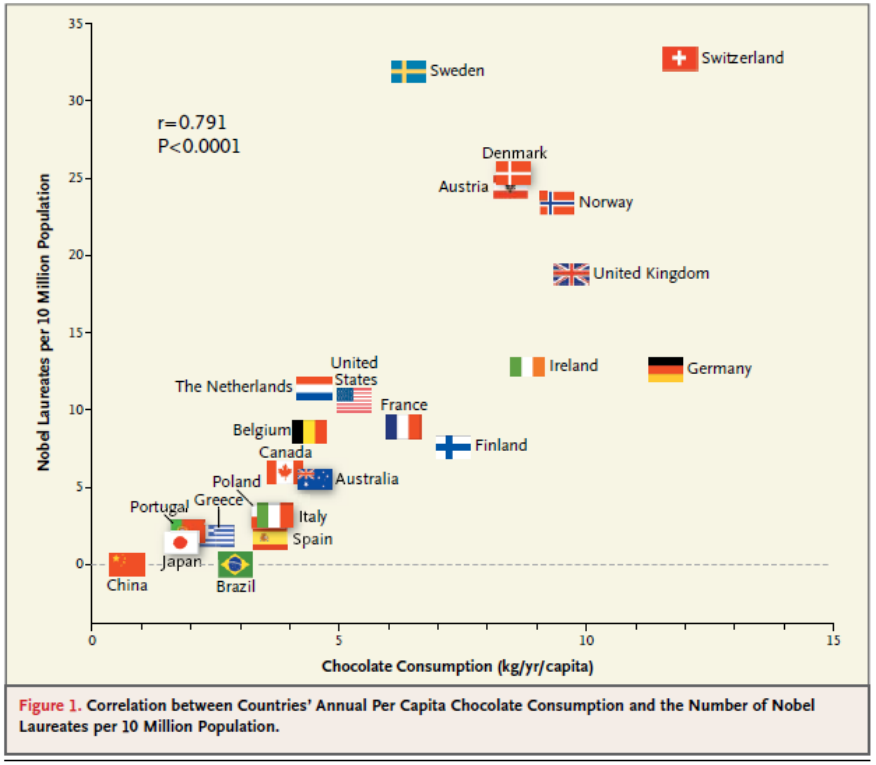


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Messerli F (2012), NEJM, 367:16

"Some scientists are reluctant to speak so blatantly about cause and effect, but in statements of hypothesis and in describing study objectives such boldness serves to keep the real goal firmly in focus and is therefore highly preferable to insipid statements about 'association' instead of 'causation'

Rothman (1986), Modern Epidemiology

Overview

1. What is causal inference?

2. How does causal inference help?
   - Better causal questions
   - Better confounding control

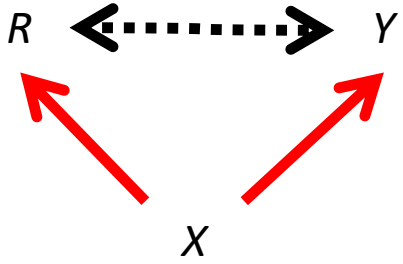
3. How can we be more careful with causal inference?
   - Or should we be more explicit?

If there is a significant correlation between two variables *R* and *Y*, then either:

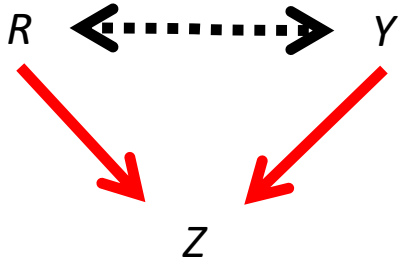
1. *R* causes *Y*

2. *Y* causes *R*

3. *R* and *Y* share a common cause *X*



4. *R* and *Y* are conditioned on a common descendent *Z*



The general principle of causal inference

- Statistical models can only tell us about association between two variables (say *R* and *Y*)

- The aim of causal inference is to infer whether this association can be given a causal interpretation (e.g. *R* causes *Y*) by:
  - defining the causal estimands
  - being explicit about the assumptions being made
  - thinking about other possible explanations for observed effects, especially confounding.

- There are now many, many methods purporting to give causally valid solutions to this problem; this session only gives an overview of some of these

A brief history of causal inference (1)

- Neyman (1923) and Fisher (1925) discussed the **potential yield** to be gained from agricultural plots under different experimental exposures.

- First introduction of the concept of random allocation as an experimental design.

Ronald Fisher
(1890-1962)

Jerzy Neyman
(1894-1981)

A brief history of causal inference (2)

- This was formalised statistically for both randomised and non-randomised studies many years later.
  - Potential outcomes
  - Rubin Causal Model (Holland 1986)

- Rubin DB (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5), 688-701.

- Rosenbaum PR and Rubin DB (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), pp41-55.

Don Rubin

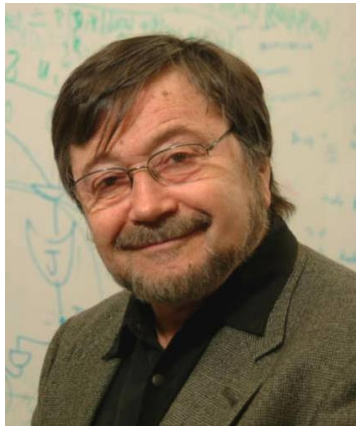
A brief history of causal inference (3)

- Extended the potential outcomes framework to longitudinal setting (repeated measures).
- This required a new methodology for estimating parameters using semi-parametric theory: the “G-family”
- Uses terminology ‘counterfactuals’ rather than potential outcomes.

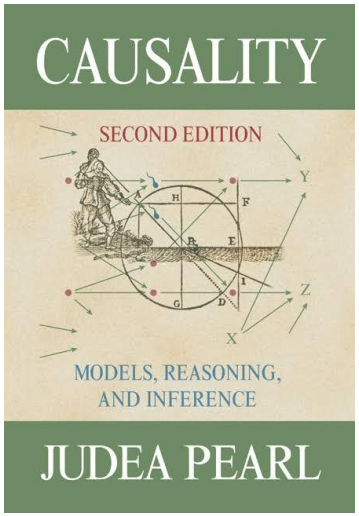
Jamie Robins

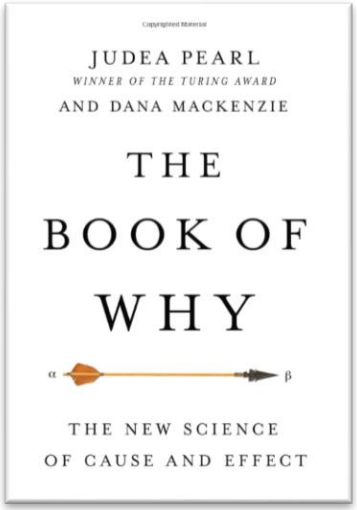
A brief history of causal inference (4)

- Developed a theory of causal and counterfactual inference based on graphical models and probabilistic reasoning.
- Derived a new method for determining relations between variables, known as ‘do-calculus’.
- Explores the link between counterfactuals and non-parametric structural equation models.



Judea Pearl





A brief history of causal inference (5)

- There is a group who argue against using the counterfactuals or potential outcomes framework.
- Dawid and colleagues propose for methods for causal inference without counterfactuals, mainly using decision theory, graphical models and stochastic modelling.

L-R: Carlo Berzuini, Phil Dawid, Vanessa Didelez

Is the terminology important?

"Personally I see the different formalisms as different 'languages'. The French language may be best for making love whereas the Italian may be suitable for singing, but both are indeed possible…"

Lauritzen: Scandinavian Journal of Statistics 2004 Vol. 31 p189

Fundamental concept of causal inference

Receive treatment

Measure outcome

Receive control

Measure outcome

Comparison of outcomes gives an
individual treatment effect

Fundamental concept of causal inference

Receive treatment

Measure outcome

Receive control

Measure outcome

Comparison of these outcomes **will not give** an individual treatment effect

Fundamental concept of causal inference

Receive treatment

Measure outcome

Receive control

Measure outcome

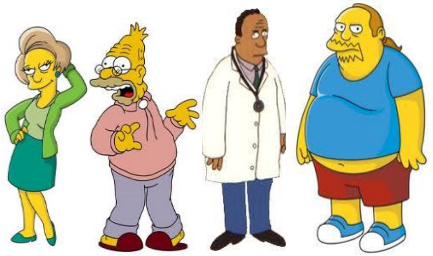
Comparison of average outcomes defines the
average treatment effect

Fundamental concept of causal inference

Receive treatment

Measure outcome

Receive control



Measure outcome

Comparison of average outcomes estimates the

average treatment effect

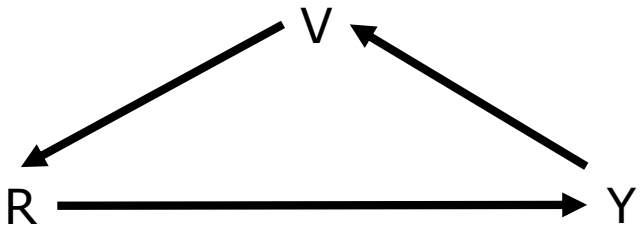
Directed Acyclic Graphs

- A causal effect of one variable on another (R: treatment, Y: outcome) is shown as:

- If *Y* is continuous variable then we could estimate the effect using linear regression:

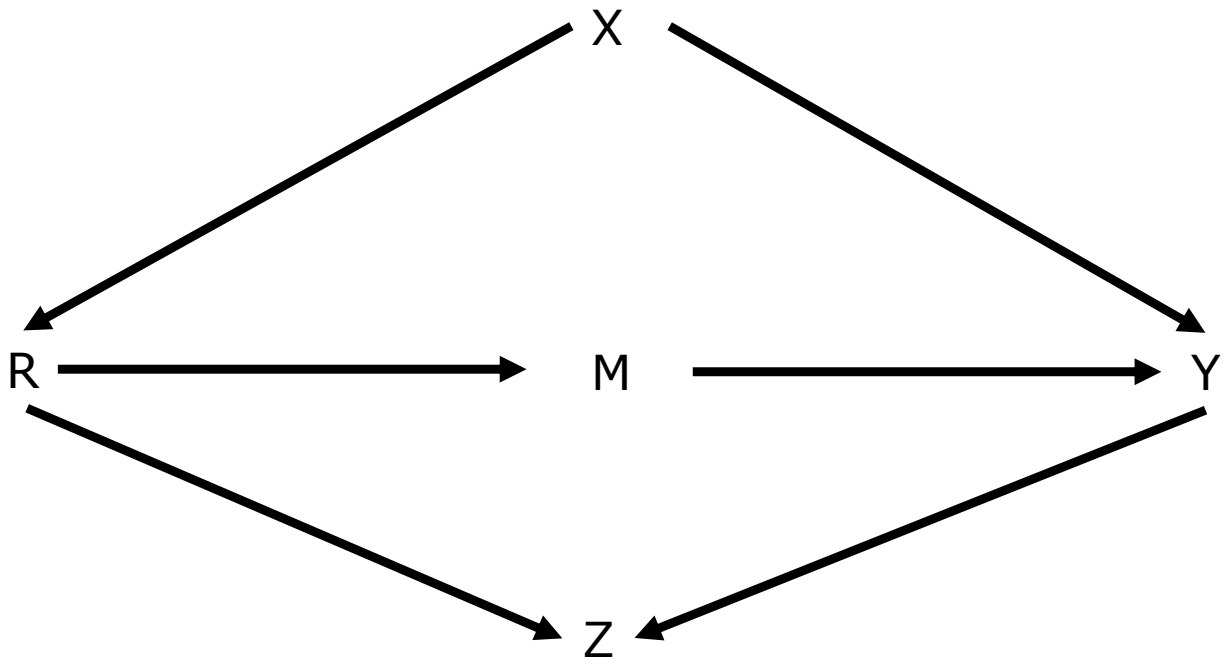
$$y_i = \beta_0 + \beta_1 r_i + \varepsilon_i,\ \varepsilon_i \sim N(0, \sigma^2)$$

- A DAG must not be cyclic, i.e. starting at treatment we should be able to get back to it



Types of variable in DAGs

- Confounding (*X*)
- Mediation (*M*)
  - Partitions effects into direct and indirect effects
- Colliders (*Z*) – biases association between their parents



Overview

1. What is causal inference?

2. How does causal inference help?
   - Better causal questions
   - Better confounding control

3. How can we be more careful with causal inference?
   - Or should we be more explicit?

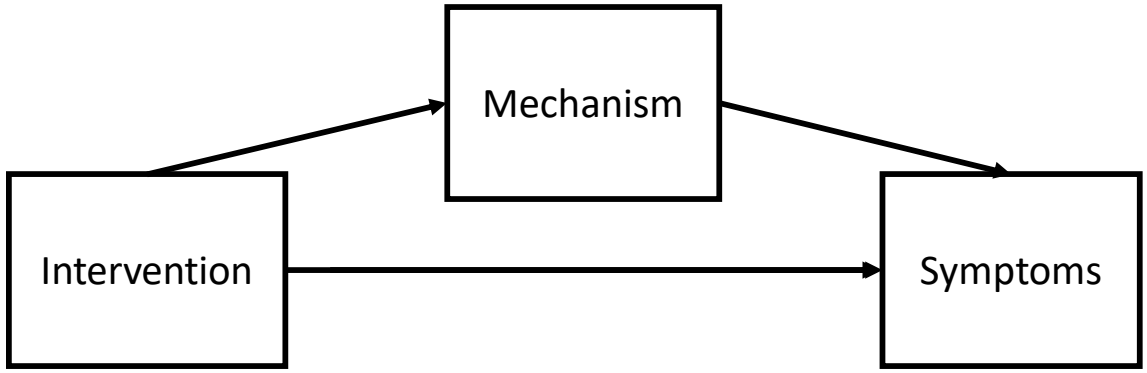
When can ATE be estimated?

- We wish to evaluate the effects of receiving a treatment compared to a suitably defined control condition
- Thus we want to use a sample of subjects from a relevant target population to compare outcomes between a treated group and a control group
- When can we do this without running into problems?
  - Randomised controlled trial (RCT): Participants are randomised to two arms (experimental treatment and control)
    - YES - provided participants adhere to their allocated treatment.
  - Observational study: Compares subjects receiving the experimental treatment with subjects under the control condition
    - SOMETIMES - only if variables that drive treatment group selection have been measured and accounted for appropriately

What are we estimating in trials?

- Interested in various measures of effect
  - Effectiveness - the benefit of a treatment policy
  - Efficacy - the benefit of actually receiving treatment
- ITT measures effectiveness as **implemented in a given trial**
- **What is the effectiveness of offering the intervention?**
- It tells us whether randomising the treatment works
  - **On average, not for an individual patient!**
  - **Regardless of whether you receive the treatment or not!**

Target mechanisms



- Target intermediate variables:
  - Some treatments target a particular intermediate variable in order to bring about change in a clinical outcome.
    - Motivational interviewing → substance use → symptoms
    - Cognitive behaviour therapy → thinking → symptoms
    - Beta blockers → blood pressure → stroke risk
    - Sleep intervention → sleep → cognition

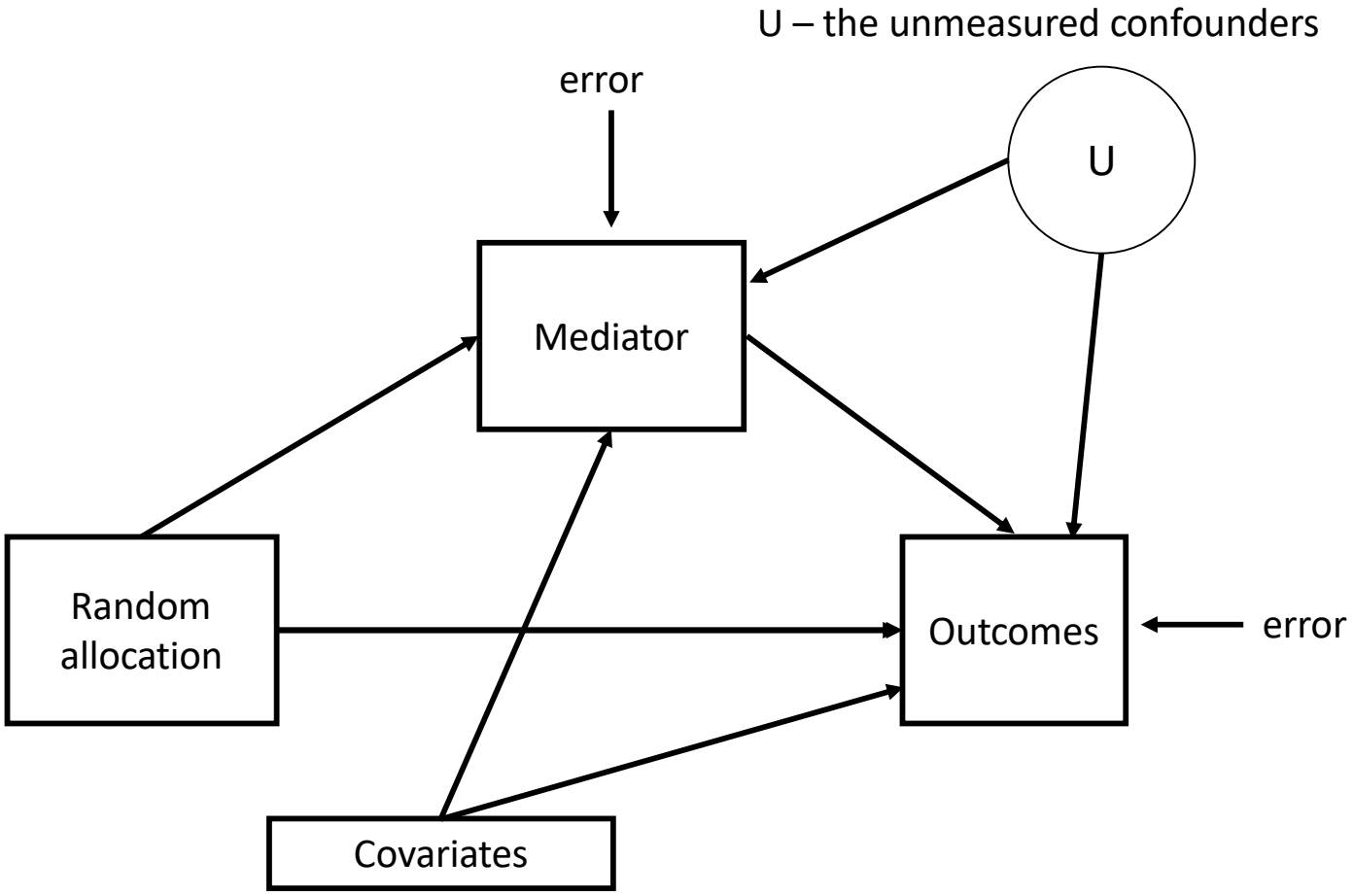
- An explanatory analysis of a trial would seek to establish that this is indeed the case using **mediation analysis**; i.e. assess the mediated path.

Mediation analysis and causal inference…

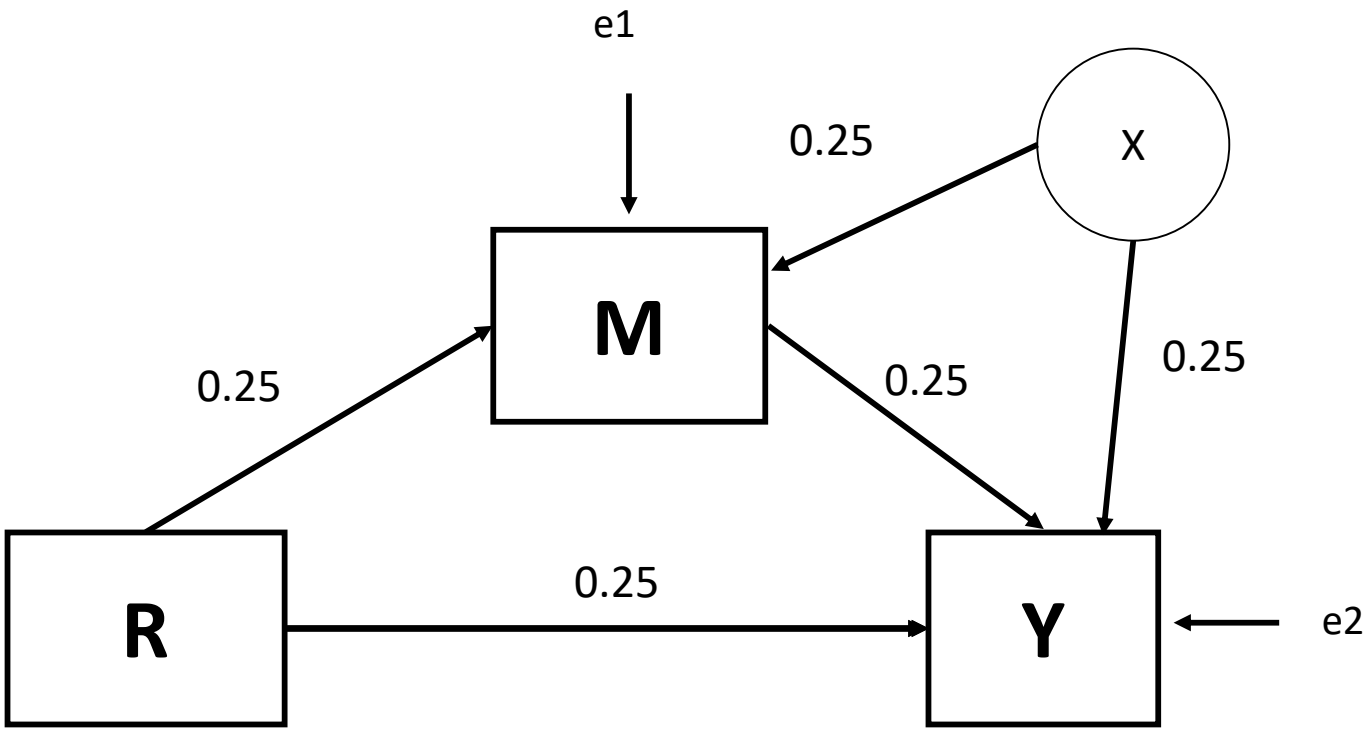
"Mediation analysis is a form of causal analysis…all too often persons conducting mediational analysis either do not realize that they are conducting causal analyses or they fail to justify the assumptions that they have made in their casual model."

David Kenny (2008), Reflections on Mediation, *Organizational Research Methods.*

The basic underlying problem: estimating valid causal effects



Example: simple mediation analysis



Total effect	**= direct effect**	+	**indirect effect**
	= 0.25	+	0.25*0.25
	= 0.25	+	0.0625

Example: regression approach stage 1

```
. regress y r
```

Source	SS	df	MS	Number of obs	=	2,000
				F(1, 1998)	=	289.30
Model	49.0074641	1	49.0074641	Prob > F	=	0.0000
Residual	338.463931	1,998	.169401367	R-squared	=	0.1265
				Adj R-squared	=	0.1260
Total	387.471395	1,999	.193832614	Root MSE	=	.41158

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
r	.3130773	.0184068	17.01	0.000	.2769787	.3491759
_cons	.3082655	.012983	23.74	0.000	.2828039	.3337272

Example: regression approach stage 2

```
. regress m r
```

Source	SS	df	MS	Number of obs	=	2,000
				F(1, 1998)	=	260.01
Model	31.8440108	1	31.8440108	Prob > F	=	0.0000
Residual	244.698233	1,998	.122471588	R-squared	=	0.1152
				Adj R-squared	=	0.1147
Total	276.542244	1,999	.138340292	Root MSE	=	.34996

m	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
r	.252368	.0156509	16.12	0.000	.2216743	.2830617
_cons	.2383794	.0110391	21.59	0.000	.21673	.2600288

Example: regression approach stage 3

```
. regress y m r
```

Source	SS	df	MS		
				Number of obs	= 2,000
				F(2, 1997)	= 1002.47
Model	194.120376	2	97.0601879	Prob > F	= 0.0000
Residual	193.351019	1,997	.096820741	R-squared	= 0.5010
				Adj R-squared	= 0.5005
Total	387.471395	1,999	.193832614	Root MSE	= .31116

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m	.7700831	.0198916	38.71	0.000	.7310728	.8090935
r	.1187329	.0147935	8.03	0.000	.0897207	.1477452
_cons	.1246936	.0109006	11.44	0.000	.1033158	.1460713

Example: regression approach stage 3

```
. regress y m r x
```

Source	SS	df	MS		
Model	260.153288	3	86.7177628	Number of obs	= 2,000
Residual	127.318106	1,996	.063786626	F(3, 1996)	= 1359.50
				Prob > F	= 0.0000
Total	387.471395	1,999	.193832614	R-squared	= 0.6714
				Adj R-squared	= 0.6709
				Root MSE	= .25256

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m	.241628	.0230313	10.49	0.000	.1964601	.2867958
r	.2540715	.0127229	19.97	0.000	.2291199	.279023
x	.2605752	.0080987	32.17	0.000	.2446924	.2764581
_cons	.2510649	.0096803	25.94	0.000	.2320804	.2700495

Example: with measurement error

```
. regress y m_star r x
```

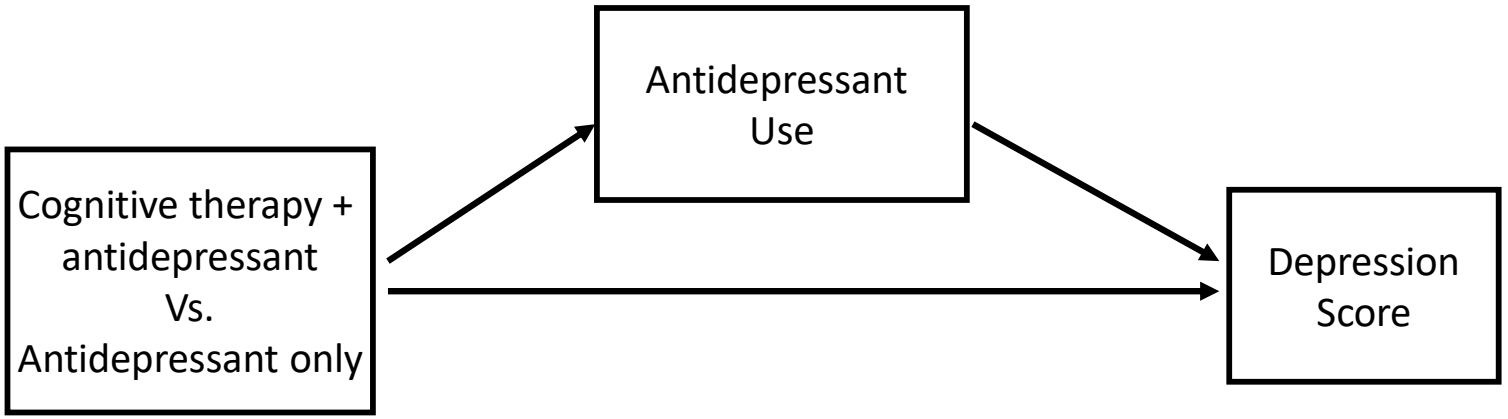
Source	SS	df	MS		
Model	254.780967	3	84.9269891	Prob > F	0.0000
Residual	132.690428	1,996	.06647817	R-squared	0.6575
				Adj R-squared	0.6570
Total	387.471395	1,999	.193832614	Root MSE	.25783

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m_star	.052463	.0105353	4.98	0.000	.0318016	.0731244
r	.3014506	.0118715	25.39	0.000	.2781688	.3247324
x	.3077627	.0063906	48.16	0.000	.2952298	.3202957
_cons	.2973219	.0084511	35.18	0.000	.280748	.3138957

“Nuisance” mediators

- Variables measured post-randomisation that we may wish to rule out having a mediated effect - essentially we want to estimate the residual direct effects and find a small indirect effect

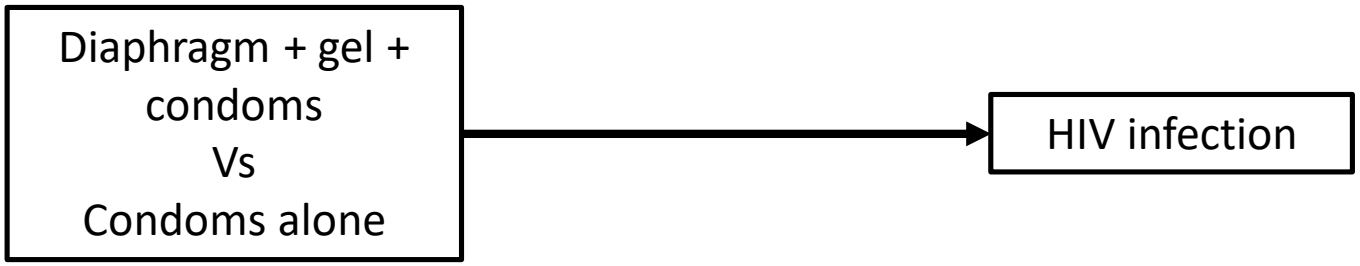
- Use of concomitant medication or interventions in treatment as usual



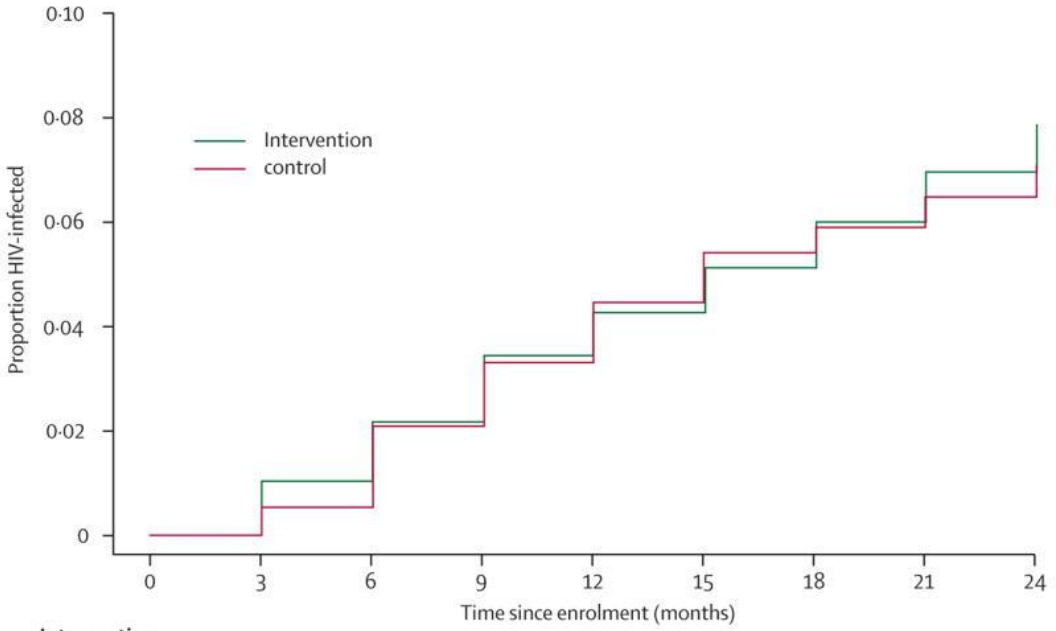
Example: HIV and condom use

- Diaphragm and lubricant gel for prevention of HIV acquisition in southern African women: a randomised controlled trial

- Padian et al. (2007) Lancet 370(9583):251-61.



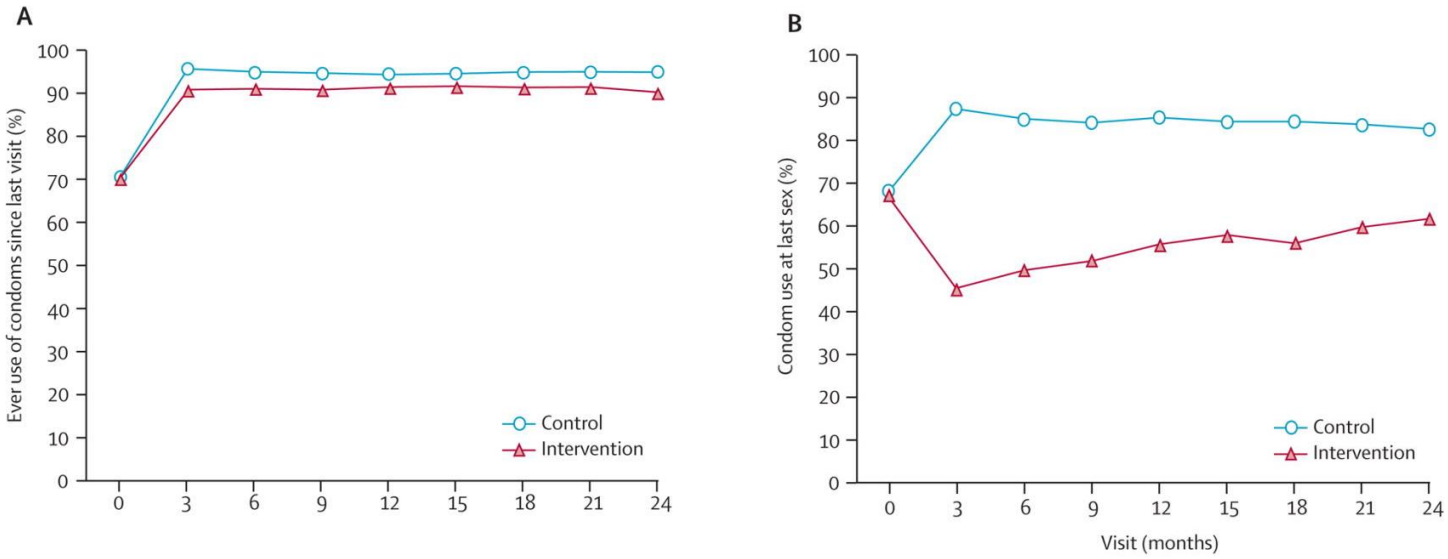
Example: ITT effect on HIV incidence



Intervention								
Number at risk	2472	2427	2381	2314	2000	1606	1234	906
Events	25	28	31	20	18	15	12	9
Control								
Number at risk	2476	2442	2385	2344	2011	1634	1248	928
Events	13	38	30	28	19	9	8	6

- Overall HIV incidence was 4.0% per 100 woman-years: 4.1% in the intervention group (n=2472) and 3.9% in the control group (n=2476), corresponding to a relative hazard of 1.05 (95% CI 0.84-1.32, intention-to-treat analysis)
- Padian et al (2007)

Example: ITT effect on condom use



- The proportion of women using condoms was significantly lower in the intervention than in the control group
  (54% vs 85% of visits, $p<0.0001$).

- Padian et al (2007)

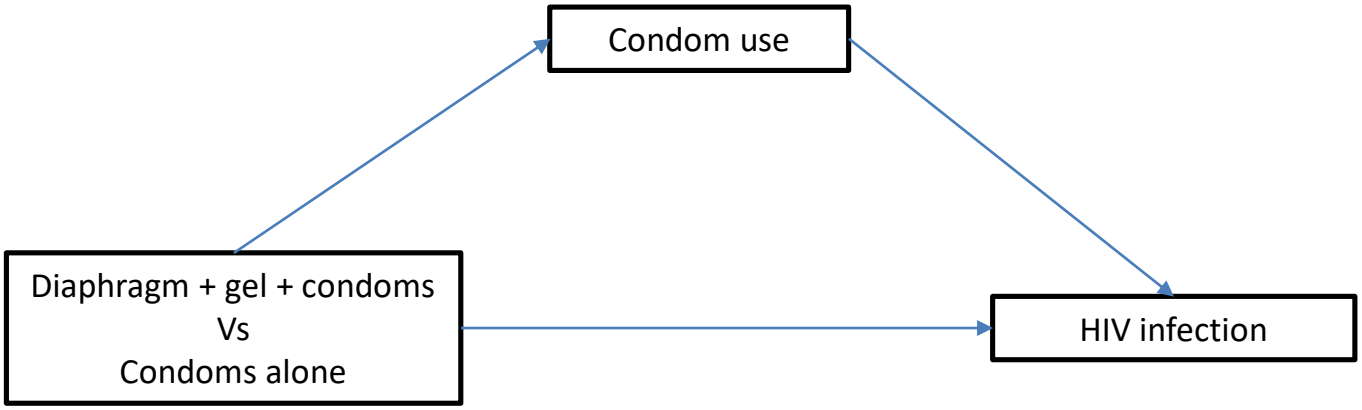
Example: mediation question

- "Shelton and Stein also ask for a disentanglement of the separate effects of condom and diaphragm use on incidence of HIV infection in the trial, part of which necessarily involves an estimation of the independent effect of condom use. We agree that such analyses are important additions to basic intention-to-treat results, and in fact, we prespecified appropriate methods to address these issues in our analytical plan and included results in the submitted paper. Regrettably, we were directed not to include these findings by both a referee and an editor of the original article."

- Jewell et al (2007), The Lancet, 370(9602):1823-1824

Example: mediation hypothesis

- A case for controlled direct effects…
- What is the direct effect of randomisation to diaphragm use on HIV infection if **everyone in the population** used condoms?
- What is the direct effect of randomisation to diaphragm use on HIV infection if **no-one in the population** used condoms?



Example: mediation analysis

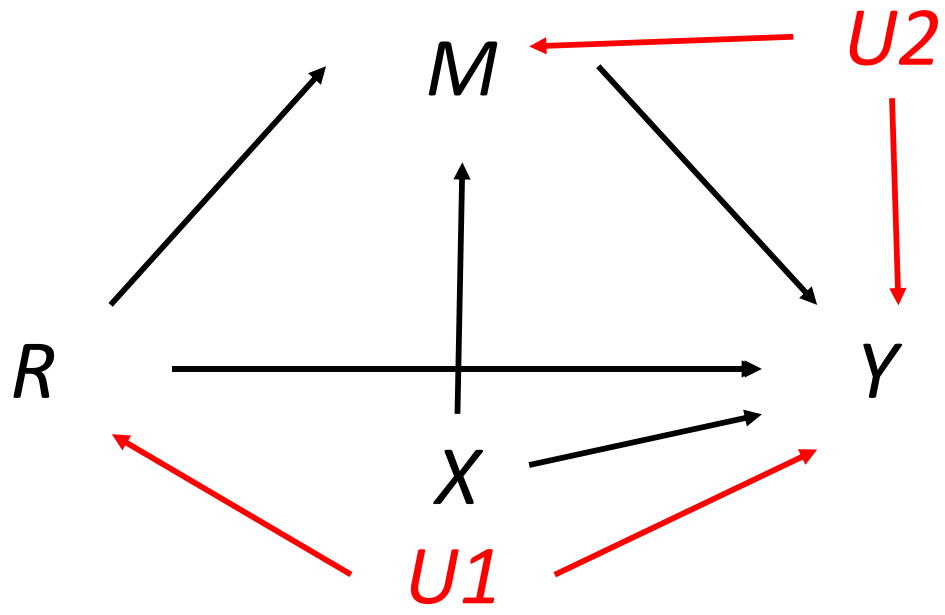
- The estimated relative risk of HIV infection for assignment to the intervention versus control group, had **all** participants been constrained to **always** use condoms, was 0.96 (95% CI 0.59–1.45).

- By contrast, the estimated analogous relative risk of HIV infection had **all** participants **never** used condoms was 0.59 (0.26–4.56).

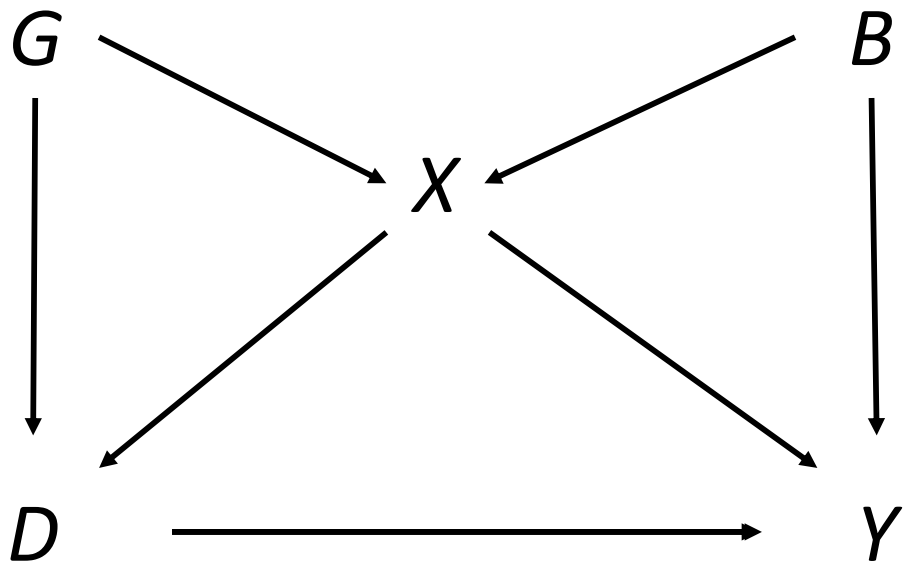
- Jewell et al (2007), The Lancet, 370(9602):1823-1824

Assumptions for identification

- Controlled direct effects require:
  - A1: no unmeasured *R*-*Y* confounding (*U1*);
  - A2: no unmeasured *M*-*Y* confounding (*U2*).

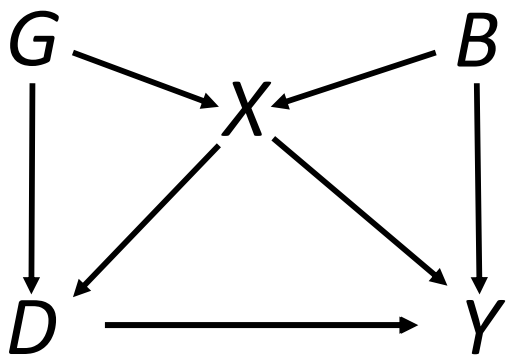


M-bias graph



M-bias graph

- What are all the paths from *D* to *Y*?



- What is the consequence of adjusting for *X* in estimating the effect of *D* on *Y*?

- What variables do we need to adjust for to estimate the effect of *D* on *Y*?

A realistically complex DAG

- To estimate the effect of obesity on PE (Pearce and Lawlor, IJE, 2017)

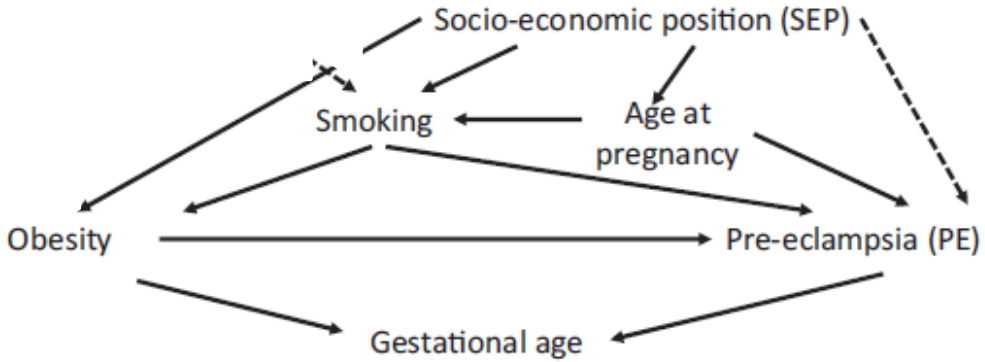


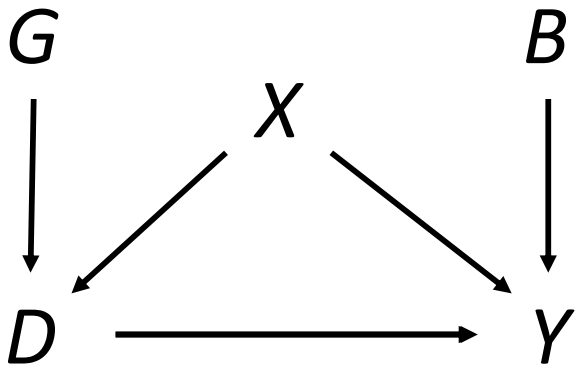
Figure 1. Illustrative example–directed acyclic graph for the hypothesis that obesity is causally related to pre-eclampsia

- Unblocked paths:
  - Obesity – Smoking – SEP - Age at pregnancy - PE
  - Obesity – SEP – Age at pregnancy - PE
  - Obesity – Smoking – Age at pregnancy - PE
  - Obesity – Smoking – PE

- Adjusting for age at pregnancy and Smoking is sufficient

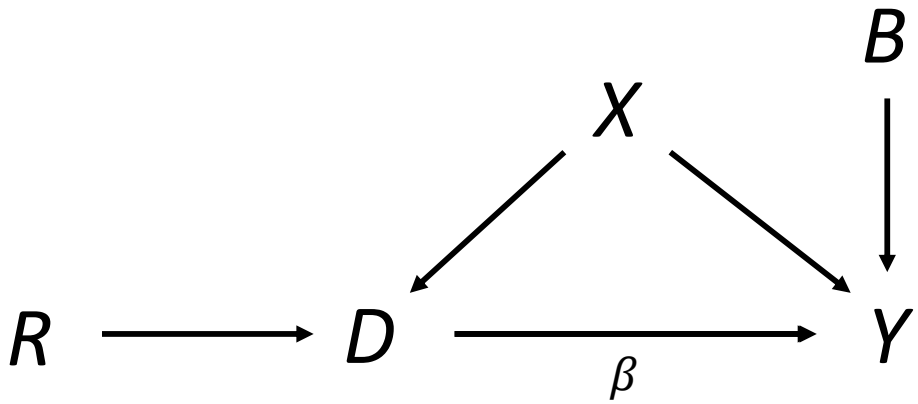
Confounding

- Confounding is a causal concept
  - Confounding of which effect?
  - Cannot causally interpret the parameters of confounders



- Association does not require any confounding adjustment
  - simply compute it from the data
- Prognostic models do not require confounding adjustment
  - Include *X* and *B*
- Propensity score models should not include predictors of exposure (variable *G*)

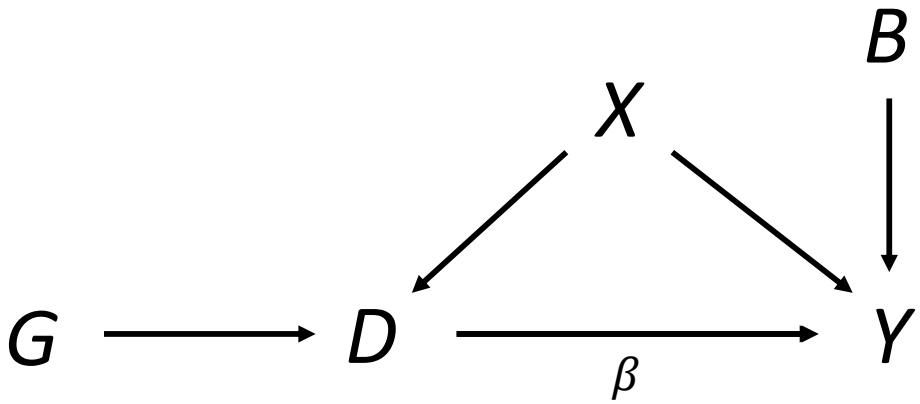
Non-compliance in trials



- *R* = randomisation
- *D* = treatment received
- *Y* = outcome
- *X* = confounder
- *B* = prognostic variable

$$\beta = \frac{\text{Effect of } R \text{ on } Y}{\text{Effect of } R \text{ on } D}$$

Mendelian randomisation



- *G* = gene
- *D* = treatment received
- *Y* = outcome
- *X* = confounder
- *B* = prognostic variable

$$\beta = \frac{\text{Effect of } G \text{ on } Y}{\text{Effect of } G \text{ on } D}$$

Overview

1. What is causal inference?

2. How does causal inference help?
   - Better causal questions
   - Better confounding control

3. How can we be more careful with causal inference?
   - Or should we be more explicit?

Should we be careful with the C-Word?

- Miguel Hernán (2018), American Journal of Public Health

COMMENTARY

The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data

Causal inference is a core task of science. However, authors and editors often refrain from

Miguel A. Hernán, MD, DrPH

You know the story: process and, inevitably, errors are 0.8 comes from a very large

Observational studies and causality

- In observational studies, we want to know about variables that can be modified or manipulated
- Define the causal effect in the population as the causal effect that would have been observed in a hypothetical trial
- Your observational effect estimate may be seriously confounded…but we know that
- The goal is causal…the analysis is associational
- This is true of randomised trials as well!

When can we use the words “causal effect”?

- Title 
- Introduction 
- Methods 
- Results 
- Discussion 

Hayes and Rockwood (2017)

"There are some **hardliners** who say that to claim the existence of cause-effect relationships (and mediation is by definition a cause-effect process), one must engage in experimental manipulation with random assignment, collect data over time or, ideally, both.

Furthermore, one must meet an overwhelming number of assumptions beyond those of linear modeling that go by such names as "sequential ignorability," "stable unit treatment value" and others, many that are quite **technical in nature or hard** or impossible to test.

Others argue that one cannot conduct a mediation analysis with merely correlational data, that moderators must be independent of presumed causes of effects, and the list of requirements goes **on and on**…"

“Hardliners”

“(see e.g., Emsley, Dunn, & White, 2010; Preacher, 2015, for a discussion

Hayes and Rockwood (2017)

"We feel that if these are taken as literal requirements rather than as just ideals or recommendations, most research would not be done because most researchers cannot meet these requirements (due to resource constraints, ethics, and a myriad list of other reasons)."

"We would rather see more imperfect work conducted and published than see research slow to a trickle because investigators don't feel that their work will satisfy all critics and pass every test for valid causal inference."

"**You can do most anything you want with your data. Most any statistical tool can provide some insight into the story you ultimately end up telling with your data."**

Hayes and Rockwood (2017), BRAT, 98:39-57

The Future for causal inference

- Training courses
- Masters level training
- European Causal Inference Meeting
  - Bermen, April 2019
- Journal of Causal Inference
- Lots of forthcoming books
- Machine Learning and artificial intelligence

Research Programme: Efficacy and Mechanisms Evaluation

Joint work and slides prepared with **Graham Dunn**, **Ian White, Andrew Pickles** and **Sabine Landau.**

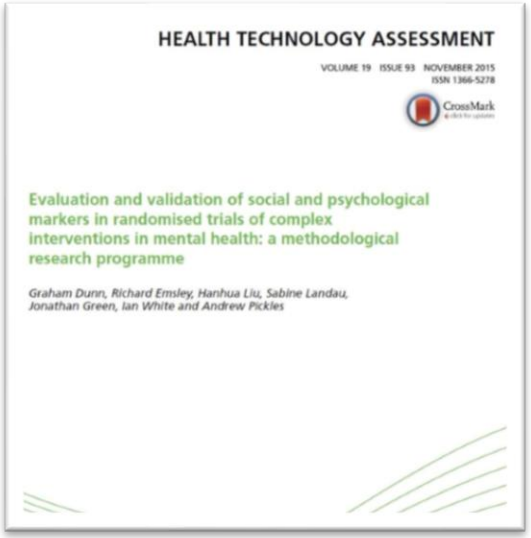
Funded by Medical Research Council Methodology Research Programmes:

- **Design and methods of explanatory (causal) analysis for randomised trials of complex interventions in mental health (2006-2009)**
  - Graham Dunn (PI), Richard Emsley, et al
- **Estimation of causal effects of complex interventions in longitudinal studies with intermediate variables (2009-2012)**
  - Richard Emsley (PI), Graham Dunn.
- **MRC Early Career Centenary Award (2012-13)**
- **Designs and analysis for the evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health (2010-12)**
  - Graham Dunn (PI), Richard Emsley, et al.
- **Developing methods for understanding mechanism in complex interventions (2013-16)**
  - Sabine Landau (PI), Richard Emsley, et al.
- **MRC NorthWest Hub for Trials Methodology Research (2013-2018)**
  - Paula Williamson (PI), Richard Emsley, et al.

Methodology report

- Dunn G, Emsley RA, Liu H, Landau S, Green J, White I and Pickles A. (2015). Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health. Health Technology Assessment 19 (93).

- Non-technical introduction and summary of our work on analysing complex interventions:
  - Introduction to causal inference
  - Mediation analysis
  - Process evaluation
  - Longitudinal extensions
  - Stratified medicine
  - Guidance and tips for trialists



Thank you for your attention

Email: richard.emsley@kcl.ac.uk