

Bias and Confounding

Christian Torp-Pedersen

November 27, 2022

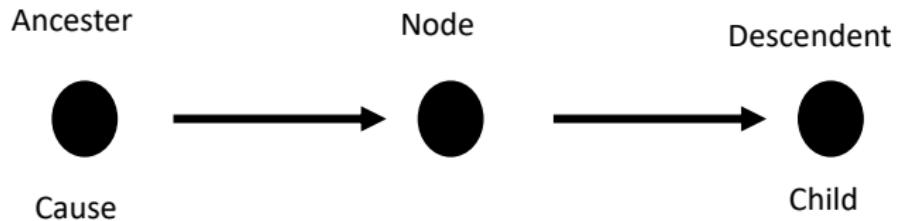


Adjustment for / Conditioning on

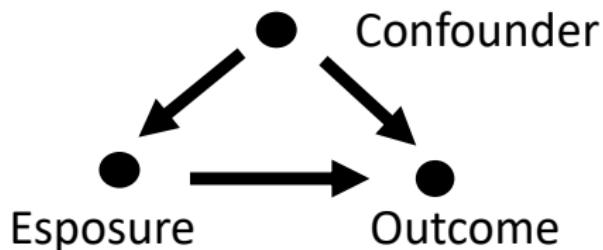
$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \text{sex} + \beta_2 * \text{age} + \beta_3 * \text{CPR} + \beta_4 * \text{time}$$

$$OddsRatio_{CPR} = e^{\beta_3}$$

Directed Acyclic Graphs

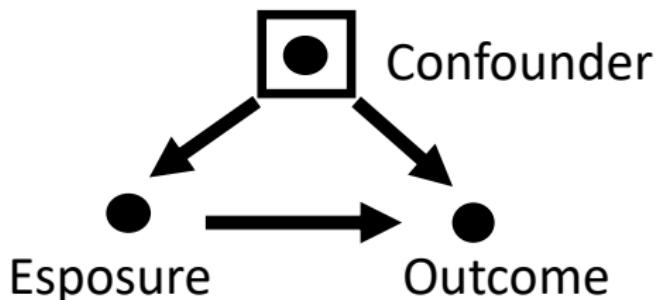


Directed Acyclic Graphs



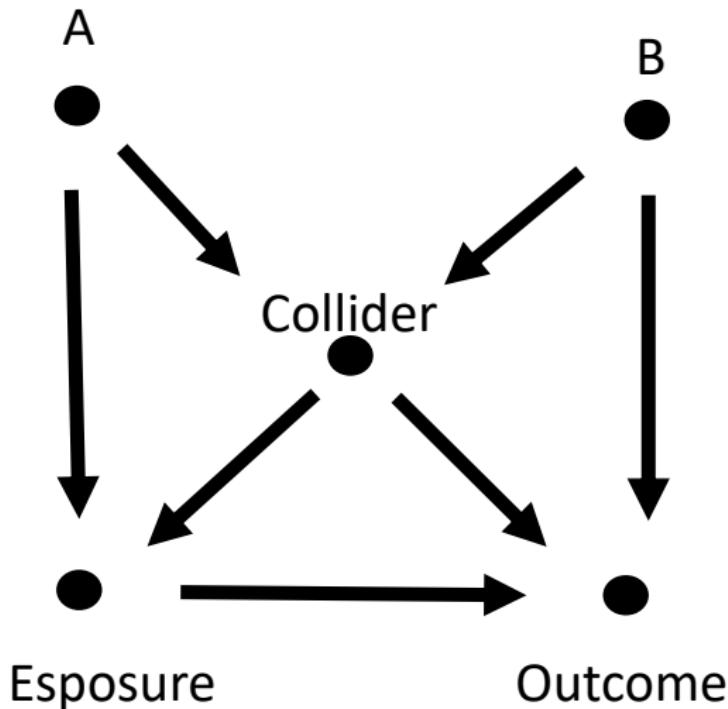
B – Confounding and backdoor pathway

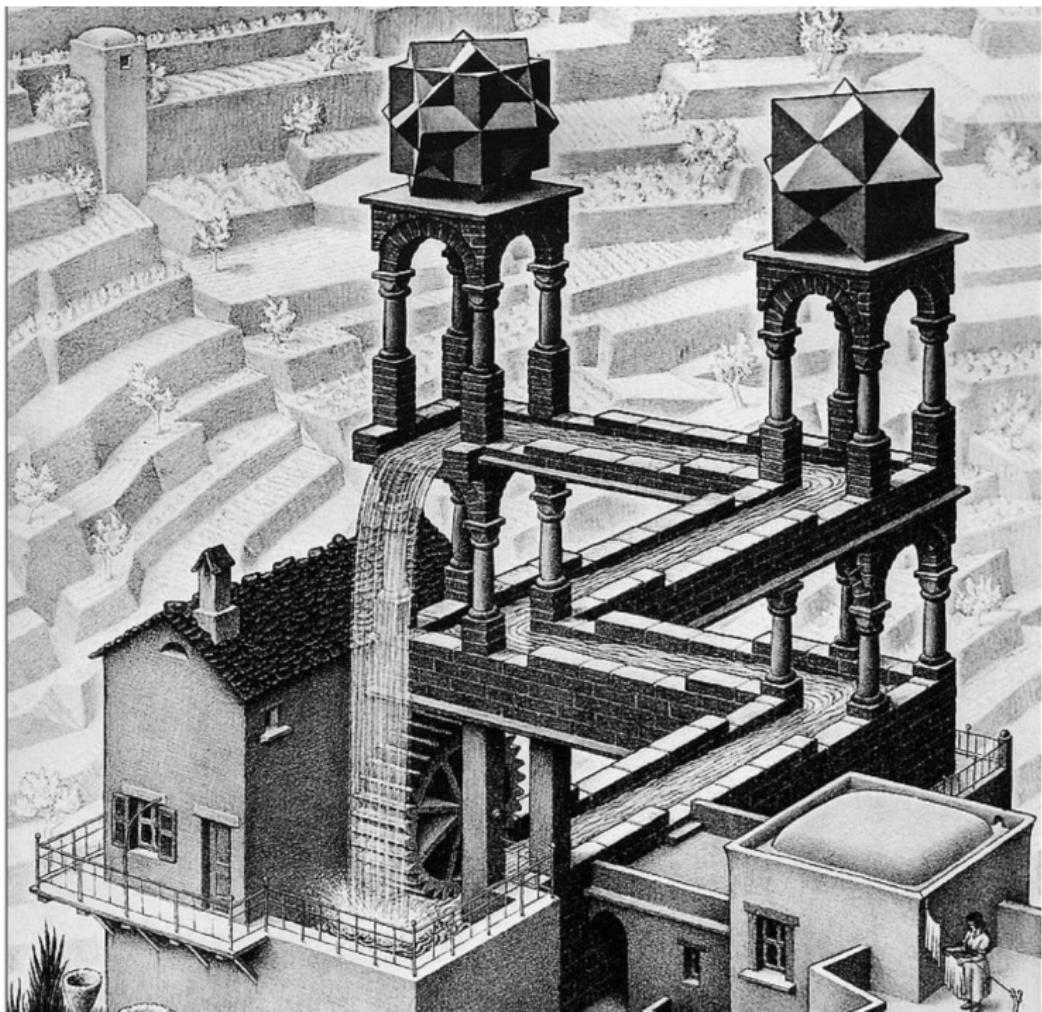
Directed Acyclic Graphs



C – Blocked backdoor pathway

Where to condition/adjust/block

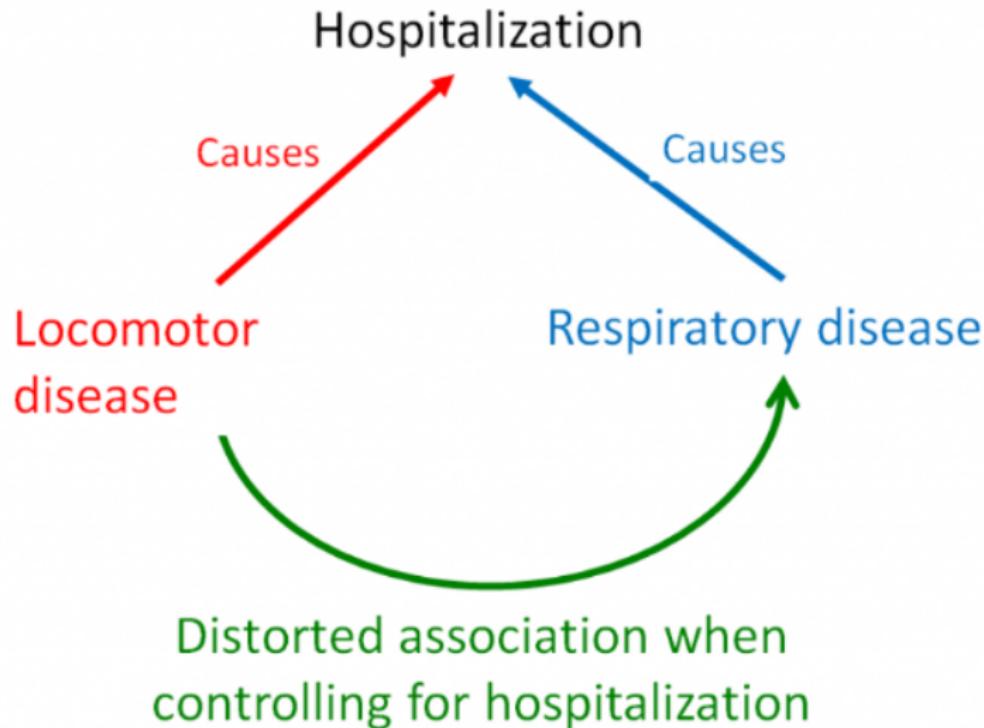




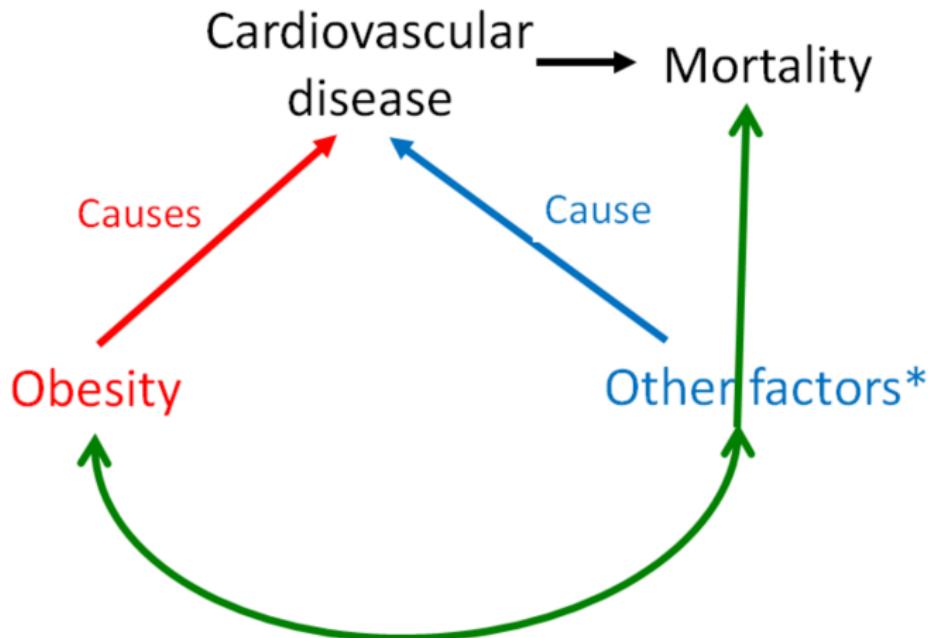
Blocking rules

1. Without conditioning a path is blocked if and only if two arrowhead collide at some node on the path.
2. A path that contains a non-collider that has been conditioned on is blocked.
3. A path that contains a collider is not blocked when the collider is conditioned on.
4. A collider that has a descendant that has been conditioned on does not block a path.

Collider Bias - Locomotor disease Respiratory disease

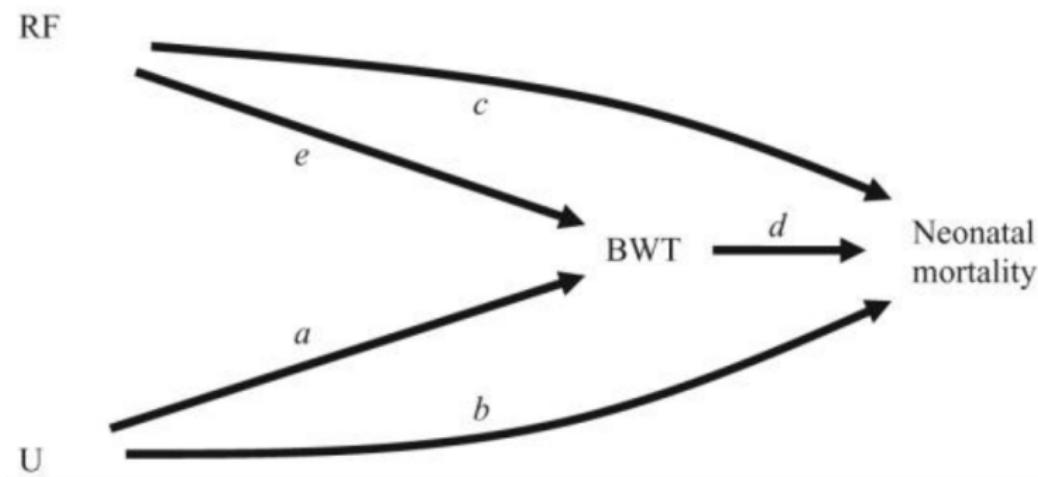


Collider Bias - Obesity Paradox



Distorted association when controlling
for cardiovascular disease

Collider Bias - Smoking Neonatal Mortality



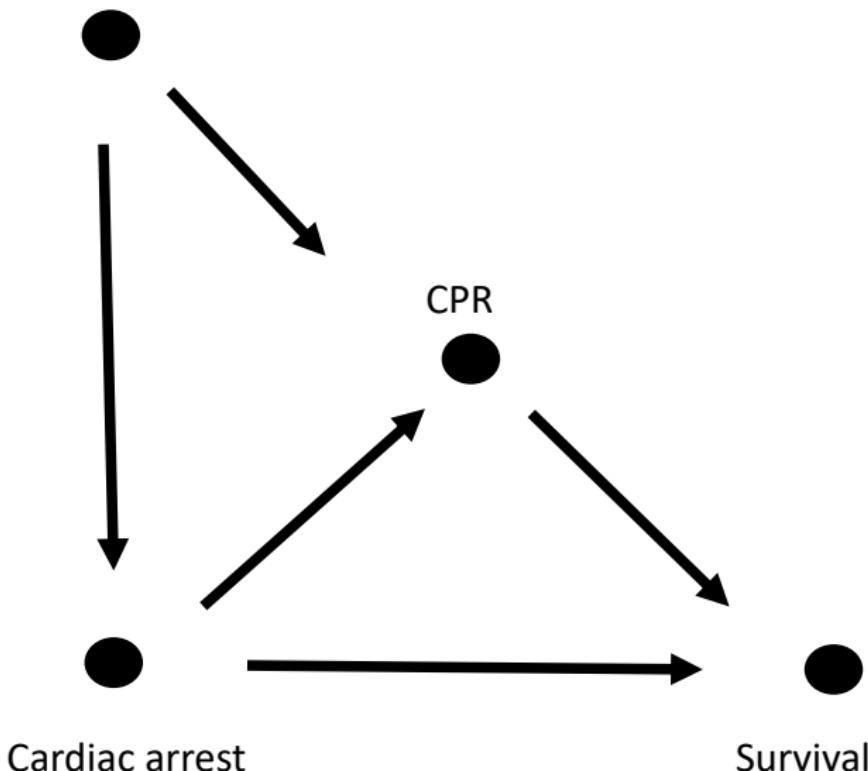
RF - Risk factor/ Smoking

U - Unknown confounder

BWT - Birth Weight

Cardiac arrest - Psychiatric disease

Psychicatric Disease



Safari File Edit View History Bookmarks Window Help

daggity.net

remote.dst.dk

Tue 18 Sep 10.13

DAGGITY v2.3

Model Examples How to ... Layout Help

Causal effect identification
Adjustment (total effect)

No adjustment is necessary to estimate the total effect of E on D.

Testable implications
The model implies the following conditional independencies:

- A $\perp\!\!\!\perp$ B
- A $\perp\!\!\!\perp$ D | E
- B $\perp\!\!\!\perp$ E
- D $\perp\!\!\!\perp$ A, B
- D $\perp\!\!\!\perp$ Z | A, B
- E $\perp\!\!\!\perp$ Z | A

Export R code

Model code

```
A 1 0-2.209,-1.526
A 1 0-2.209,1.526
D 0 0-1.450,1.182
E 0 0-2.209,1.597
Z 1 0-0.309,-0.082
A E 0-0.791,-1.645
A D 0 0.686,-0.496
A D
```

Diagram style
classic SEM-like

View mode
normal moral graph correlation graph

Coloring
cause paths biasing paths ancestral structure

Effect analysis
atomic direct effects

Legend
exposure outcome ancestor of exposure ancestor of outcome ancestor of exposure and outcome adjusted variable unobserved (latent) other variable causal path biasing path

Summary
exposure(s) E
outcome(s) D
covariates 3
causal paths 1



Interaction / Effect Modification

The Phenomenon that the importance of one variable is dependent on another

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \text{sex} + \beta_2 * \text{age} + \beta_3 * \text{CPR} + \beta_4 * \text{time} + \beta_5 * \text{sex} * \text{CPR}$$

This model is fine for interpreting whether there is interaction. To examine the importance of the interaction it is fruitful to establish three alternative variables:

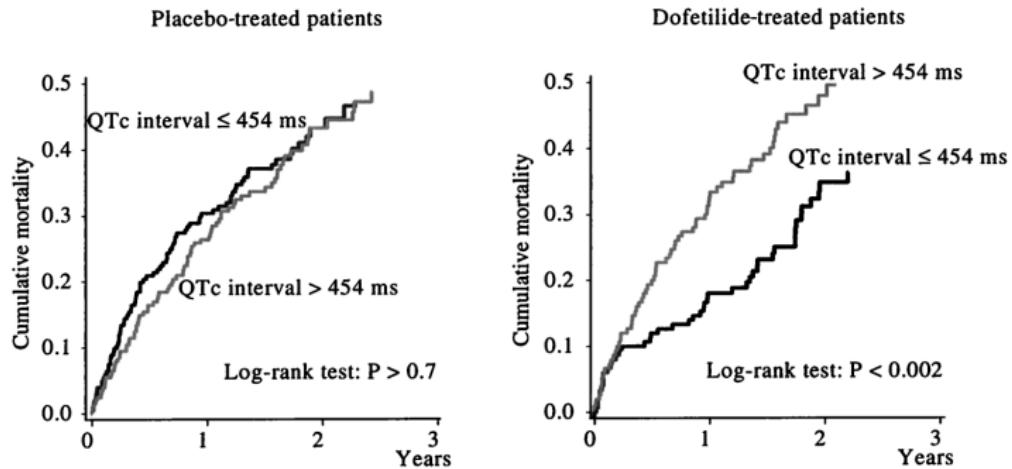
$\text{CPR}_{Male} = 1$ when male and CPR given, otherwise 0

$\text{CPR}_{Female} = 1$ when female and CPR given, otherwise 0

SEX

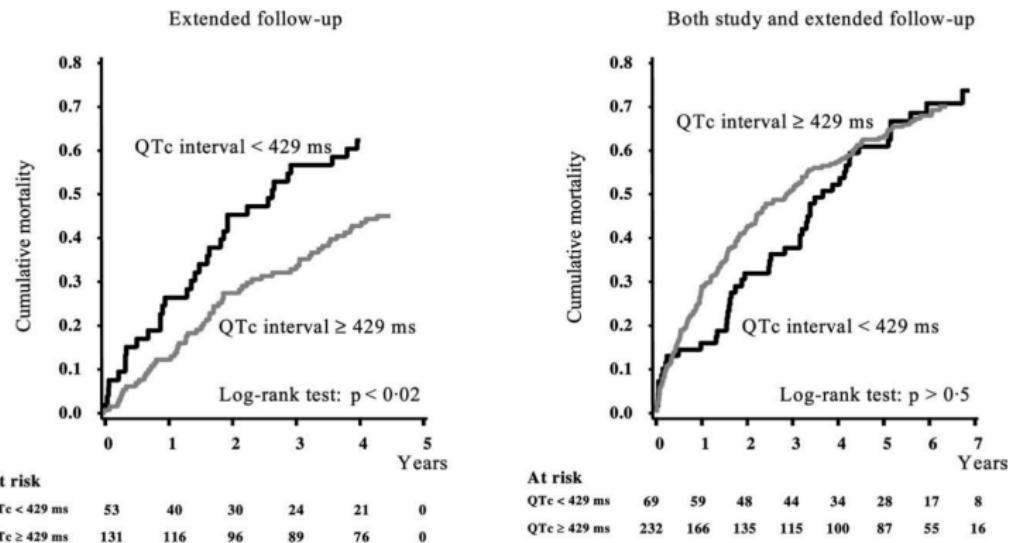
The two models with these alternative three variables represent the same interaction

Dofetilide QTc interaction



Brendorp, Circulation 2001

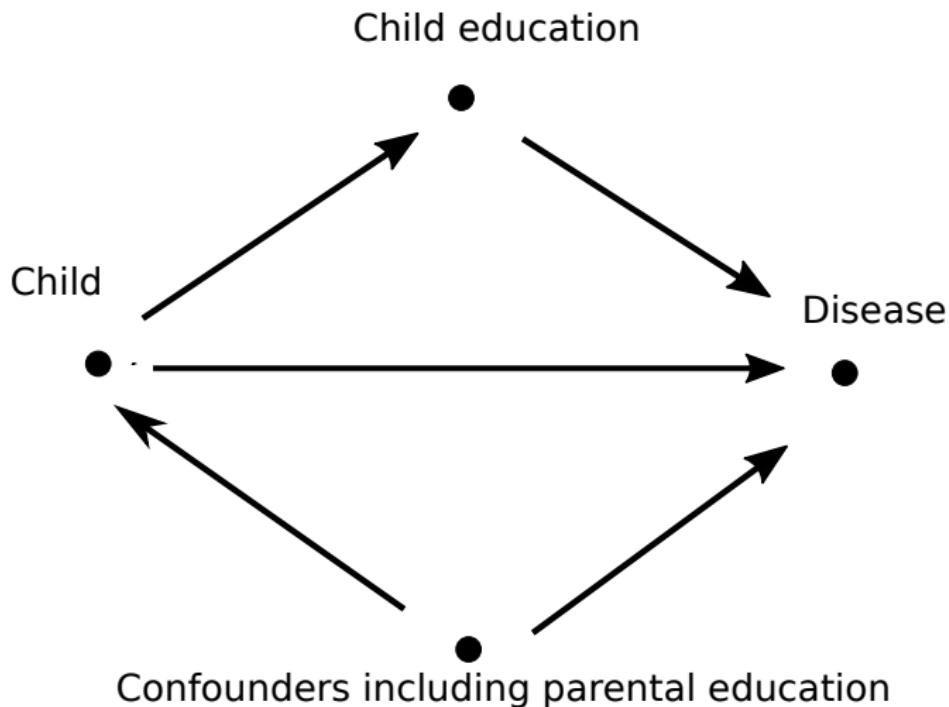
After dofetilide discontinuation



Brendorp, European Heart Journal 2003



Mediation



Mediation Analysis

Total effect - $TE = E[Y(1) - Y(0)]$

TE measures the expected increase in the outcome Y as X changes from $X=0$ to $X=1$, while the mediator is allowed to track the change in X as dictated by the function $M = g(X, 2)$.

Controlled direct effect - $CDE(m) = E[Y(1,m) - Y(0,m)]$

CDE measures the expected increase in the outcome Y as X changes from $X = 0$ to $X = 1$, while the mediator is fixed at a pre-specified level $M = m$ uniformly over the entire population

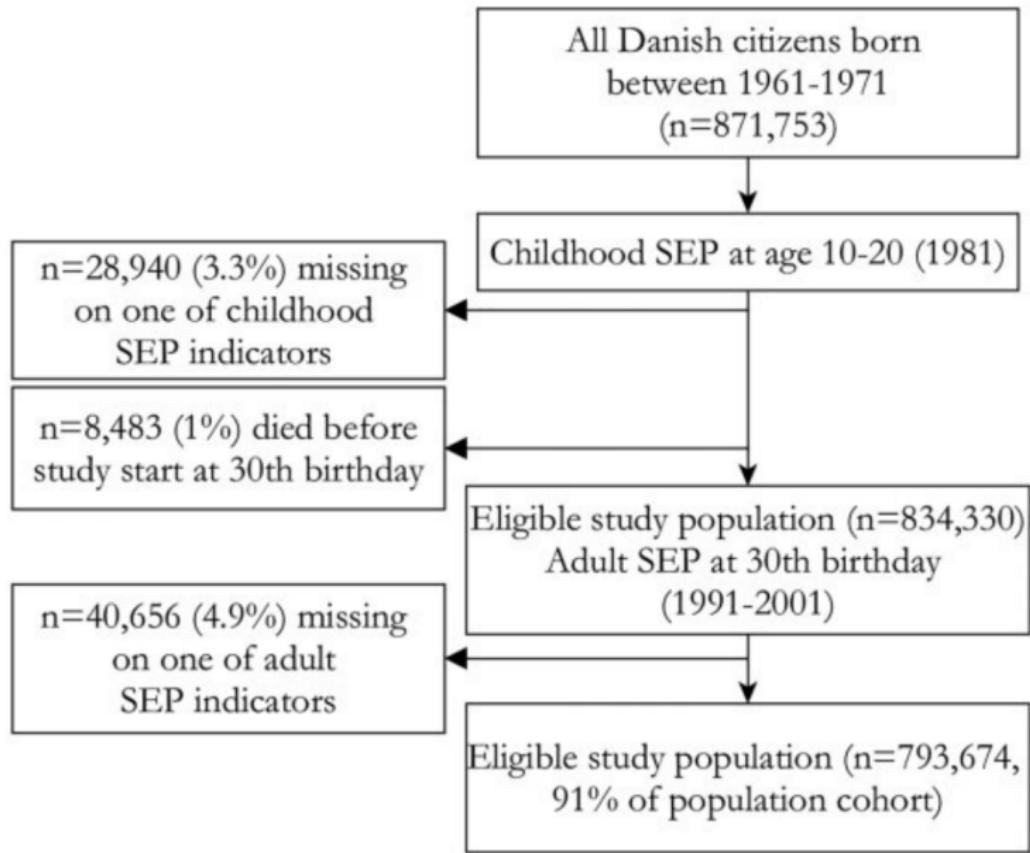
Natural direct effect - $NDE = E[Y(1,M(0)) - Y(0,M(0))]$

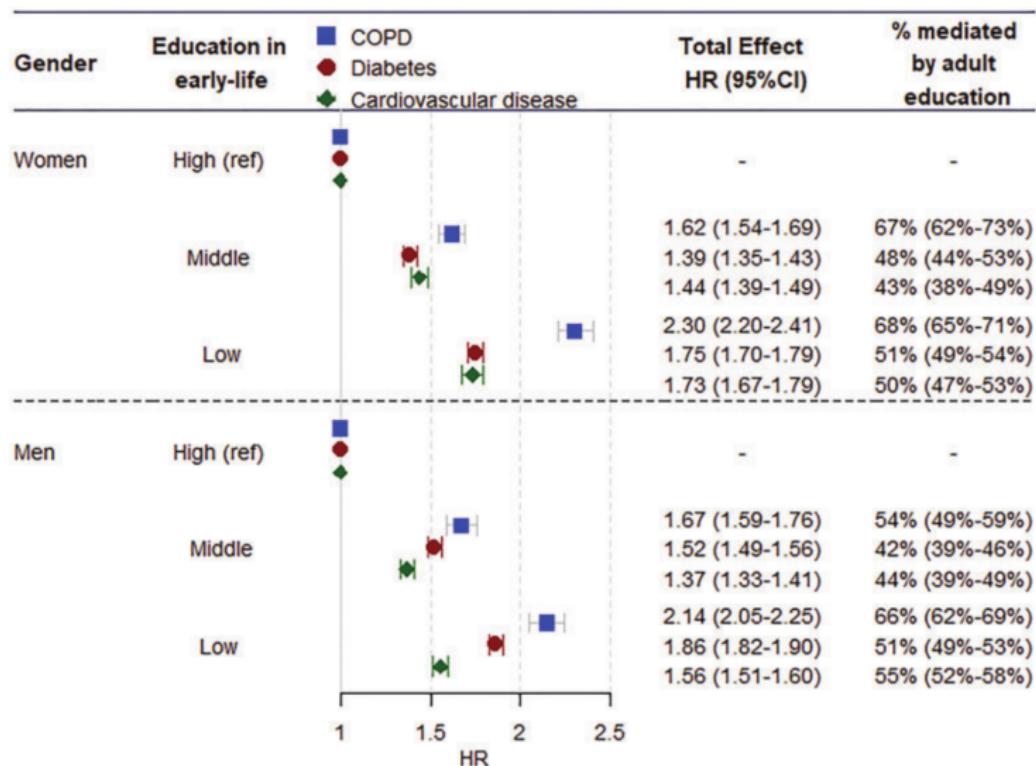
NDE measures the expected increase in Y as X changes from $X = 0$ to $X = 1$, while setting the mediator variable to whatever value it would have obtained under $X = 0$, i.e., before the change.

Natural indirect effect - $NIE = E[Y(0,M(1)) - Y(0,M(0))]$

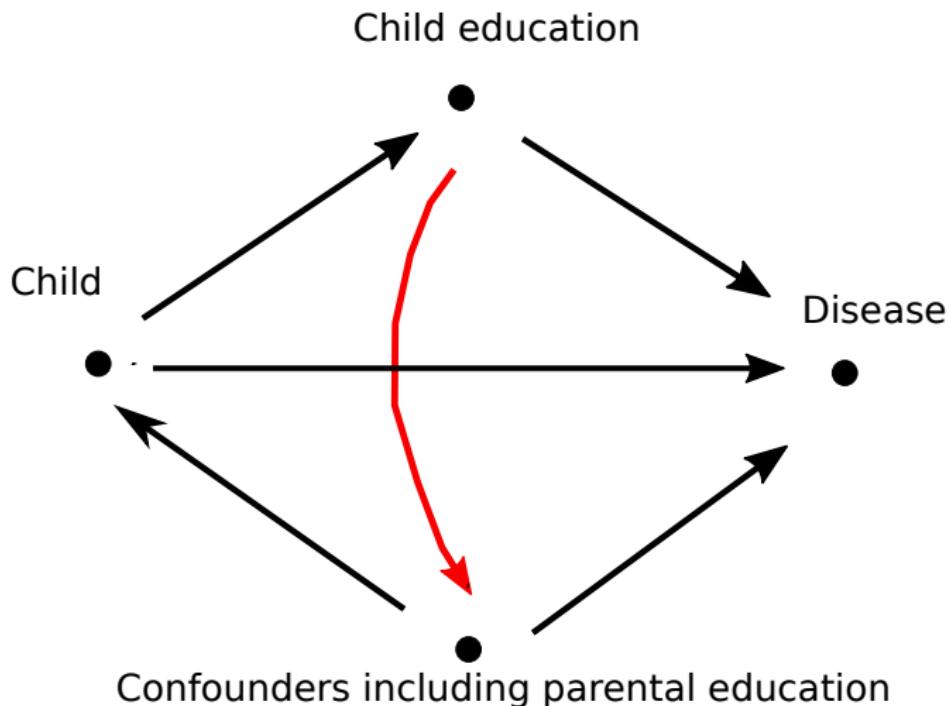
NIE measures the expected increase in Y when the X is held constant, at $X = 1$, and M changes to whatever value it would have attained (for each individual) under $X = 1$.

The difference $TE - NDE$ measures the extent to which mediation is necessary for explaining the effect, while the NIE measures the





Mediation? Interaction?



Bias

- ▶ Selection bias
- ▶ Information bias
- ▶ Confounding bias

Selection Bias

Prevalence incidence (Neyman) bias	A late look at those with a disease or condition will miss early problems and those that have died
Admission rate (Berkson) bias	A hospital-based study of the relation between a disease and some exposure will be biased if patients with the disease are more or less admitted to hospital depending on the exposure of interest
Immortal lifetime bias	When future events are included as baseline data those that have the future event will be immortal until the time when the future data were recorded
Unmasking (detection signal) bias	An innocent exposure may become associated with disease if it triggers search for a disease
Volunteer bias	Individuals volunteering for studies or seeking early help for symptoms may be more healthy than non-volunteers or latecomers
Response bias	People who agree to take part in a study have different characteristics from those that do not, and this distorts the results when making conclusions about the whole population
Withdrawal bias	If patients that discontinue a study differ importantly from those that remain in a study the final result may be severely distorted, in particular when only measurements at the end of the study, such as rhythm control can enter the analyses
Channelling bias	The propensity of 'sicker' or selected patients to be prescribed disproportionately the newer and perceived to be more potent medications differentially

Information Bias

Recall bias	Information that relies on patient memory may be influenced by their condition. If a relation between a disease and a symptom is available to the patient that may help the patient remember a condition
Insensitive measure bias	If the measurement used in a study does not detect what it is supposed to detect and underestimation of that measurement will be the result
Regression dilution bias	If a measurement is inaccurate the relation between the measurement and outcome is weakened. For comparison of continuous variables, the slope will be reduced
Follow-up bias	If follow-up depends on the presence of a condition this can create a false relation between a condition and a disease, the direction depending on whether the condition improves or worsens follow-up
Assessment bias	The assessment and thus collected data on a subject is influence by other factors
Interviewer bias	If an interviewer is aware of the subject's health status, this may influence the questions asked, or how they are asked, which consequently affects the response



166 Giovanni da Venezia

