PATHWAYS

Current Methods in Mediation Analysis

Bianca De Stavola and Rhian Daniel, LSHTM

IEU Symposium on Statistical Methods for Epigenetic Change
Bristol, 6 June 2014

Website http://pathways.lshtm.ac.uk
Email pathways@lshtm.ac.uk
Twitter @pathwaysNCRM

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

National Centre for Research Methods

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Mediation

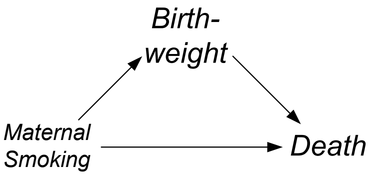
- In many research contexts we might be interested in the extent to which the effect of some exposure X on some outcome Y is mediated by an intermediate variable M.

Mediation

- In many research contexts we might be interested in the extent to which the effect of some exposure X on some outcome Y is mediated by an intermediate variable M.
- In other words we are interested in the study of mediation.

Example I

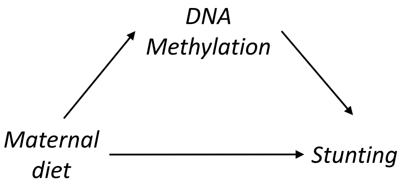
Maternal smoking, birth weight, and infant mortality



For example, how much of the effect of maternal smoking on infant mortality is due to its effect on birth weight?

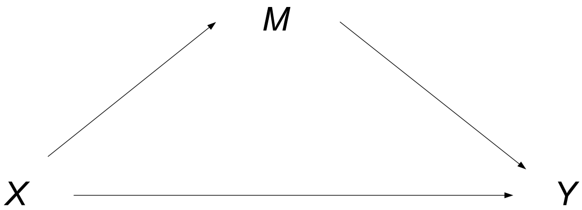
Example II

Maternal nutrition at conception, methylation and stunting



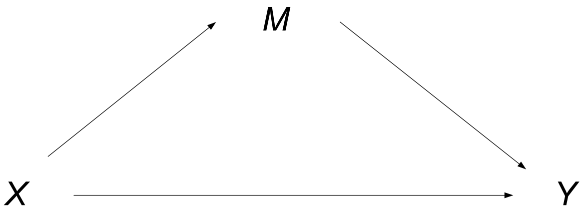
Study of Gambian infants (Dominiguez-Salas *et al.*, *Nature Comm*, 2014)

Typical setting



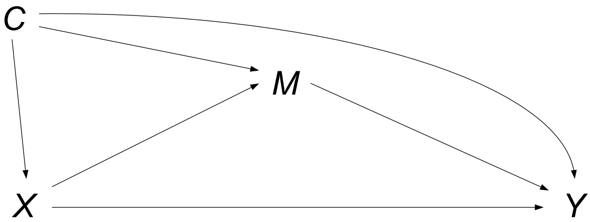
- Write X for the exposure, M for the mediator and Y for the outcome.

Typical setting



- Write X for the exposure, M for the mediator and Y for the outcome.
- Let M and Y be continuous.

Typical setting

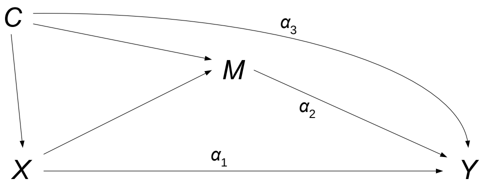


- Write X for the exposure, M for the mediator and Y for the outcome.
- Let M and Y be continuous.
- Let's explicitly include confounders C.

Outline

Outline

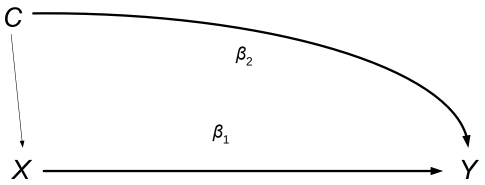
Combination of simple least squares regressions



Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

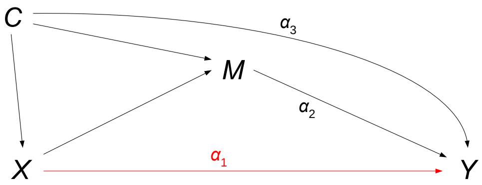
Combination of simple least squares regressions



Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

Combination of simple least squares regressions

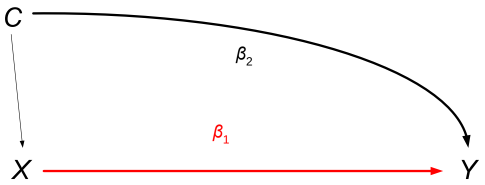


Consider two regression models:

$$E(Y|C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),

Combination of simple least squares regressions

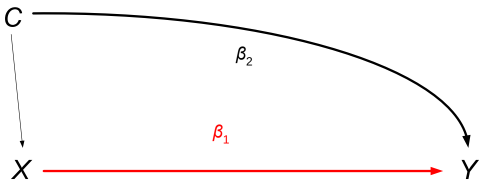


Consider two regression models:

$$E(Y|C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,

Combination of simple least squares regressions

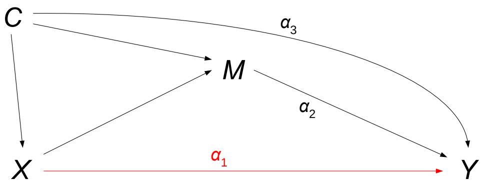


Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

Combination of simple least squares regressions

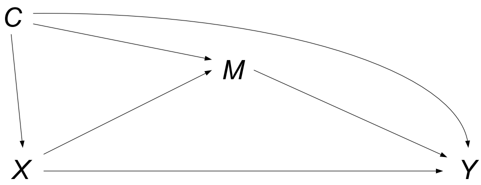


Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

Combination of simple least squares regressions

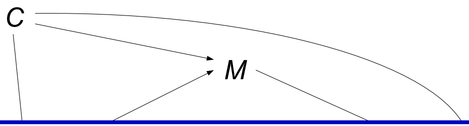


Consider two regression models:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

Combination of simple least squares regressions



- Estimation via ordinary least squares.
- Various options (delta method, bootstrapping) to obtain SE for the indirect effect.

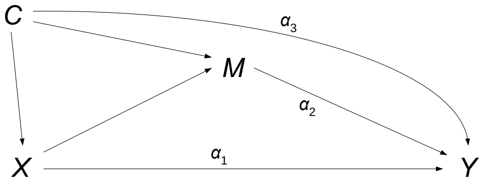
$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

Outline

A linear Structural Equation Model

Wright, 1921

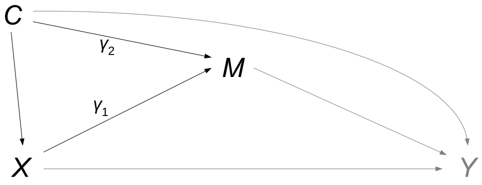


Alternatively, consider a (linear) structural equations model:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

A linear Structural Equation Model

Wright, 1921



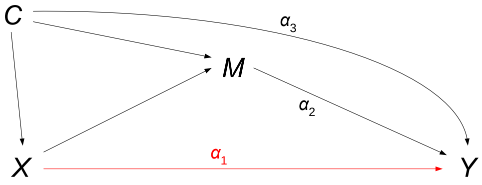
Alternatively, consider a (linear) structural equations model:

$$E(Y|C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

$$E(M|C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

A linear Structural Equation Model

Wright, 1921



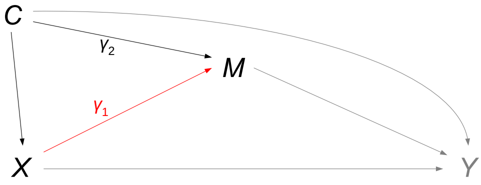
Alternatively, consider a (linear) structural equations model:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,

A linear Structural Equation Model

Wright, 1921



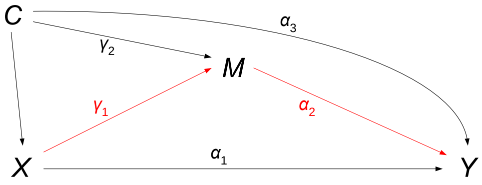
Alternatively, consider a (linear) structural equations model:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,
- and now $\gamma_1 \alpha_2$ is the indirect effect.

A linear Structural Equation Model

Wright, 1921



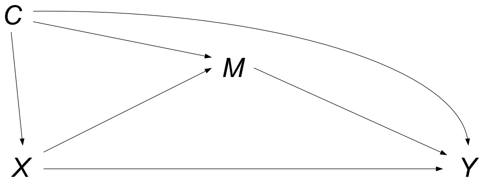
Alternatively, consider a (linear) structural equations model:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,
- and now $\gamma_1 \alpha_2$ is the indirect effect.

A linear Structural Equation Model

Wright, 1921



Alternatively, consider a (linear) structural equations model:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,
- and now $\gamma_1\alpha_2$ is the indirect effect.

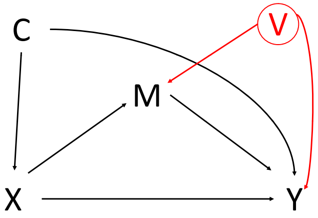
Outline

Limitations of these approaches

1. Lack of generality: Definitions are specific to simple linear models (in particular no X-M interactions).

Limitations of these approaches

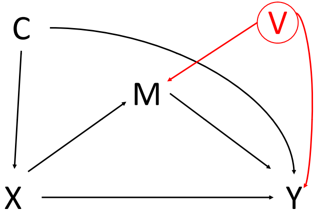
1. Lack of generality: Definitions are specific to simple linear models (in particular no X-M interactions).
2. Identifiability: often not appreciated that unaccounted confounders V of the M–Y relationship:



would bias the partitioning of direct/indirect effects.

Limitations of these approaches

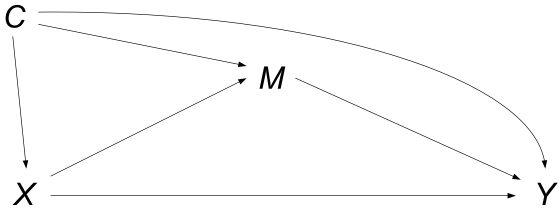
1. Lack of generality: Definitions are specific to simple linear models (in particular no X-M interactions).
2. Identifiability: often not appreciated that unaccounted confounders V of the M–Y relationship:



would bias the partitioning of direct/indirect effects.

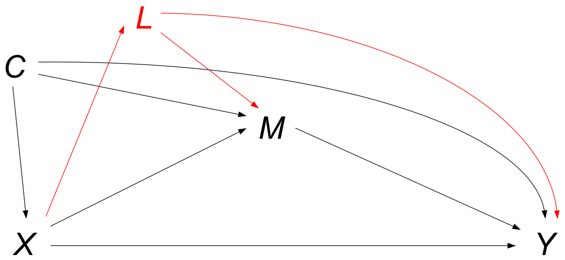
3. Intermediate confounding.

Problem 3: intermediate confounding



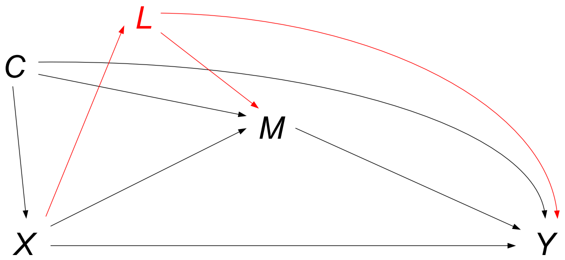
- Intermediate confounders L are common causes of M and Y that are affected by X.

Problem 3: intermediate confounding



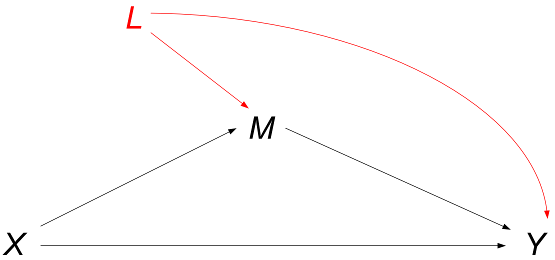
- Intermediate confounders L are common causes of M and Y that are affected by X.

Problem 3: intermediate confounding



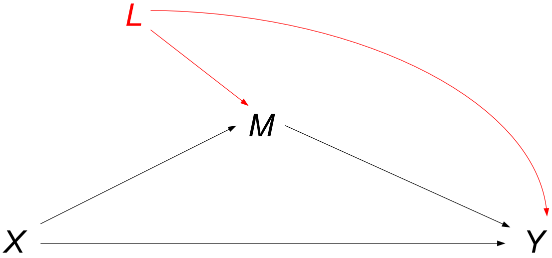
- Such L are problematic.

Problem 3: intermediate confounding



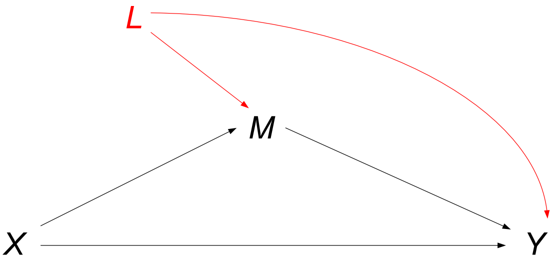
- Such L are problematic.
- Let us ignore C for simplicity, and, let us even ignore the arrow from X to L at first, ie L is NOT an intermediate confounder in this diagram for now…

Problem 3: intermediate confounding



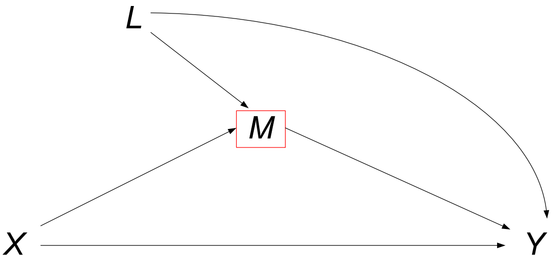
- L is a confounder for the M-Y relation and therefore cannot be ignored.

Problem 3: intermediate confounding



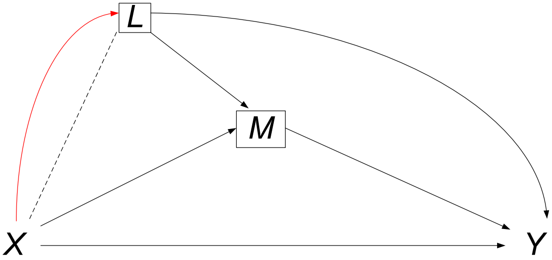
- *L* is a confounder for the *M*-*Y* relation and therefore cannot be ignored.
- However conditioning on *M* (in the model for *Y*) induces an association between *X* and *L* even if there was none there before (and would alter an existing association).

Problem 3: intermediate confounding



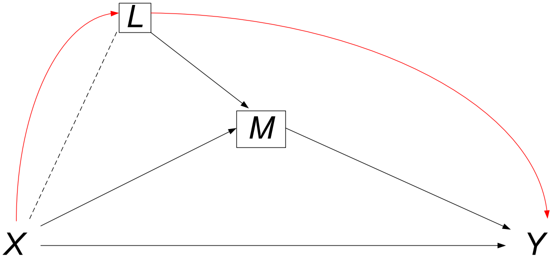
- L is a confounder for the M-Y relation and therefore cannot be ignored.
- However conditioning on M (in the model for Y) induces an association between X and L even if there was none there before (and would alter an existing association).
- Hence we should also condition on L . . .

Problem 3: intermediate confounding



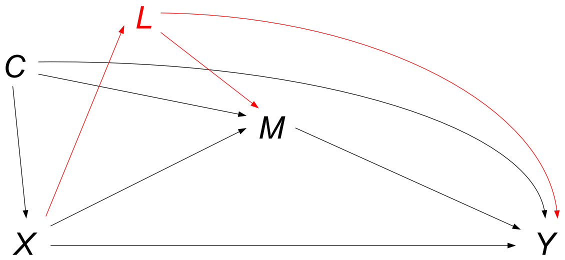
- But this is NOT a solution when L is affected by X.

Problem 3: intermediate confounding



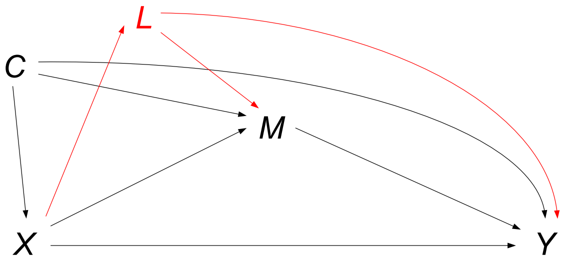
- But this is NOT a solution when L is affected by X.
- Since we block part of the direct effect (unmediated by M).

Problem 3: intermediate confounding



- But this is NOT a solution when L is affected by X.
- Since we block part of the direct effect (unmediated by M).
- Thus standard regression cannot be used when there is intermediate confounding.

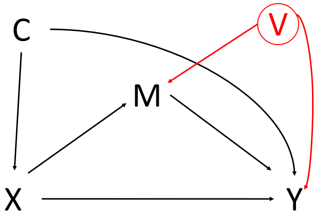
Problem 3: intermediate confounding



- But this is NOT a solution when *L* is affected by *X*.
- Since we block part of the direct effect (unmediated by *M*).
- Thus standard regression cannot be used when there is intermediate confounding.
- (SEMs could deal with this, but only for linear models for *L*, *M* and *Y* ...).

Limitations of these approaches

1. Lack of generality: Definitions are specific to simple linear models (in particular no X-M interactions).
2. Identifiability: often not appreciated that unaccounted confounders V of the M–Y relationship:

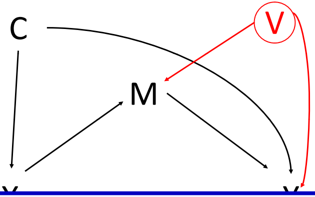


would bias the partitioning of direct/indirect effects.

3. Intermediate confounding.

Limitations of these approaches

1. Lack of generality: Definitions are specific to simple linear models (in particular no X-M interactions).
2. Identifiability: often not appreciated that unaccounted confounders V of the M–Y relationship:



More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.

Outline

Outline

Counterfactuals to the rescue!

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.

Counterfactuals to the rescue!

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.

Counterfactuals to the rescue!

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.

Counterfactuals to the rescue!

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

Counterfactuals to the rescue!

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

These counterfactuals are central to the (model-free) definitions of direct/indirect effects in causal inference.

Total causal effect

- The total causal effect of X on Y expressed as a mean difference is

$$\text{TCE} = E\{Y(1)\} - E\{Y(0)\}.$$

Total causal effect

- The total causal effect of X on Y expressed as a mean difference is

$$\text{TCE} = E\{Y(1)\} - E\{Y(0)\}.$$

- Note that this can also be written as

$$\text{TCE} = E[Y\{1, M(1)\}] - E[Y\{0, M(0)\}].$$

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y when M is controlled at m, expressed as a mean difference is

$$\text{CDE}(m) = E\{Y(1, m)\} - E\{Y(0, m)\}.$$

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y when M is controlled at m, expressed as a mean difference is

$$\text{CDE}\,(m) = E\left\{Y\,(1,m)\right\} - E\left\{Y\,(0,m)\right\}.$$

- This (as always with a causal contrast) is a comparison of two hypothetical worlds.

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y when M is controlled at m, expressed as a mean difference is

$$\text{CDE}\,(m) = E\,\{Y\,(1, m)\} - E\,\{Y\,(0, m)\}\,.$$

- This (as always with a causal contrast) is a comparison of two hypothetical worlds.
- In the first, X is set to 1, and in the second X is set to 0. In both worlds, M is set to m.

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y when M is controlled at m, expressed as a mean difference is

$$\text{CDE}(m) = E\{Y(1, m)\} - E\{Y(0, m)\}.$$

- This (as always with a causal contrast) is a comparison of two hypothetical worlds.
- In the first, X is set to 1, and in the second X is set to 0. In both worlds, M is set to m.
- By keeping M fixed at m, we are getting at the direct effect of X, unmediated by M.

Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.

Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this turns out not to be possible using this definition of a controlled direct effect.

Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this turns out not to be possible using this definition of a controlled direct effect.
- For this reason, it is useful to have a different definition of a direct effect.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- The natural direct effect of X on Y expressed as a mean difference is

$$\text{NDE} = E\left[Y\{1, M(0)\}\right] - E\left[Y\{0, M(0)\}\right].$$

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- The natural direct effect of X on Y expressed as a mean difference is

$$\text{NDE} = E\left[Y\left\{1, M\left(0\right)\right\}\right] - E\left[Y\left\{0, M\left(0\right)\right\}\right].$$

- This is a comparison of two hypothetical worlds.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- The natural direct effect of X on Y expressed as a mean difference is

$$\text{NDE} = E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to 1, and in the second X is set to 0. In both worlds, M is set to $M(0)$, the value it would take if X were set to 0.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- The natural direct effect of X on Y expressed as a mean difference is

$$\text{NDE} = E\left[Y\left\{1, M\left(0\right)\right\}\right] - E\left[Y\left\{0, M\left(0\right)\right\}\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to 1, and in the second X is set to 0. In both worlds, M is set to $M(0)$, the value it would take if X were set to 0.
- Since M is the same (*within* subject) in both worlds, we are still getting at the direct effect of X.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- The natural direct effect of X on Y expressed as a mean difference is

$$\text{NDE} = E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to 1, and in the second X is set to 0. In both worlds, M is set to $M(0)$, the value it would take if X were set to 0.
- Since M is the same (*within* subject) in both worlds, we are still getting at the direct effect of X.
- If no individual-level interaction between X and M, $\text{CDE}(m) = \text{NDE} \;\; \forall m$.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

- This is a comparison of two hypothetical worlds.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\{1, M(1)\}\right] - E\left[Y\{1, M(0)\}\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, M is set to $M(1)$ and in the second M is set to $M(0)$. In both worlds, X is set to 1.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, M is set to $M(1)$ and in the second M is set to $M(0)$. In both worlds, X is set to 1.
- X is allowed to influence Y only through its influence on M. Thus it is an indirect effect through M.

Effect decomposition

Now we see that the sum of the natural direct and indirect effects is

$$\begin{aligned} \text{NDE} + \text{NIE} = & \; E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] \\ & + E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right] \\ & = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] = \text{TCE}, \end{aligned}$$

the total causal effect, as desired.

Effect decomposition

Now we see that the sum of the natural direct and indirect effects is

$$\begin{aligned}\text{NDE} + \text{NIE} &= E\left[Y\{1, M(0)\}\right] - E\left[Y\{0, M(0)\}\right] \\ &\quad + E\left[Y\{1, M(1)\}\right] - E\left[Y\{1, M(0)\}\right] \\ &= E\left[Y\{1, M(1)\}\right] - E\left[Y\{0, M(0)\}\right] = \text{TCE},\end{aligned}$$

the total causal effect, as desired.

Effect decomposition

Now we see that the sum of the natural direct and indirect effects is

$$\begin{aligned} \text{NDE} + \text{NIE} &= E\left[Y\{1, M(0)\}\right] - E\left[Y\{0, M(0)\}\right] \\ &+ E\left[Y\{1, M(1)\}\right] - E\left[Y\{1, M(0)\}\right] \\ &= E\left[Y\{1, M(1)\}\right] - E\left[Y\{0, M(0)\}\right] = \text{TCE}, \end{aligned}$$

the total causal effect, as desired.

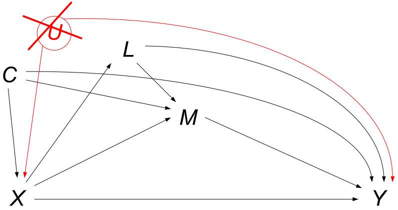
What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.

What next?

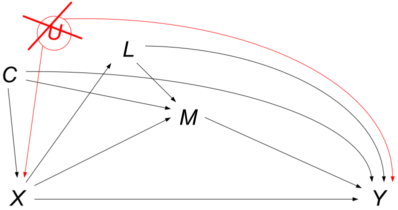
- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- As well as technical assumptions of no interference and consistency, there are no unmeasured confounding assumptions, and more...

Assumptions for identification: TCE



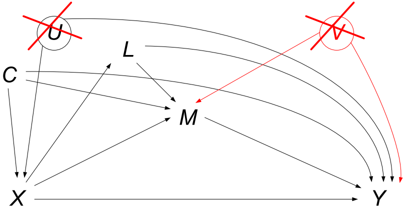
- No unmeasured confounding of the X–Y relationship.

Assumptions for identification: CDE



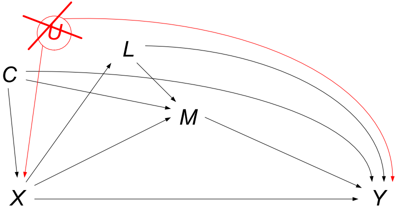
- No unmeasured confounding of the X–Y or M–Y relationships.

Assumptions for identification: CDE



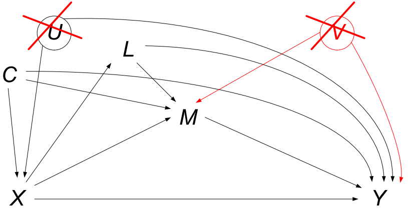
- No unmeasured confounding of the X–Y or M–Y relationships.

Assumptions for identification: NDE, NIE



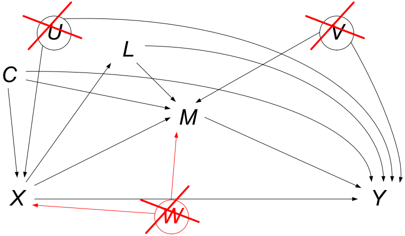
- No unmeasured confounding of the X–Y, M–Y, or X–M relationships.

Assumptions for identification: NDE, NIE



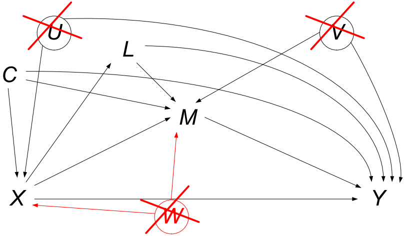
- No unmeasured confounding of the X–Y, M–Y, or X–M relationships.

Assumptions for identification: NDE, NIE



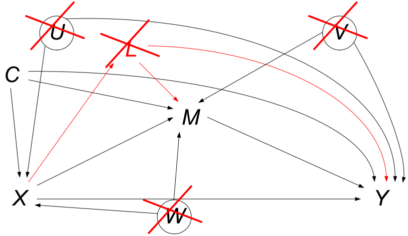
- No unmeasured confounding of the X–Y, M–Y, or X–M relationships.

Assumptions for identification: NDE, NIE



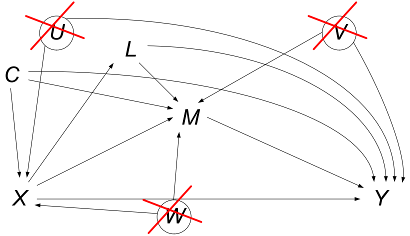
- No unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- AND, in addition, either:

Assumptions for identification: NDE, NIE



- No unmeasured confounding of the *X*–*Y*, *M*–*Y*, or *X*–*M* relationships.
- AND, in addition, either:
  - No intermediate confounding, or

Assumptions for identification: NDE, NIE



- No unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- AND, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al, 2006).

Outline

G-computation formula for the CDE
Robins 1986

- Let's look at how the CDE is estimated:

$$\mathrm{CDE}(m) = E\{Y(1,m)\} - E\{Y(0,m)\}$$
$$= \int E(Y|C=c, X=1, L=l, M=m)\, f_{L|C,X}(l|c,1)\, f_C(c) dl\, dc$$
$$- \int E(Y|C=c, X=0, L=l, M=m)\, f_{L|C,X}(l|c,0)\, f_C(c) dl\, dc$$

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\begin{aligned}\text{CDE}(m) &= E\{Y(1,m)\} - E\{Y(0,m)\}\\ &= \int E(Y|C=c, X=1, L=l, M=m)\, f_{L|C,X}(l|c,1)\, f_C(c)dl\, dc\\ &- \int E(Y|C=c, X=0, L=l, M=m)\, f_{L|C,X}(l|c,0)\, f_C(c)dl\, dc\end{aligned}$$

- This is the g-computation formula.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}(m) = E\{Y(1,m)\} - E\{Y(0,m)\}$$
$$= \int E(Y|C=c, X=1, L=l, M=m)\, f_{L|C,X}(l|c,1)\, f_C(c) dl\, dc$$
$$- \int E(Y|C=c, X=0, L=l, M=m)\, f_{L|C,X}(l|c,0)\, f_C(c) dl\, dc$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y|C,X,L,M$ and $L|C,X$.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$
\begin{aligned}
&\text{CDE}(m) = E\{Y(1,m)\} - E\{Y(0,m)\} \\
&= \int E(Y|C=c, X=1, L=l, M=m)\, f_{L|C,X}(l|c,1)\, f_C(c)\,dl\,dc \\
&- \int E(Y|C=c, X=0, L=l, M=m)\, f_{L|C,X}(l|c,0)\, f_C(c)\,dl\,dc
\end{aligned}
$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y|C,X,L,M$ and $L|C,X$.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\begin{aligned} \text{CDE}(m) &= E\{Y(1,m)\} - E\{Y(0,m)\} \\ &= \int E(Y|C=c, X=1, L=l, M=m)\, f_{L|C,X}(l|c,1)\, f_C(c) dl\, dc \\ &- \int E(Y|C=c, X=0, L=l, M=m)\, f_{L|C,X}(l|c,0)\, f_C(c) dl\, dc \end{aligned}$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y|C,X,L,M$ and $L|C,X$.
- Both models can be completely flexible: they can include non-linearities and interactions.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}(m) = E\{Y(1,m)\} - E\{Y(0,m)\}$$

$$= \int E(Y|C=c, X=1, L=l, M=m)\, f_{L|C,X}(l|c,1)\, f_C(c)\, dl\, dc$$

$$- \int E(Y|C=c, X=0, L=l, M=m)\, f_{L|C,X}(l|c,0)\, f_C(c)\, dl\, dc$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y|C,X,L,M$ and $L|C,X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L|C,X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M \mid C, X, L$ is now required.

G-computation formula for the NDE/NIE

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M \mid C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$
\begin{aligned}
E\{Y(1,m) - Y(0,m) \mid C = c, M(0) = m\} \\
= E\{Y(1,m) - Y(0,m) \mid C = c\}.
\end{aligned}
$$

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M \mid C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$E\{Y(1,m) - Y(0,m) \mid C = c, M(0) = m\}$$
$$= E\{Y(1,m) - Y(0,m) \mid C = c\}.$$

- This can also be carried out in Stata's `gformula` command.

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L \mid C, X$:

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L|C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L|C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),

Alternative semiparametric approches

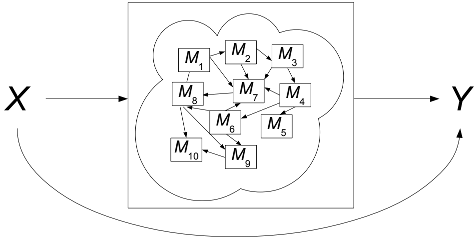
- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L|C,X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L|C,X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L|C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L|C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

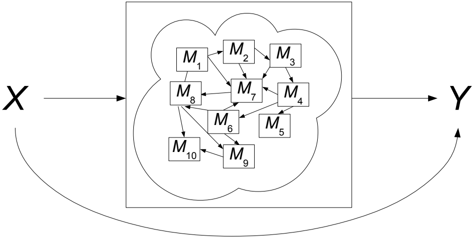
Outline

Multiple mediators



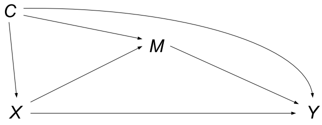
- Often there are many mediators of interest, eg many metabolites potentially mediating the relationship between CVD SNPs and CVD.

Multiple mediators



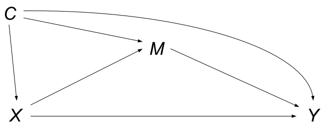
- Often there are many mediators of interest, eg many metabolites potentially mediating the relationship between CVD SNPs and CVD.
- Unless these do not causally affect one another (unlikely), and if we are interested in path-specific effects, this makes things much more complicated (Daniel et al, under review).

Interactions



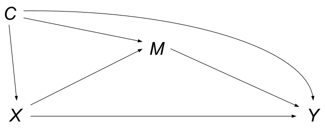
- Interest may lie not in mediation *per se* but in interactions between X and M, eg SNP and DNA methylation.

Interactions



- Interest may lie not in mediation *per se* but in interactions between X and M, eg SNP and DNA methylation.
- Standard approaches (regression of Y on X, M, $X * M$ and confounders) are then essentially attempting to estimate the CDE at each m and assess whether these CDEs are all the same.

Interactions



- Interest may lie not in mediation *per se* but in interactions between X and M, eg SNP and DNA methylation.
- Standard approaches (regression of Y on X, M, $X * M$ and confounders) are then essentially attempting to estimate the CDE at each m and assess whether these CDEs are all the same.
- But if there are unmeasured confounding of M and Y, for example, this would lead to bias in these estimates and, potentially, to misleading conclusions about the presence and magnitude of any interaction.

Outline

Quick summary

- Traditional approaches are somewhat limited, and too vague for these limitations to be always apparent.

Quick summary

- Traditional approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.

Quick summary

- Traditional approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.

Quick summary

- Traditional approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.

Quick summary

- Traditional approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the traditional approach even if we didn't realise it.

Quick summary

- Traditional approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the traditional approach even if we didn't realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Outline

References

Baron, R. M. and Kenny, D. A. (1986)
The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations
Journal of Personality and Social Psychology, 51:1173–1182.

Wright, S. (1921)
Correlation and causation: Part I - Method of Path Coefficients
Journal of Agriculture Research, 20:557–575.

Pearl, J. (2001)
Direct and indirect effects
Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 411–420.

References

- Robins, J. M. and Greenland, S. (1992):
  Identifiability and exchangeability for direct and indirect effects
  Epidemiology, 3:143–155.
- VanderWeele, T. J. and Vansteelandt, S. (2010)
  Odds ratios for mediation analysis for a dichotomous outcome
  American Journal of Epidemiology, 172:1339–1348.
- Petersen, M. L., Sinisi, S. E. and van der Laan, M. J. (2006)
  Estimation of direct causal effects
  Epidemiology, 17:276–284.

References

Robins, J. M. (1986)
A new approach to causal inference in mortality studies with a sustained exposure period – application to control of the healthy worker survivor effect
Mathematical Modelling, 7:1393–1512.

Daniel, R. M., De Stavola, B. L. and Cousens, S. N. (2011)
gformula: Estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula
The Stata Journal, 11:479–517.

References

Ten Have, T. R. and Joffe, M. (2012)
A review of causal estimation of effects in mediation analyses
Statistical Methods in Medical Research, 21:77–107.

Imai, K., Keele, L. and Yamamoto, T. (2010)
Identification, inference, and sensitivity analysis for causal mediation effects
Statistical Science, 25:51–71.

Daniel, R. M., De Stavola, B. L., Cousens, S. N. and Vansteelandt, S. (under review)
Causal mediation analysis with multiple mediators
Biometrics.