Causal mediation analysis with multiple causally-ordered mediators

Rhian Daniel and Bianca De Stavola

Centre for Statistical Methodology
London School of Hygiene and Tropical Medicine

Meeting on Developments and challenges in observational data analysis
Manchester 17th October 2012

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

— Evidence suggests that increased alcohol intake causes increased SBP.
— What is the mechanism?

— Evidence suggests that increased alcohol intake causes increased SBP.
— What is the mechanism?
— One possible pathway is through GGT, a liver enzyme.

— Evidence suggests that increased alcohol intake causes increased SBP.
— What is the mechanism?
— One possible pathway is through GGT, a liver enzyme.
— How much of the effect of alcohol on SBP is via GGT?

— Evidence suggests that increased alcohol intake causes increased SBP.
— What is the mechanism?
— One possible pathway is through GGT, a liver enzyme.
— How much of the effect of alcohol on SBP is via GGT?
— This is the standard mediation question.

— Another possible pathway is through BMI.
— Can we further decompose the effect?

— Another possible pathway is through BMI.
— Can we further decompose the effect?
— But BMI affects GGT: causally-ordered mediators.

— Another possible pathway is through BMI.
— Can we further decompose the effect?
— But BMI affects GGT: causally-ordered mediators.
— Can we decompose the effect of alcohol on SBP into an effect via BMI only,

— Another possible pathway is through BMI.
— Can we further decompose the effect?
— But BMI affects GGT: causally-ordered mediators.
— Can we decompose the effect of alcohol on SBP into an effect via BMI only, via GGT only,

— Another possible pathway is through BMI.
— Can we further decompose the effect?
— But BMI affects GGT: causally-ordered mediators.
— Can we decompose the effect of alcohol on SBP into an effect via BMI only, via GGT only, via both BMI and GGT,

— Another possible pathway is through BMI.
— Can we further decompose the effect?
— But BMI affects GGT: causally-ordered mediators.
— Can we decompose the effect of alcohol on SBP into an effect via BMI only, via GGT only, via both BMI and GGT, and via neither?

— More generally, how do we de ne and estimate path-speci c effects through n causally-ordered mediators?

— More generally, how do we de ne and estimate path-speci c effects through n causally-ordered mediators?
— Focus has either been on setting where mediators don't affect each other (MacKinnon 2000, Preacher & Hayes 2008)

— More generally, how do we de ne and estimate path-speci c effects through n causally-ordered mediators?

— Focus has either been on setting where mediators don't affect each other (MacKinnon 2000, Preacher & Hayes 2008)

— Or on effects that don't decompose (Avin et al 2005, Albert & Nelson 2011).

1 Mediation analysis with one mediator
- Traditional approach
- Limitations of the traditional approach
- Contributions from causal inference

2 Mediation analysis with multiple mediators

3 Back to the motivating example

4 Summary

1 Mediation analysis with one mediator
- Traditional approach
- Limitations of the traditional approach
- Contributions from causal inference

2 Mediation analysis with multiple mediators

3 Back to the motivating example

4 Summary

1 Mediation analysis with one mediator
- Traditional approach
- Limitations of the traditional approach
- Contributions from causal inference

2 Mediation analysis with multiple mediators

3 Back to the motivating example

4 Summary

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

- Write X for the exposure, M for the mediator and Y for the outcome.
- Explicitly include confounders C.

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is interpreted as the direct effect (not via M),

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is interpreted as the direct effect (not via M),
- ${}_1$ is interpreted as the total effect,

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is interpreted as the direct effect (not via M),
- ${}_1$ is interpreted as the total effect,
- and thus ${}_1 \quad {}_1$ is the indirect effect (via M).

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is interpreted as the direct effect (not via M),
- ${}_1$ is interpreted as the total effect,
- and thus ${}_1 \quad {}_1$ is the indirect effect (via M).

Consider two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is interpreted as the direct effect (not via M),
- ${}_1$ is interpreted as the total effect,
- and thus ${}_1 \quad {}_1$ is the indirect effect (via M).

Alternatively, consider the two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

Alternatively, consider the two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

Alternatively, consider the two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is (as before) the direct effect,

Alternatively, consider the two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is (as before) the direct effect,
- and now ${}_1 \; {}_2$ is the indirect effect.

Alternatively, consider the two regression models:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- ${}_1$ is (as before) the direct effect,
- and now ${}_1 \; {}_2$ is the indirect effect.

Alternatively, consider the two regression models:

$$E(Y \mid C; X; M) = \quad {}_0 + \quad {}_1 X + \quad {}_2 M + \quad {}_3^T C$$

$$E(M \mid C; X) = \quad {}_0 + \quad {}_1 X + \quad {}_2^T C$$

- ${}_1$ is (as before) the direct effect,
- and now ${}_1 \quad {}_2$ is the indirect effect.
- These (`difference', `product') are equivalent:

  ${}_1 \qquad {}_1 = \quad {}_1 \quad {}_2$.

1 Mediation analysis with one mediator
- Traditional approach
- Limitations of the traditional approach
- Contributions from causal inference

2 Mediation analysis with multiple mediators

3 Back to the motivating example

4 Summary

Consider the regression model that includes $_1$:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

Consider the regression model that includes $_1$:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

- $_1$ is just a parameter in an associational model.

Consider the regression model that includes $_1$:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

- $_1$ is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.

Consider the regression model that includes $_1$:

$$E(Y \mid C; X; M) = \ _0 + \ _1 X + \ _2 M + \ _3^T C$$

- $_1$ is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.
- Under what conditions does this correspond to a direct effect?

Consider the regression model that includes $_1$:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

- $_1$ is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.
- Under what conditions does this correspond to a direct effect?
- Does this question even make sense without a de nition (other than $_1$) of `direct effect'?

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- These are exible for C.

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_{31} C_1 + {}_{32} C_2 + {}_{33} C_1 C_2$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_{31} C_1 + {}_{32} C_2 + {}_{33} C_1 C_2$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_{31} C_1 + {}_{32} C_2 + {}_{33} C_1 C_2$$

- These are exible for C.
- We can include interactions between confounders, and non-linearities for the confounders.

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_{31} C_1 + {}_{32} \exp(C_2)$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_{31} C_1 + {}_{32} C_1^2 + {}_{33} C_2$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_{31} C_1 + {}_{32} \log(C_2) + {}_{33} \sqrt{C_1 C_2}$$

- These are exible for C.
- We can include interactions between confounders, and non-linearities for the confounders.

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- These are exible for C.
- We can include interactions between confounders, and non-linearities for the confounders.
- But the exibility ends here.

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C + {}_4 XM$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C$$

- What about interactions between X and other terms?

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C + {}_4^T XC$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C + {}_3^T XC$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C + {}_3^T XC$$

- What about interactions between X and other terms?

Consider the regression models again:

$$E(Y \mid C; X; M) = {}_0 + {}_1 X + {}_2 M + {}_3^T C + {}_4 X^2$$

$$E(Y \mid C; X) = {}_0 + {}_1 X + {}_2^T C + {}_3 X^2$$

$$E(M \mid C; X) = {}_0 + {}_1 X + {}_2^T C + {}_3 X^2$$

- What about interactions between X and other terms?
- Non-linearities for X?

Consider the regression models again:

$$E(Y \mid C; X; M) = \text{logit}^{-1}(\beta_0 + \beta_1 X + \beta_2 M + \beta_3^T C)$$

$$E(Y \mid C; X) = \text{logit}^{-1}(\alpha_0 + \alpha_1 X + \alpha_2^T C)$$

$$E(M \mid C; X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?

Consider the regression models again:

$$E(Y \mid C; X; M) = \text{logit}^{-1}(\beta_0 + \beta_1 X + \beta_2 M + \beta_3^T C + \beta_4 XM)$$

$$E(Y \mid C; X) = \text{logit}^{-1}(\alpha_0 + \alpha_1 X + \alpha_2^T C + \alpha_3^T XC)$$

$$E(M \mid C; X) = \gamma_0 + \gamma_1 X + \gamma_2^T C + \gamma_3 X^2$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?
- Not clear how `difference' and `product' methods extend.

- Thus far, we have considered confounders C that are (potentially) common causes of X, M and Y.

- Thus far, we have considered confounders C that are (potentially) common causes of X, M and Y.
- Intermediate confounders L are common causes of M and Y that are affected by X.

- Thus far, we have considered confounders C that are (potentially) common causes of X, M and Y.
- Intermediate confounders L are common causes of M and Y that are affected by X.
- It is unclear how the traditional approaches can be used when there is intermediate confounding.

1 Mediation analysis with one mediator
- Traditional approach
- Limitations of the traditional approach
- Contributions from causal inference

2 Mediation analysis with multiple mediators

3 Back to the motivating example

4 Summary

- 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 $Y(x)$ 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 $Y(x)$ 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 $Y(x)$ 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.

- 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.

- Many (subtly different) counterfactual de nitions of direct/indirect effects have been proposed.

- Many (subtly different) counterfactual de nitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).

- Many (subtly different) counterfactual de nitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).

- Many (subtly different) counterfactual de nitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

- Many (subtly different) counterfactual de nitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

- The total causal effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x vs x is

$$\mathrm{TCE}(c; x; x) = E\{Y(x)|C = c\} - E\{Y(x)|C = c\}:$$

- The total causal effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x vs x is

$$TCE(c; x; x) = E\{Y(x) \mid C = c\} - E\{Y(x) \mid C = c\}:$$

- Note that this can also be written as

$$TCE(c; x; x) = E[Y\{x, M(x)\} \mid C = c] - E[Y\{x, M(x)\} \mid C = c]:$$

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x vs x is

$$\mathrm{CDE}(c; x; x\ ; m) = E\{Y(x\ ; m) \mid C = c\} - E\{Y(x; m) \mid C = c\}:$$

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x vs x is

$$\mathrm{CDE}(c; x; x\ ; m) = E\{Y(x\ ; m) \mid C = c\} - E\{Y(x; m) \mid C = c\}:$$

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

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x vs x is

$$\mathrm{CDE}(c; x; x\ ; m) = E\{Y(x\ ; m) \mid C = c\} - E\{Y(x; m) \mid C = c\}:$$

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

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x vs x is

$$\mathrm{CDE}(c; x; x\ ; m) = E\{Y(x\ ; m) \mid C = c\} - E\{Y(x; m) \mid C = c\}:$$

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

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

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our de nition of a controlled direct effect.

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our de nition of a controlled direct effect.
- The controlled direct effect could be different for each value of m.

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our de nition of a controlled direct effect.
- The controlled direct effect could be different for each value of m.
- But the total causal effect is independent of m.

- Even with no mediation, an interaction would mean that

$$\mathrm{TCE}(c; x; x\ ) \neq \mathrm{CDE}(c; x; x\ ; m):$$

- Even with no mediation, an interaction would mean that

$$\mathrm{TCE}(c; x; x\ ) \neq \mathrm{CDE}(c; x; x\ ; m):$$

- But in this case, the indirect effect would intuitively be zero.

- Even with no mediation, an interaction would mean that

$$\mathrm{TCE}(c; x; x\ ) \neq \mathrm{CDE}(c; x; x\ ; m):$$

- But in this case, the indirect effect would intuitively be zero.
- And so we cannot hope to nd a de nition of a controlled indirect effect (CIE) such that

$$\mathrm{TCE}(c; x; x\ ) = \mathrm{CDE}(c; x; x\ ; m) + \mathrm{CIE}(c; x; x\ ; m):$$

- For this reason, it is useful to have a different de nition of a direct effect.

- For this reason, it is useful to have a different de nition of a direct effect.
- The pure natural direct effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{PNDE}(c; x; x\ ) = E[Y\{x\ ; M(x)\} \mid C = c] - E[Y\{x; M(x)\} \mid C = c]:$$

- For this reason, it is useful to have a different de nition of a direct effect.
- The pure natural direct effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{PNDE}(c; x; x\ ) = E[Y\{x\ ; M(x)\} \mid C = c] - E[Y\{x; M(x)\} \mid C = c]:$$

- This is a comparison of two hypothetical worlds.

- For this reason, it is useful to have a different de nition of a direct effect.
- The pure natural direct effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{PNDE}(c; x; x) = E[Y\{x, M(x)\} \mid C = c] - E[Y\{x, M(x)\} \mid C = c].$$

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

- For this reason, it is useful to have a different de nition of a direct effect.
- The pure natural direct effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{PNDE}(c; x; x) = E[Y\{x, M(x)\} \mid C = c] - E[Y\{x, M(x)\} \mid C = c].$$

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

- For this reason, it is useful to have a different de nition of a direct effect.
- The pure natural direct effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{PNDE}(c; x; x\ ) = E[Y\{x\ ; M(x)\} \mid C = c] - E[Y\{x; M(x)\} \mid C = c]:$$

- This is a comparison of two hypothetical worlds.
- In the rst, X is set to x, and in the second X is set to x. In both worlds, M is set to $M(x)$, the value it would take if X were set to x.
- 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, $\mathrm{CDE}(c; x; x\ ; m) = \mathrm{PNDE}(c; x; x\ )\ \ 8m$.

- The advantage of de ning the pure natural direct effect in this way, is that it leads to a natural indirect effect.

- The advantage of de ning the pure natural direct effect in this way, is that it leads to a natural indirect effect.
- The total natural indirect effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{TNIE}(c; x; x) = E[Y\{x, M(x)\} \mid C = c] - E[Y\{x, M(x)\} \mid C = c]:$$

- The advantage of de ning the pure natural direct effect in this way, is that it leads to a natural indirect effect.
- The total natural indirect effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{TNIE}(c; x; x) = E[Y\{x; M(x)\} \mid C = c] - E[Y\{x; M(x)\} \mid C = c].$$

- This is a comparison of two hypothetical worlds.

- The advantage of de ning the pure natural direct effect in this way, is that it leads to a natural indirect effect.
- The total natural indirect effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{TNIE}(c; x; x) = E[Y\{x; M(x)\} \mid C = c] - E[Y\{x; M(x)\} \mid C = c]:$$

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

- The advantage of de ning the pure natural direct effect in this way, is that it leads to a natural indirect effect.
- The total natural indirect effect of X on Y, conditional on $C = c$, as a mean difference comparing x vs x is

$$\mathrm{TNIE}(c; x; x) = E[Y\{x, M(x)\} \mid C = c] - E[Y\{x, M(x)\} \mid C = c].$$

- This is a comparison of two hypothetical worlds.
- In the rst, M is set to $M(x)$, the value it would take if X were set to x and in the second M is set to $M(x)$, the value it would take if X were set to x. In both worlds, X is set to x.
- X is allowed to in uence Y only through its in uence on M. Thus it is an indirect effect through M.

Now we see that the sum of the pure natural direct effect and total natural indirect effect is

$$
\begin{aligned}
&\mathrm{PNDE}(c;x;x\ ) + \mathrm{TNIE}(c;x;x\ ) \\
&\quad = E[Y\{x\ ;M(x)\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\quad + E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x\ ;M(x)\}|C=c] \\
&\quad = E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \mathrm{TCE}(c;x;x\ );
\end{aligned}
$$

the total causal effect, as desired.

Now we see that the sum of the pure natural direct effect and total natural indirect effect is

$$
\begin{aligned}
&\mathrm{PNDE}(c;x;x\ ) + \mathrm{TNIE}(c;x;x\ ) \\
&\quad = E[Y\{x\ ;M(x)\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\quad + E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x\ ;M(x)\}|C=c] \\
&\quad = E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \mathrm{TCE}(c;x;x\ );
\end{aligned}
$$

the total causal effect, as desired.

Now we see that the sum of the pure natural direct effect and total natural indirect effect is

$$
\begin{aligned}
&\mathrm{PNDE}(c;x;x\ ) + \mathrm{TNIE}(c;x;x\ ) \\
&\quad = E[Y\{x\ ;M(x)\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\quad + E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x\ ;M(x)\}|C=c] \\
&\quad = E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \mathrm{TCE}(c;x;x\ );
\end{aligned}
$$

the total causal effect, as desired.

In our de nition of the PNDE, we controlled M at $M(x)$:

$$\mathrm{PNDE}(c; x; x\ ) = E[Y\{x\ ; M(x)\} \mid C = c] - E[Y\{x; M(x)\} \mid C = c]$$

Why not $M(x\ )$?

In our de nition of the PNDE, we controlled M at $M(x)$:

$$\mathrm{PNDE}(c;x;x\ ) = E[Y\{x\ ;M(x)\}|C=c] \quad E[Y\{x;M(x)\}|C=c]$$

Why not $M(x\ )$?

And in our de nition of the TNIE, we controlled X at $x\ :$

$$\mathrm{TNIE}(c;x;x\ ) = E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x\ ;M(x)\}|C=c]$$

Why not x?

We could instead have de ned a total natural direct effect:

$$\mathrm{TNDE}(c; x; x\ ) = E[Y\ f x\ ; M(x\ )g\ jC = c] \qquad E[Y\ f x; M(x\ )g\ jC = c]$$

We could instead have de ned a total natural direct effect:

$$\mathrm{TNDE}(c; x; x\ ) = E[Y\ f\, x\ ; M(x\ )g\, jC = c] \qquad E[Y\ f\, x; M(x\ )g\, jC = c]$$

and a pure natural indirect effect:

$$\mathrm{PNIE}(c; x; x\ ) = E[Y\ f\, x; M(x\ )g\, jC = c] \qquad E[Y\ f\, x; M(x)g\, jC = c]$$

The TCE also decomposes into the TNDE+PNIE:

$$
\begin{aligned}
&\mathrm{TNDE}(c;x;x\ ) + \mathrm{PNIE}(c;x;x\ ) \\
&\quad = E[Y\ f\,x\ ;M(x\ )g\,jC = c] \quad E[Y\ f\,x;M(x\ )g\,jC = c] \\
&\quad\ + E[Y\ f\,x;M(x\ )g\,jC = c] \quad E[Y\ f\,x;M(x)g\,jC = c] \\
&\quad = E[Y\ f\,x\ ;M(x\ )g\,jC = c] \quad E[Y\ f\,x;M(x)g\,jC = c] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \mathrm{TCE}(c;x;x\ ):
\end{aligned}
$$

The TCE also decomposes into the TNDE+PNIE:

$$
\begin{aligned}
&\mathrm{TNDE}(c;x;x\ ) + \mathrm{PNIE}(c;x;x\ ) \\
&\quad = E[Y\ f\,x\ ;M(x\ )g\,jC = c] \quad E[Y\ f\,x;M(x\ )g\,jC = c] \\
&\qquad + E[Y\ f\,x;M(x\ )g\,jC = c] \quad E[Y\ f\,x;M(x)g\,jC = c] \\
&\qquad = E[Y\ f\,x\ ;M(x\ )g\,jC = c] \quad E[Y\ f\,x;M(x)g\,jC = c] \\
&\qquad\qquad = \mathrm{TCE}(c;x;x\ ):
\end{aligned}
$$

The TCE also decomposes into the TNDE+PNIE:

$$
\begin{aligned}
&\mathrm{TNDE}(c;x;x\ ) + \mathrm{PNIE}(c;x;x\ ) \\
&\quad = E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x;M(x\ )\}|C=c] \\
&\qquad + E[Y\{x;M(x\ )\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\qquad = E[Y\{x\ ;M(x\ )\}|C=c] \quad E[Y\{x;M(x)\}|C=c] \\
&\qquad\qquad = \mathrm{TCE}(c;x;x\ ):
\end{aligned}
$$

VanderWeele (Epidemiology, in press) has recently suggested a three-way decomposition:

$$\mathrm{TCE}(c; x; x\ ) = \mathrm{PNDE}(c; x; x\ ) + \mathrm{PNIE}(c; x; x\ ) + \mathrm{Remainder}(c; x; x\ )$$

VanderWeele (Epidemiology, in press) has recently suggested a three-way decomposition:

$$\mathrm{TCE}(c; x; x\ ) = \mathrm{PNDE}(c; x; x\ ) + \mathrm{PNIE}(c; x; x\ ) + \mathrm{Remainder}(c; x; x\ )$$

where, for binary M and X (so $x = 0$, $x\ = 1$), he shows that

$$\begin{aligned}\mathrm{Remainder}(c; 0; 1) = {} & E[\{M(1) \quad M(0)\} \\ & \{Y(1; 1) \quad Y(1; 0) \quad Y(0; 1) + Y(0; 0)\} \mid C = c]\end{aligned}$$

VanderWeele (Epidemiology, in press) has recently suggested a three-way decomposition:

$$\mathrm{TCE}(c; x; x\ ) = \mathrm{PNDE}(c; x; x\ ) + \mathrm{PNIE}(c; x; x\ ) + \mathrm{Remainder}(c; x; x\ )$$

where, for binary M and X (so $x = 0$, $x\ = 1$), he shows that

$$\mathrm{Remainder}(c; 0; 1) = E[\{M(1) \quad M(0)\} \{Y(1; 1) \quad Y(1; 0) \quad Y(0; 1) + Y(0; 0)\} \mid C = c]$$

This is non-zero only if there is mediation and interaction; thus he calls the remainder term a mediated interaction.

VanderWeele (Epidemiology, in press) has recently suggested a three-way decomposition:

$$\mathrm{TCE}(c; x; x\ ) = \mathrm{PNDE}(c; x; x\ ) + \mathrm{PNIE}(c; x; x\ ) + \mathrm{Remainder}(c; x; x\ )$$

where, for binary M and X (so $x = 0$, $x\ = 1$), he shows that

$$\mathrm{Remainder}(c; 0; 1) = E[\{M(1) \quad M(0)\} \{Y(1; 1) \quad Y(1; 0) \quad Y(0; 1) + Y(0; 0)\} \mid C = c]$$

This is non-zero only if there is mediation and interaction; thus he calls the remainder term a mediated interaction.

The total natural direct/indirect effects include the mediated interaction; the pure effects exclude it.

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

- Given clear de nitions of the estimands we would like to estimate, we can give assumptions under which they can be identi ed from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.

- Given clear de nitions of the estimands we would like to estimate, we can give assumptions under which they can be identi ed from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three avours:

- Given clear de nitions of the estimands we would like to estimate, we can give assumptions under which they can be identi ed from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three avours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x; m)$ with the actual outcome Y, for certain subjects.

- Given clear de nitions of the estimands we would like to estimate, we can give assumptions under which they can be identi ed from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three avours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x; m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identi ed by consistency, can be estimated by borrowing information across subjects.

- Given clear de nitions of the estimands we would like to estimate, we can give assumptions under which they can be identi ed from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three avours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x;m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identi ed by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more ef ciently.

- Consistency for $Y(x)$:

$$Y = Y(x) \text{ if } X = x$$

- Consistency for $Y(x)$:

$$Y = Y(x) \text{ if } X = x$$

- Conditional exchangeability given C for X wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

- Consistency for $Y(x)$:

$$Y = Y(x) \text{ if } X = x$$

- Conditional exchangeability given C for X wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

- Essentially, this means no unmeasured confounding of the X–Y relationship.

- Consistency for $Y(x;m)$:

$$Y = Y(x;m) \text{ if } X = x \text{ and } M = m$$

- Consistency for $Y(x;m)$:

$$Y = Y(x;m) \text{ if } X = x \text{ and } M = m$$

- Sequential conditional exchangeability given C for X wrt Y and given $C;X;L$ for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

$$Y(x;m) \perp\!\!\!\perp M \mid C;X;L \quad \forall x;m$$

- Consistency for $Y(x;m)$:

$$Y = Y(x;m) \text{ if } X = x \text{ and } M = m$$

- Sequential conditional exchangeability given C for X wrt Y and given $C;X;L$ for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

$$Y(x;m) \perp\!\!\!\perp M \mid C;X;L \quad \forall x;m$$

- Essentially, this means no unmeasured confounding of the X–Y or M–Y relationships.

- Consistency for $Y(x;m)$:

$$Y = Y(x;m) \text{ if } X = x \text{ and } M = m$$

- Sequential conditional exchangeability given C for X wrt Y and given $C;X;L$ for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

$$Y(x;m) \perp\!\!\!\perp M \mid C;X;L \quad \forall x;m$$

- Essentially, this means no unmeasured confounding of the X–Y or M–Y relationships.

- Consistency for $Y(x;m)$, $M(x)$ and $Y\{x;M(x^*)\}$:

$$Y = Y(x;m) \text{ if } X = x \text{ and } M = m; \quad M = M(x) \text{ if } X = x;$$

$$Y = Y\{x;M(x^*)\} \text{ if } X = x \text{ and } M = M(x^*):$$

- Consistency for $Y(x;m)$, $M(x)$ and $Y\{x;M(x)\}$:

$$Y = Y(x;m) \text{ if } X = x \text{ and } M = m; \quad M = M(x) \text{ if } X = x;$$

$$Y = Y\{x;M(x)\} \text{ if } X = x \text{ and } M = M(x):$$

- Sequential conditional exchangeability given C for X wrt Y, given $C;X;L$ for M wrt Y, and given C for X wrt M:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x; \quad Y(x;m) \perp\!\!\!\perp M \mid C;X;L \quad \forall x;m;$$

$$M(x) \perp\!\!\!\perp X \mid C \quad \forall x:$$

- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.

- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.

- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.

- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:

- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or

- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, 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).

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

$$\mathrm{CDE}(c;x;x^{*};m) =$$

$$E\{Y(x^{*};m)|C=c\} - E\{Y(x;m)|C=c\}$$

$$= \int E(Y|C=c;X=x^{*};L=l;M=m)\, f_{L|C;X}(l|c;x^{*})\, dl$$

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

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

$$\begin{aligned}
&\mathrm{CDE}(c;x;x^{*};m) = \\
&\quad E\{Y(x^{*};m)\,|\,C=c\} - E\{Y(x;m)\,|\,C=c\} \\
&= \int E(Y\,|\,C=c;X=x^{*};L=l;M=m)\,f_{L|C;X}(l\,|\,c;x^{*})\,dl \\
&\quad - \int E(Y\,|\,C=c;X=x;L=l;M=m)\,f_{L|C;X}(l\,|\,c;x)\,dl
\end{aligned}$$

- This is the g-computation formula.

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

$$\mathrm{CDE}(c;x;x^{*};m) =$$

$$E\{Y(x^{*};m)|C=c\} - E\{Y(x;m)|C=c\}$$

$$= \int E(Y|C=c;X=x^{*};L=l;M=m)\, f_{L|C;X}(l|c;x^{*})\,dl$$

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

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

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

$$\mathrm{CDE}(c;x;x^{*};m) =$$

$$E\{Y(x^{*};m)|C=c\} - E\{Y(x;m)|C=c\}$$

$$= \int E(Y|C=c;X=x^{*};L=l;M=m)\, f_{L|C;X}(l|c;x^{*})\, dl$$

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

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

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

$$\mathrm{CDE}(c;x;x^{*};m) =$$

$$E\{Y(x^{*};m)|C=c\} - E\{Y(x;m)|C=c\}$$

$$= \int E(Y|C=c;X=x^{*};L=l;M=m)\, f_{L|C;X}(l|c;x^{*})\, dl$$

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

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

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

$$\mathrm{CDE}(c;x;x\ ;m) =$$

$$\left\{ E\{Y(x\ ;m)|C=c\} - E\{Y(x;m)|C=c\} \right\}$$

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

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

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

- In the absence of intermediate confounding, non-linearities and interactions, this becomes $_1$ as earlier.

- In the absence of intermediate confounding, non-linearities and interactions, this becomes $_1$ as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.

- In the absence of intermediate confounding, non-linearities and interactions, this becomes $_1$ as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.

- The g-computation formula can similarly be used to estimate the PNDE, TNDE, PNIE and TNIE, with further modelling and assumptions.

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

- The g-computation formula can similarly be used to estimate the PNDE, TNDE, PNIE and TNIE, 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:

$$
\begin{aligned}
&E\{Y(x^*; m) \quad Y(x; m) \mid C = c; M(x) = m\} \\
&\quad = E\{Y(x^*; m) \quad Y(x; m) \mid C = c; M(x^*) = m\} \\
&\qquad = E\{Y(x^*; m) \quad Y(x; m) \mid C = c\}:
\end{aligned}
$$

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

- 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.

- 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 $L \mid C; X$:

- 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 $L \mid C; X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),

- 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 $L \mid C; X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),

- 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 $L \mid C; X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other avours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

1 Mediation analysis with one mediator
- Traditional approach
- Limitations of the traditional approach
- Contributions from causal inference

2 Mediation analysis with multiple mediators

3 Back to the motivating example

4 Summary

- For two mediators, we need the following counterfactuals: $M_1(x)$, $M_2(x; m_1)$, $Y(x; m_1; m_2)$, de ned in the obvious way.

- For two mediators, we need the following counterfactuals: $M_1(x)$, $M_2(x; m_1)$, $Y(x; m_1; m_2)$, de ned in the obvious way.
- And now the complex composite counterfactual

$$Y[x_a; M_1(x_b); M_2 f x_c; M_1(x_d) g]:$$

- For two mediators, we need the following counterfactuals: $M_1(x)$, $M_2(x; m_1)$, $Y(x; m_1; m_2)$, de ned in the obvious way.
- And now the complex composite counterfactual

$$Y[x_a; M_1(x_b); M_2 f x_c; M_1(x_d) g]:$$

We can de ne appropriate path-speci c effects by changing these arguments.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-}x_b x_c x_d = E\,(Y\,[1; M_1\,(x_b)\,; M_2\,f\,x_c; M_1\,(x_d)g]$$
$$Y\,[0; M_1\,(x_b)\,; M_2\,f\,x_c; M_1\,(x_d)g])$$

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-}x_b x_c x_d = E\,(Y\,[1; M_1\,(x_b)\,; M_2\,f\,x_c; M_1\,(x_d)g]$$
$$Y\,[0; M_1\,(x_b)\,; M_2\,f\,x_c; M_1\,(x_d)g])$$

This is what makes it a direct effect.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-}x_b x_c x_d = E\,(Y\,[1; M_1\,(x_b)\,; M_2\,f\,x_c; M_1\,(x_d)g] \\ Y\,[0; M_1\,(x_b)\,; M_2\,f\,x_c; M_1\,(x_d)g])$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-000} = E(Y[1; M_1(0); M_2\{0; M_1(0)\}] \\ Y[0; M_1(0); M_2\{0; M_1(0)\}])$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-001} = E(Y[1; M_1(0); M_2\{0; M_1(1)\}] \quad Y[0; M_1(0); M_2\{0; M_1(1)\}])$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-010} = E\left(Y\left[1; M_1(0); M_2\{1; M_1(0)\}\right] - Y\left[0; M_1(0); M_2\{1; M_1(0)\}\right]\right)$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-011} = E\left(Y\left[1; M_1(0); M_2\{1; M_1(1)\}\right] - Y\left[0; M_1(0); M_2\{1; M_1(1)\}\right]\right)$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-100} = E(Y[1; M_1(1); M_2\{0; M_1(0)\}] \quad Y[0; M_1(1); M_2\{0; M_1(0)\}])$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-101} = E(Y[1; M_1(1); M_2\{0; M_1(1)\}] \quad Y[0; M_1(1); M_2\{0; M_1(1)\}])$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-110} = E(Y[1; M_1(1); M_2\{1; M_1(0)\}] \quad Y[0; M_1(1); M_2\{1; M_1(0)\}])$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

For simplicity, let's drop conditioning on $C = c$ (it is implied everywhere) and let's consider binary X so that $x = 0$ and $x = 1$.

Effects through neither M_1 nor M_2 (direct effect):

$$\text{NDE-}111 = E\left(Y\left[1; M_1(1); M_2\{1; M_1(1)\}\right] - Y\left[0; M_1(1); M_2\{1; M_1(1)\}\right]\right)$$

This is what makes it a direct effect.

And these can be set at any values, so 8 choices.

Similarly, path-speci c effects through M_1 only:

$$\mathrm{NPSE}_1\text{-}x_a x_c x_d = E\,(Y\,[x_a; M_1\,(1)\,M_2\,f\,x_c; M_1\,(x_d)g]$$
$$Y\,[x_a; M_1\,(0)\,M_2\,f\,x_c; M_1\,(x_d)g])$$

Similarly, path-speci c effects through M_1 only:

$$\mathrm{NPSE}_1\text{-}x_a x_c x_d = E\,(Y\,[x_a; M_1\,(1)\,M_2\,f\,x_c; M_1\,(x_d)g]$$
$$Y\,[x_a; M_1\,(0)\,M_2\,f\,x_c; M_1\,(x_d)g])$$

This is what makes it a path-speci c effect through M_1 only.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}x_a x_c x_d = E\,(Y\,[x_a; M_1\,(1)\,M_2\,f\,x_c; M_1\,(x_d)g]$$
$$Y\,[x_a; M_1\,(0)\,M_2\,f\,x_c; M_1\,(x_d)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}000 = E\,(Y\,[0; M_1(1)\,M_2\,f\,0; M_1(0)g]$$
$$Y\,[0; M_1(0)\,M_2\,f\,0; M_1(0)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}001 = E\,(Y\,[0; M_1(1)\,M_2\,f\,0; M_1(1)g]$$
$$Y\,[0; M_1(0)\,M_2\,f\,0; M_1(1)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}010 = E\,(Y\,[0; M_1(1)\,M_2\,f\,1; M_1(0)g]$$
$$Y\,[0; M_1(0)\,M_2\,f\,1; M_1(0)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}011 = E\,(Y\,[0; M_1(1)\,M_2\,f\,1; M_1(1)g]$$
$$Y\,[0; M_1(0)\,M_2\,f\,1; M_1(1)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}100 = E\,(Y\,[1; M_1(1)\,M_2\,f\,0; M_1(0)g]$$
$$Y\,[1; M_1(0)\,M_2\,f\,0; M_1(0)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}101 = E\,(Y\,[1; M_1(1)\,M_2\,f\,0; M_1(1)g]$$
$$Y\,[1; M_1(0)\,M_2\,f\,0; M_1(1)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}110 = E\,(Y\,[1; M_1(1)\,M_2\,f\,1; M_1(0)g]$$
$$Y\,[1; M_1(0)\,M_2\,f\,1; M_1(0)g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_1 only:

$$\text{NPSE}_1\text{-}111 = E\,(Y\,[1; M_1(1)\, M_2\, f\, 1; M_1(1) g] \\ Y\,[1; M_1(0)\, M_2\, f\, 1; M_1(1) g])$$

This is what makes it a path-speci c effect through M_1 only.

And these can again be set at any of 8 combinations.

Similarly, path-speci c effects through M_2 only:

$$\mathrm{NPSE}_2\text{-}x_a x_b x_d = E\,(Y\,[x_a; M_1\,(x_b)\,M_2\,f\,1; M_1\,(x_d)g]$$
$$Y\,[x_a; M_1\,(x_b)\,M_2\,f\,0; M_1\,(x_d)g])$$

Similarly, path-speci c effects through M_2 only:

$$\text{NPSE}_2\text{-}x_a x_b x_d = E\,(Y\,[x_a; M_1\,(x_b)\,M_2\,f\,1; M_1\,(x_d)g]$$
$$Y\,[x_a; M_1\,(x_b)\,M_2\,f\,0; M_1\,(x_d)g])$$

This is what makes it a path-speci c effect through M_2 only.