Counterfactual-based mediation analysis
Workshop 1

Rhian Daniel
London School of Hygiene and Tropical Medicine

CIMPOD
27th February, 2017

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Outline

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

1. Setting the scene
   - Introduction
   - Traditional approach
   - Causal inference gets involved
   - —Estimands
   - —Assumptions
   - —Identification
   - Interventional effects
2. Case study
3. Q&A
4. Wrapping up
5. References

Outline

1. Setting the scene
   - Introduction
   - Traditional approach
   - Causal inference gets involved
   - —Estimands
   - —Assumptions
   - —Identification
   - Interventional effects
2. Case study
3. Q&A
4. Wrapping up
5. References

Outline

Mediation

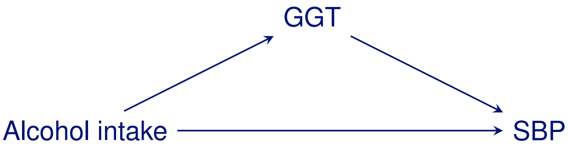
- For nearly a century, statisticians, and researchers in many different substantive disciplines, have been attempting to address questions concerning mediation.

Mediation

- For nearly a century, statisticians, and researchers in many different substantive disciplines, have been attempting to address questions concerning mediation.
  [Wright 1921, 1934; Baron and Kenny 1986; Robins and Greenland 1992; Pearl 2001; Cole and Hernán 2002; VanderWeele and Vansteelandt 2009; VanderWeele 2015.]

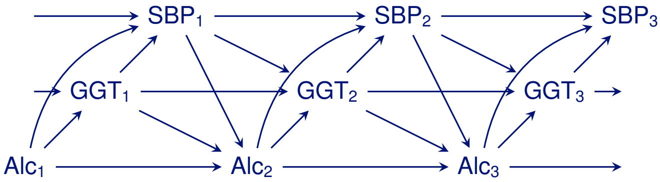
Mediation

- For nearly a century, statisticians, and researchers in many different substantive disciplines, have been attempting to address questions concerning mediation.
  [Wright 1921, 1934; Baron and Kenny 1986; Robins and Greenland 1992; Pearl 2001; Cole and Hernán 2002; VanderWeele and Vansteelandt 2009; VanderWeele 2015.]

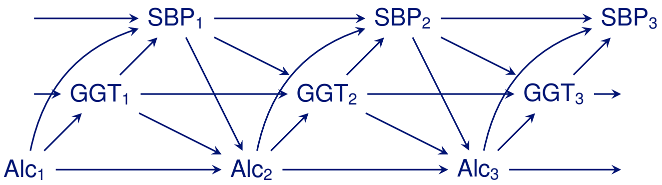


- For example (today's case study), how much of the effect of alcohol consumption on systolic blood pressure is via GGT (gamma-glutamyl transpeptidase), a blood enzyme?

(Of course, things are rarely this simple. . .)



(Of course, things are rarely this simple. . .)



Aalen OO, RK, Gran JM, Kouyos R, Lange T (2014)
Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms
SMMR, 25(5):2294–314.

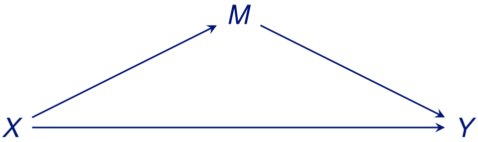
VanderWeele TJ, Tchetgen Tchetgen EJ
Mediation analysis with time-varying exposures and mediators
JRSS B, in press.

Outline

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Traditional approach

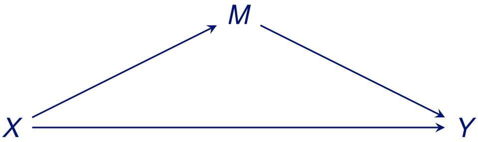
Path tracing rules [Wright 1934]



- Originally, mediation analysis was only attempted using linear models.

Traditional approach

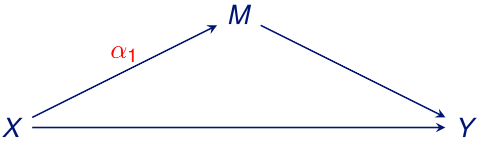
Path tracing rules [Wright 1934]



- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

Traditional approach

Path tracing rules [Wright 1934]

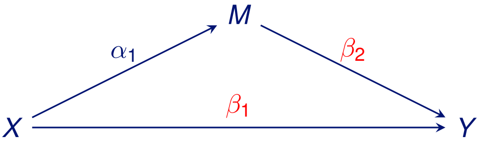


- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$

Traditional approach

Path tracing rules [Wright 1934]

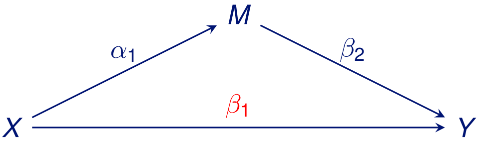


- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$
$$E(Y|X, M) = \beta_0 + \beta_1 X + \beta_2 M$$

Traditional approach

Path tracing rules [Wright 1934]



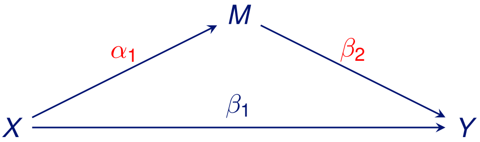
- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$
$$E(Y|X, M) = \beta_0 + \beta_1 X + \beta_2 M$$

- β_1 would then be labelled the direct effect.

Traditional approach

Path tracing rules [Wright 1934]



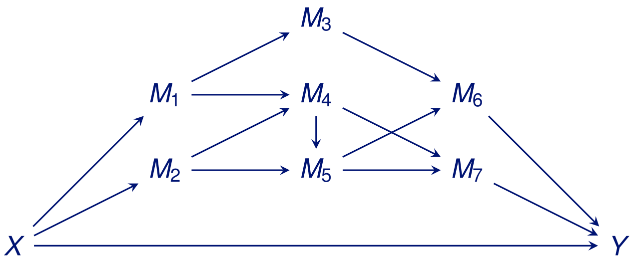
- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$
$$E(Y|X,M) = \beta_0 + \beta_1 X + \beta_2 M$$

- β_1 would then be labelled the direct effect.
- And $\alpha_1\beta_2$ the indirect effect.

More complex diagrams

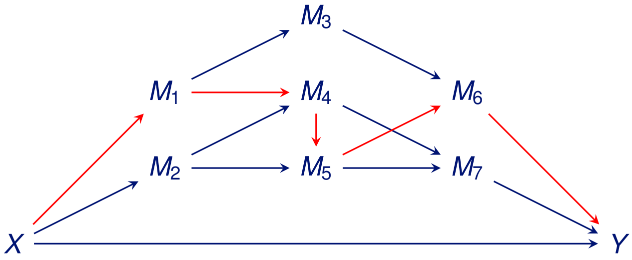
Path tracing rules [Wright 1934]



- This simple method extends to arbitrarily complex diagrams, as long as all models are simple linear regressions (with no interaction terms).

More complex diagrams

Path tracing rules [Wright 1934]



- This simple method extends to arbitrarily complex diagrams, as long as all models are simple linear regressions (with no interaction terms).
- The path-specific effect along a particular pathway is equal to the product of the coefficients along that path.

Outline

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Causal inference 'investigates'

- In the early 1990s, the 'causal inference' school became interested in this area [Robins and Greenland 1992].

Causal inference 'investigates'

- In the early 1990s, the 'causal inference' school became interested in this area [Robins and Greenland 1992].
- Mediation is a causal concept: associations are symmetric, but mediation implies an ordered sequence.

Causal inference 'investigates'

- In the early 1990s, the 'causal inference' school became interested in this area [Robins and Greenland 1992].
- Mediation is a causal concept: associations are symmetric, but mediation implies an ordered sequence.
- Core principles of causal inference: (1) what is the estimand? (2) under what assumptions can it be identified? (3) are there more flexible estimation methods than currently used?

Potential outcomes and mediators

- 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.

Potential outcomes and mediators

- 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.

Potential outcomes and mediators

- 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.

Potential outcomes and mediators

- 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.

Potential outcomes and mediators

- 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 hypothetical quantities were used to create model-free definitions of direct/indirect effects that match our intuition.

Outline

Controlled direct effect

Pearl, 2001

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

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

Controlled direct effect

Pearl, 2001

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

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

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

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y when M is controlled at m, expressed as a marginal 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 (or more) hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both situations, 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 marginal 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 (or more) hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both situations, M is set to m.
- By keeping M fixed at m, we are getting at a direct effect of X, unmediated by M.

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y when M is controlled at m, expressed as a marginal 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 (or more) hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both situations, M is set to m.
- By keeping M fixed at m, we are getting at a direct effect of X, unmediated by M.
- In our example, it is the change in mean SBP if everyone vs noone drinks, with everyone having their GGT fixed to a common value, m.

Natural direct effect

Pearl 2001; Robins and Greenland 1992

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

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

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of X on Y expressed as a marginal 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 again a comparison of two hypothetical situations.

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of X on Y expressed as a marginal 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 again a comparison of two hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both, M is set to $M(0)$, its value 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 marginal 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 again a comparison of two hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both, M is set to $M(0)$, its value if X were set to 0.
- Since M is the same (*within* subject) in both situations, we are also intuitively getting at a direct effect of X.

Natural direct effect

Pearl 2001; Robins and Greenland 1992

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

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

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

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of X on Y expressed as a marginal 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 again a comparison of two hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both, M is set to $M\left(0\right)$, its value if X were set to 0.
- Since M is the same (*within* subject) in both situations, we are also intuitively getting at a direct effect of X.
- If no individual-level interaction between X and M, $\text{CDE}\left(m\right) = \text{NDE} \;\; \forall m$.
- It is the change in mean SBP if everyone vs noone drinks, with each individual's GGT fixed at what it would have been for that person under no drinking.

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- 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 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 situations.

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- The natural indirect effect of X on Y is

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

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

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- 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 situations.
- In the first, M is set to $M(1)$ and in the second M is set to $M(0)$. In both, X is set to 1.
- X is allowed to influence Y only through its influence on M. Thus it intuitively corresponds to an indirect effect through M.

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- 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 situations.
- In the first, M is set to $M(1)$ and in the second M is set to $M(0)$. In both, X is set to 1.
- X is allowed to influence Y only through its influence on M. Thus it intuitively corresponds to an indirect effect through M.
- It is the change in mean SBP we would see if we changed everyone's GGT from its non-drinking level to its drinking level, whilst fixing the exposure to 'drinking'.

Effect decomposition

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 of X on Y.

Effect decomposition

The sum of the natural direct and indirect effects is

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

the total causal effect of X on Y.

Effect decomposition

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 of X on Y.

Effect decomposition

The sum of the natural direct and indirect effects is

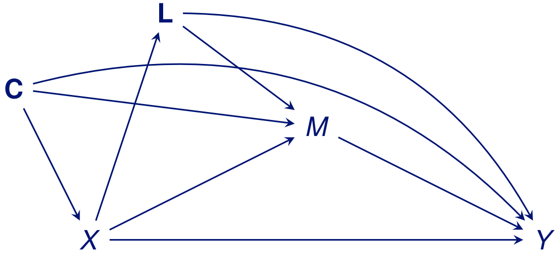
$$\begin{aligned} \text{NDE} + \text{NIE} = & \; E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] \\ & + E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right] \\ & = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] = \text{TCE},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

the total causal effect of X on Y.

Note that such a sensible decomposition is not possible using the CDE.

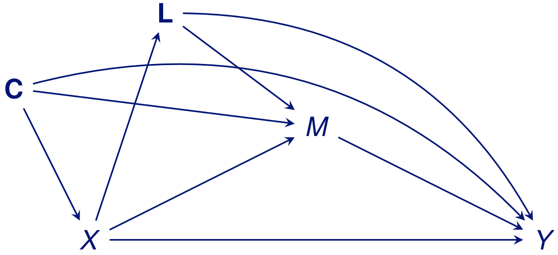
Outline

Assumptions for identification (1)



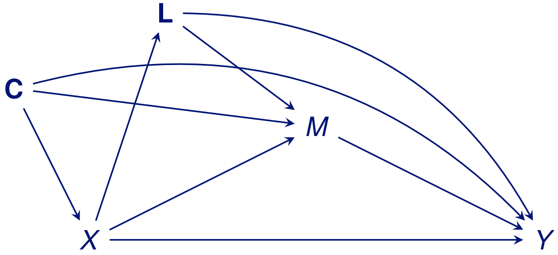
- Consider the setting with baseline confounders **C** and intermediate confounders **L**.

Assumptions for identification (1)



- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.

Assumptions for identification (1)

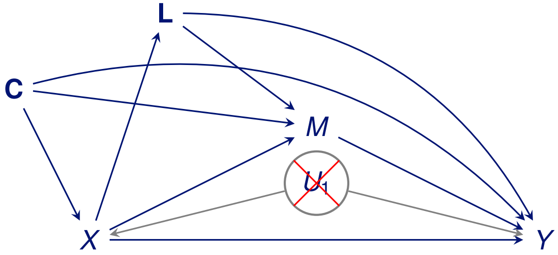


- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.
- Then there are sequential conditional exchangeability assumptions:

$$Y(x,m) \perp\!\!\!\perp X \mid \mathbf{C}=\mathbf{c} \ , \forall x, m, \mathbf{c}$$

$$Y(x,m) \perp\!\!\!\perp M \mid \mathbf{C}=\mathbf{c}, X=x, \mathbf{L}=\mathbf{l} \ , \forall x, m, \mathbf{c}, \mathbf{l}$$

Assumptions for identification (1)

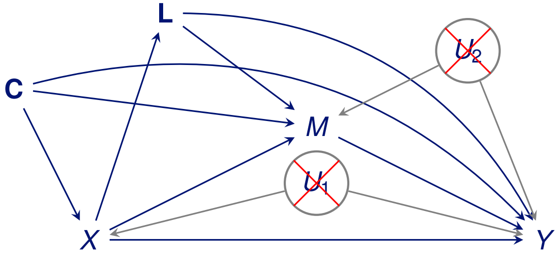


- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.
- Then there are sequential conditional exchangeability assumptions:

$$Y(x, m) \perp\!\!\!\perp X \,|\mathbf{C} = \mathbf{c} \ , \forall x, m, \mathbf{c}$$

$$Y(x, m) \perp\!\!\!\perp M \,|\mathbf{C} = \mathbf{c}, X = x, \mathbf{L} = \mathbf{l} \ , \forall x, m, \mathbf{c}, \mathbf{l}$$

Assumptions for identification (1)

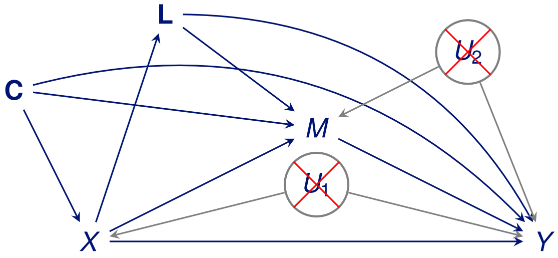


- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.
- Then there are sequential conditional exchangeability assumptions:

$$Y(x,m) \perp\!\!\!\perp X \,|\, \mathbf{C}=\mathbf{c} \;, \forall x, m, \mathbf{c}$$

$$Y(x,m) \perp\!\!\!\perp M \,|\, \mathbf{C}=\mathbf{c}, X=x, \mathbf{L}=\mathbf{l} \;, \forall x, m, \mathbf{c}, \mathbf{l}$$

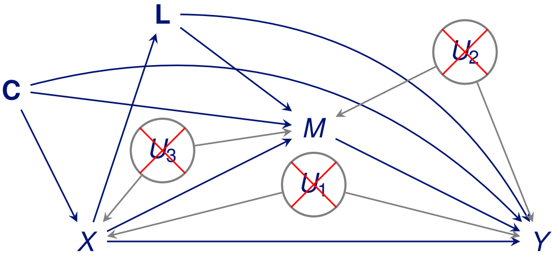
Assumptions for identification (2)



- And:

$$M(x) \perp\!\!\!\perp X \,|\, \mathbf{C} = \mathbf{c} \;, \forall x, \mathbf{c}$$

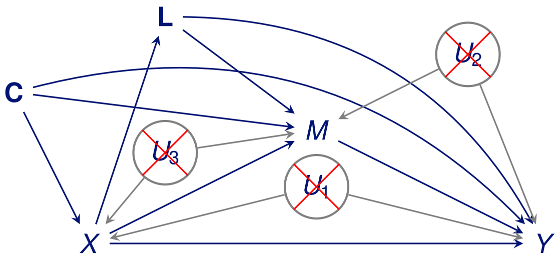
Assumptions for identification (2)



- And:

$$M(x) \perp\!\!\!\perp X \,|\, \mathbf{C} = \mathbf{c} \;,\; \forall x, \mathbf{c}$$

Assumptions for identification (2)

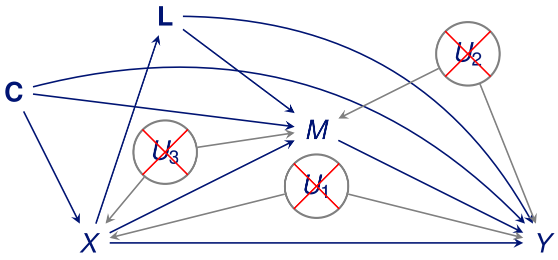


- And:

$$M(x) \perp\!\!\!\perp X \,|\, \mathbf{C} = \mathbf{c} \;,\; \forall x, \mathbf{c}$$

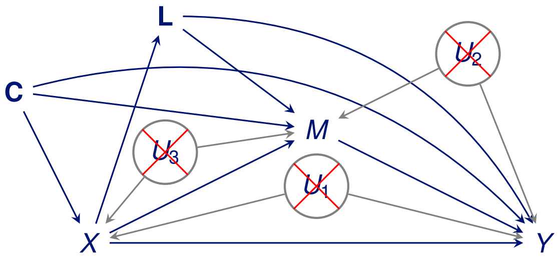
This much, we would probably expect!

Assumptions for identification (3)



- Perhaps surprisingly, these assumptions (although sufficient for the CDE) are not enough for NDE/NIE.

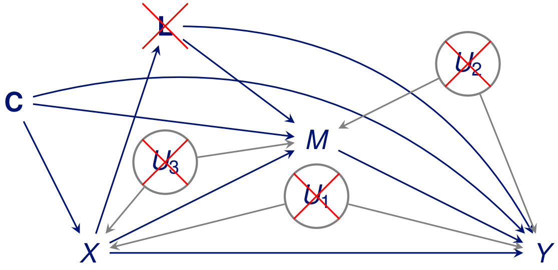
Assumptions for identification (3)



- Perhaps surprisingly, these assumptions (although sufficient for the CDE) are not enough for NDE/NIE.
- In addition, we need something such as the cross-world independence assumption:

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\, \mathbf{C} = \mathbf{c} \;, \forall x, m, x^*, \mathbf{c}$$

Assumptions for identification (3)



- Perhaps surprisingly, these assumptions (although sufficient for the CDE) are not enough for NDE/NIE.
- In addition, we need something such as the cross-world independence assumption:

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\, \mathbf{C} = \mathbf{c} \;,\; \forall x, m, x^*, \mathbf{c}$$

- This implies (but is not implied by, ie it is stronger than) no **L**.

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\mathbf{C} = \mathbf{c} \;, \forall x, m, x^*, \mathbf{c}$$

rules out intermediate confounders **L**.

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\, \mathbf{C} = \mathbf{c} \;, \forall x, m, x^*, \mathbf{c}$$

  rules out intermediate confounders **L**.

- In fact, a slightly weaker assumption, which does not rule out **L** is sufficient:

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

  [Petersen et al 2006]

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m)\,|\mathbf{C}=\mathbf{c}\ ,\ \forall x,m,x^*,\mathbf{c}$$

  rules out intermediate confounders **L**.

- In fact, a slightly weaker assumption, which does not rule out **L** is sufficient:

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

  [Petersen et al 2006]

- Both assumptions are very strong, and not even a hypothetical experiment exists in which they would hold by design.
  [Richardson and Robins 2013]

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m)\,|\mathbf{C}=\mathbf{c}\,,\,\forall x,m,x^*,\mathbf{c}$$

  rules out intermediate confounders **L**.
- In fact, a slightly weaker assumption, which does not rule out **L** is sufficient:

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

  [Petersen et al 2006]
- Both assumptions are very strong, and not even a hypothetical experiment exists in which they would hold by design. [Richardson and Robins 2013]
- Even the Petersen assumption places strong parametric restrictions on the relationship between **L** and Y, which can essentially only hold in linear models with no non-linearities involving **L**. [De Stavola et al 2015]

Outline

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.
- First we write:

$$E[Y\{x, M(x^*)\}] =$$

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}, M(x^*)=m\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.
- First we write:

$$E[Y\{x, M(x^*)\}] =$$

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}, M(x^*)=m\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By the cross-world independence assumption, this is equal to:

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.
- First we write:

$$E[Y\{x, M(x^*)\}] =$$
$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}, M(x^*)=m\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By the cross-world independence assumption, this is equal to:

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By conditional exchangeability, this is:

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- The hard work is now done.

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x,M=m,\mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*,\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x,M=m,\mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*,\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- The hard work is now done.
- By substituting different values for x and x^*, we can re-write the NDE and the NIE using only functions of aspects of the distribution of the observed data.

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- The hard work is now done.
- By substituting different values for x and x^*, we can re-write the NDE and the NIE using only functions of aspects of the distribution of the observed data.
- Plug-in or alternative (semiparametric) estimation could then be used. Many many proposals have been made!

Summary so far

- Mediation analysis is not new.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.

Summary so far

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving unambiguous counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.

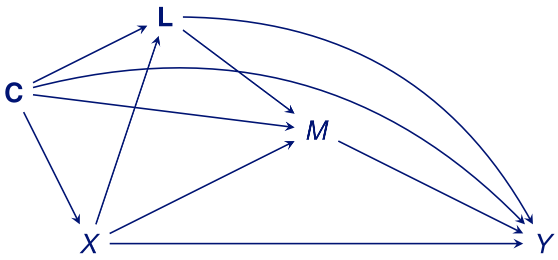
Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving unambiguous counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.
- The identification expressions can be used to estimate direct and indirect effects in the presence of non-linearities, and thus have greatly increased the flexibility of mediation analysis.

Summary so far

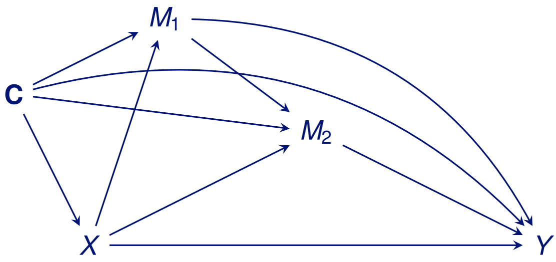
- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving unambiguous counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.
- The identification expressions can be used to estimate direct and indirect effects in the presence of non-linearities, and thus have greatly increased the flexibility of mediation analysis.
- However, it is plagued by the strength of the cross-world/Petersen assumptions; in particular, the fact that these assumptions almost rules out intermediate confounding even when measured.

Consequences for multiple mediators



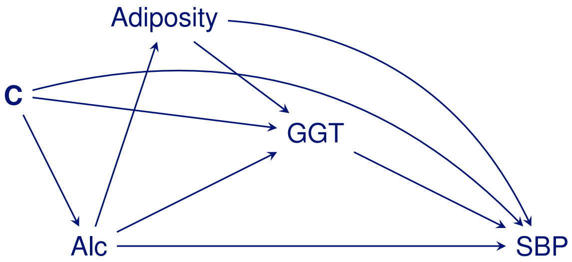
- For the same reason that in general we can't have **L** . . .

Consequences for multiple mediators



- For the same reason that in general we can't have **L** . . .
- . . . settings involving multiple mediators are also problematic.

Consequences for multiple mediators



- For the same reason that in general we can't have **L** . . .
- . . . settings involving multiple mediators are also problematic.
- eg in our motivating example.

Outline

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on **C**.

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

  where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on **C**.

- The randomised interventional analogue of the NIE of X on Y expressed as a marginal mean difference is

$$\text{RIA-NIE} = E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\}.$$

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

  where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on **C**.

- The randomised interventional analogue of the NIE of X on Y expressed as a marginal mean difference is

$$\text{RIA-NIE} = E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\}.$$

- The RIA-NDE is the effect on the mean SBP of changing everyone's drinking status, whilst leaving each subject's GGT at a random draw from the distribution of GGT given that subject's background confounder levels, amongst the non-drinkers.

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on $\mathbf{C}$.

- The randomised interventional analogue of the NIE of X on Y expressed as a marginal mean difference is

$$\text{RIA-NIE} = E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\}.$$

- The RIA-NDE is the effect on the mean SBP of changing everyone's drinking status, whilst leaving each subject's GGT at a random draw from the distribution of GGT given that subject's background confounder levels, amongst the non-drinkers.
- The RIA-NIE is the effect on mean SBP of shifting the GGT distribution given confounders from that seen in non-drinkers to that seen in drinkers, whilst setting everyone's exposure to 'drinking'.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.
- If not, then the stronger **C** predicts M, the smaller the difference between NDE and RIA-NDE.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.
- If not, then the stronger **C** predicts M, the smaller the difference between NDE and RIA-NDE.
- RIA effects correspond to interventions that could in principle be done.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.
- If not, then the stronger **C** predicts M, the smaller the difference between NDE and RIA-NDE.
- RIA effects correspond to interventions that could in principle be done.
- However, RIA-NDE + RIA-NIE =

$$E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

which is **NOT** in general equal to the total causal effect!

Summary

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.

Summary

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.

Summary

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.
- However, this endeavour has been limited by the extremely strong and untestable cross-world assumption.

Summary

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.
- However, this endeavour has been limited by the extremely strong and untestable cross-world assumption.
- This has effectively prohibited flexible multiple mediation analyses, even though applied problems frequently involve multiple mediators.

Summary

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.
- However, this endeavour has been limited by the extremely strong and untestable cross-world assumption.
- This has effectively prohibited flexible multiple mediation analyses, even though applied problems frequently involve multiple mediators.
- Interventional effects are perhaps the way forward, since they don't require this cross-world assumption.

Outline

Data description

- We now turn to the case study.

Data description

- We now turn to the case study.
- The dataset for this case study (`Pseudo_Izhevsk.dta`) has been simulated to be similar to, but a simplified version of, the data from the Izhevsk Family Study.

Data description

- We now turn to the case study.
- The dataset for this case study (`Pseudo_Izhevsk.dta`) has been simulated to be similar to, but a simplified version of, the data from the Izhevsk Family Study.
- A case–control study to study the effects of extreme alcohol consumption on mortality in men in Izhevsk, Russia.

Data description

- We now turn to the case study.
- The dataset for this case study (`Pseudo_Izhevsk.dta`) has been simulated to be similar to, but a simplified version of, the data from the Izhevsk Family Study.
- A case–control study to study the effects of extreme alcohol consumption on mortality in men in Izhevsk, Russia.
- We'll analyse simulated data that mimic the population-based controls, and use these men to estimate the effect of drinking more than 10L of ethanol in the previous year on SBP, and the extent to which this effect is mediated by GGT.

Data description

- We now turn to the case study.
- The dataset for this case study (`Pseudo_Izhevsk.dta`) has been simulated to be similar to, but a simplified version of, the data from the Izhevsk Family Study.
- A case–control study to study the effects of extreme alcohol consumption on mortality in men in Izhevsk, Russia.
- We'll analyse simulated data that mimic the population-based controls, and use these men to estimate the effect of drinking more than 10L of ethanol in the previous year on SBP, and the extent to which this effect is mediated by GGT.
- Background confounders: age, SES, smoking status (never, ex, current). Intermediate confounder: BMI.

Data description

- We now turn to the case study.
- The dataset for this case study (`Pseudo_Izhevsk.dta`) has been simulated to be similar to, but a simplified version of, the data from the Izhevsk Family Study.
- A case–control study to study the effects of extreme alcohol consumption on mortality in men in Izhevsk, Russia.
- We'll analyse simulated data that mimic the population-based controls, and use these men to estimate the effect of drinking more than 10L of ethanol in the previous year on SBP, and the extent to which this effect is mediated by GGT.
- Background confounders: age, SES, smoking status (never, ex, current). Intermediate confounder: BMI.
- For simplicity for this workshop, we have dropped the variable containing the number of cigarettes smoked per day, and we haven't simulated any data to be missing (whereas in the paper, we used single stochastic imputation for the missing values).

Tasks

Question 1

Familiarise yourselves with the dataset and check the distribution of BMI and GGT. Might log transformations be sensible?

For help with Stata syntax, see `CaseStudy1_Q1.do`.

Question 2

Investigate, using traditional mediation analysis, the extent to which the effect of alcohol on SBP is mediated by GGT.

You should take into account the background confounders age, SES and smoking, but you should ignore BMI for now, since it is an intermediate confounder (we will come back to it in Question 4).

For help with Stata syntax, see `CaseStudy1_Q2.do`.

Tasks

Question 3

(a) Now repeat the same analysis using the `paramed` command in Stata.

You may need to start by installing `paramed`:

```
findit paramed
```

The syntax for continuous outcome `y`, continuous mediator `m`, binary exposure `x`, and background confounders `c1` and `c2`, with both models simple linear regression, is:

```
paramed y, avar(x) mvar(m) a0(0) a1(1)
m(3) yreg(linear) mreg(linear) cvars(c1 c2)
nointeraction
```

For more help with the Stata syntax, see `CaseStudy1_Q3.do`.

Tasks

Question 3 (cont'd)

(b) Now repeat the same analysis, but this time allowing there to be an exposure–mediator interaction. This can be done simply by removing `nointeraction` from the command in part (a).

Do you understand the output? Does the interaction seem important? Do you understand why the `nde` was not given in the output for part (a)?

For more help with the Stata syntax, see `CaseStudy1_Q3.do`.

Question 4

We now deal with BMI, the intermediate confounder (L).

You may want to consult `CaseStudy1_Q4.do` from the beginning.

Since things are getting a bit complex now, with 3 models, and since we wish to include interactions in some/all of these models, we proceed now by Monte Carlo simulation, rather than analytically.

The general idea is as follows:

Tasks

Question 4 (cont'd)

(1) Fit a model for `logBMI` given `alc`, `age`, `SES` and `smoke`.

(2) Simulate two values of `logBMI` for each individual: one had their exposure been 1, and one had their exposure been 0, i.e. $L(1)$ and $L(0)$. These simulations need to be stochastic, so remember to add `e(rmse)*rnormal()`.

(3) Do the same for `logGGT`, so that you simulate $M(1)$ and $M(0)$ for each individual. [The model will include `logBMI`, and so when you simulate $M(1)$, use $L(1)$ in place of L, and when you simulate $M(0)$, use $L(0)$ in place pf L.]

(4) Finally, fit a model for `SBP` given all other variables, and use this model to predict $Y(1, M(1))$, $Y(1, M(0))$ and $Y(0, M(0))$ for each individual. Eg when predicting $Y(1, M(0))$ you will use 1 in place of X, $L(1)$ in place of L and $M(0)$ in place of M.

Tasks

Question 4 (cont'd)

(5) Take differences of these three predicted potential outcomes for each individual as follows:

$$\widehat{\mathrm{OE}}_i = Y(1, M(1)) - Y(0, M(0))$$

$$\widehat{\mathrm{NDE}}_i = Y(1, M(0)) - Y(0, M(0))$$

$$\widehat{\mathrm{NIE}}_i = Y(1, M(1)) - Y(1, M(0))$$

(6) Finally, take the average of these individual differences over all individuals to obtain the MC estimates of the OE, NDE and NIE.

Tasks

Question 4 (cont'd)

A few additional things to note:

(A) We can reduce the MC error in our estimates by increasing the sample size for which we predict all the potential outcomes.

(B) For inference, we use the bootstrap; that is why we include all our code into a 'program', which can then be called by Stata's `bootstrap` command.

(C) It might be sensible to start by trying the MC simulation procedure for the two analyses we've already carried out, i.e. ignoring BMI, first without the *XM* interaction, and then with it. Then, in a third step, try adding the intermediate confounder.

For more help with the Stata syntax, see `CaseStudy1_Q4.do`.

Outline

Outline

Summary of the afternoon

- Questions concerning mediation are often posed and tie in with our intuition on what it means to ‘understand mechanism’.

Summary of the afternoon

- Questions concerning mediation are often posed and tie in with our intuition on what it means to ‘understand mechanism’.
- Traditional mediation methods (‘product’ or ‘difference’) suffer from the same vagueness that has plagued all informal statistical methods for causal inference. What exactly is being estimated? Under what assumptions is our estimation method successful?

Summary of the afternoon

- Questions concerning mediation are often posed and tie in with our intuition on what it means to ‘understand mechanism’.
- Traditional mediation methods (‘product’ or ‘difference’) suffer from the same vagueness that has plagued all informal statistical methods for causal inference. What exactly is being estimated? Under what assumptions is our estimation method successful?
- Traditional mediation methods are also limited to simple linear models.

Summary of the afternoon

- Questions concerning mediation are often posed and tie in with our intuition on what it means to ‘understand mechanism’.
- Traditional mediation methods (‘product’ or ‘difference’) suffer from the same vagueness that has plagued all informal statistical methods for causal inference. What exactly is being estimated? Under what assumptions is our estimation method successful?
- Traditional mediation methods are also limited to simple linear models.
- The causal inference literature, using counterfactuals, has clarified what we might mean by ‘direct’ and ‘indirect’ effects, but there isn’t just one possibility.

Summary of the afternoon

- Questions concerning mediation are often posed and tie in with our intuition on what it means to 'understand mechanism'.
- Traditional mediation methods ('product' or 'difference') suffer from the same vagueness that has plagued all informal statistical methods for causal inference. What exactly is being estimated? Under what assumptions is our estimation method successful?
- Traditional mediation methods are also limited to simple linear models.
- The causal inference literature, using counterfactuals, has clarified what we might mean by 'direct' and 'indirect' effects, but there isn't just one possibility.
- It has led to clear assumptions under which these can be identified, and a myriad methods for estimation, reaching far beyond two simple linear models.

Summary of the afternoon (cont'd)

- Today we have focussed on the fully-parametric approach, both analytic and using MC simulation.

Summary of the afternoon (cont'd)

- Today we have focussed on the fully-parametric approach, both analytic and using MC simulation.
- Today we have focussed only on the setting with a continuous outcome and mediator, and with a single mediator of interest.

Summary of the afternoon (cont'd)

- Today we have focussed on the fully-parametric approach, both analytic and using MC simulation.
- Today we have focussed only on the setting with a continuous outcome and mediator, and with a single mediator of interest.
- In tomorrow's workshop, we turn to mediation analysis with multiple mediators, and we'll look at a setting with a binary outcome.

Summary of the afternoon (cont'd)

- Today we have focussed on the fully-parametric approach, both analytic and using MC simulation.
- Today we have focussed only on the setting with a continuous outcome and mediator, and with a single mediator of interest.
- In tomorrow's workshop, we turn to mediation analysis with multiple mediators, and we'll look at a setting with a binary outcome.
- See Tyler VanderWeele's (2015) wonderful book for the many many topics we have not been able to cover: semiparametric estimation methods, time-to-event outcomes, three- and four-way decompositions, etc.

Outline

References: traditional approach

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

Wright, S. (1934) The method of path coefficients
Annals of Mathematical Statistics, 5(3)161–215.

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.

References: early investigations by causal inference

Robins, J.M. and Greenland, S. (1992)
Identifiability and exchangeability for direct and indirect effects.
Epidemiology, 3:143–155.

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

Cole, S.R. and Hernán, M.A. (2002)
Fallibility in estimating direct effects.
International Journal of Epidemiology, 31(1):163–165.

References: more on identification and estimation

Robins, J. (1999)

Testing and estimation of direct effects by reparameterizing directed acyclic graphs with structural nested models.

In *Computation, Causation, and Discovery*, C. Glymour & G. Cooper, eds. Menlo Park, CA, Cambridge: AAAI Press/The MIT Press, pp. 349–405.

Petersen, M.L., Sinisi, S.E. and van der Laan, M.J. (2006)

Estimation of direct causal effects.

Epidemiology, 17:276–284.

VanderWeele, T.J. and Vansteelandt, S. (2009)

Conceptual issues concerning mediation, interventions and composition.

Statistics and its Interface, 2:457–468.

References: more on identification and estimation

VanderWeele, T.J. (2009)
Marginal structural models for the estimation of direct and indirect effects.
Epidemiology, 20:18–26.

Joffe, M. and Greene, T. (2009)
Related causal frameworks for surrogate outcomes.
Biometrics, 65:530–538.

Tchetgen Tchetgen, E.J. (2013)
Inverse odds ratio-weighted estimation for causal mediation.
Statistics in Medicine, 32:4567–80.

References: key textbook on causal mediation analysis

VanderWeele, T.J. (2015)
Explanation in Causal Inference: Methods for Mediation and Interaction.
Oxford University Press.

References: on interventional effects

- Vanderweele, T.J., Vansteelandt, S. and Robins, J.M. (2014)
  Effect decomposition in the presence of an exposure-induced mediator-outcome confounder.
  Epidemiology, 25:300–306.

- Vansteelandt, S. and Daniel, R.M. (2017)
  Interventional effects for mediation analysis with multiple mediators.
  Epidemiology, 28(2):258–265.