Why we Should Rethink Causal Mediation, and What to Do Instead

Vanessa Didelez
Leibniz Institute for Prevention Research & Epidemiology – BIPS

UCL, 8 June 2026
Causal Inference Interest Group (CIIG)

Overview

Part 1: Preliminaries & Motivation

Part 2: Research questions: Lost in mediation

Part 3: Non-manipulable (‘natural’) causal mediation effects

Part 4: Separable (In)Direct Causal Effects

Conclusions

Part 1:

Preliminaries & Motivation

Why Causal Mediation Analysis?

Typical context:

$A =$ treatment
$Y =$ outcome

and other things occur in between

— co-interventions / rescue medication
— imperfect adherence
— measurements of biomarkers / surrogate outcomes
— death

Why Causal Mediation Analysis?

Can do lots of cool maths/stats...

But what does it mean (if anything)? What is the practical relevance?

Depends on:

- What is the research question (RQ)?
- Intended use to inform (which?) decision making / actions?
- Under what assumption is the RQ “answerable” (identification)
  — and are the assumptions defendable in data context?

“Understanding Mechanisms?”

“Understanding” is not a research question

- How do you verify “understanding”?
  e.g. does it lead to developing actions that improve health?

“Understanding” should be implementable in the sense of:

- We will get an idea what to do to improve patients’ health, or
- We will get new ideas for what else to before we can improve patients’ health

Causal inference:

- Support decision making ... *(Dawid, 2021:JCI)*
- ... by answering ‘what if’ questions *(Hernan & Robins, 2020:book)*

“It is *Only* Because...”

Example:
Effect of new treatment (compared to standard) is small / large “only because” of some unintended consequences
– e.g. quality of care, adherence, switching, rescue medication...

This is also not a research question

- How would it affect actions / decisions if we knew that it really is “only because”, or if we knew that it really is NOT “only because”?
- Can / would we remove the unintended consequences?

⇒ More transparency by making actionable aspects of RQ clearer

Notation

A treatment / exposure — assumed **randomised** (mostly)

A_1, A_2 sequence of treatments / exposures

Y (primary) outcome

M something measured inbetween, perhaps a 'mediator'
$\longrightarrow$ typically **not randomised**

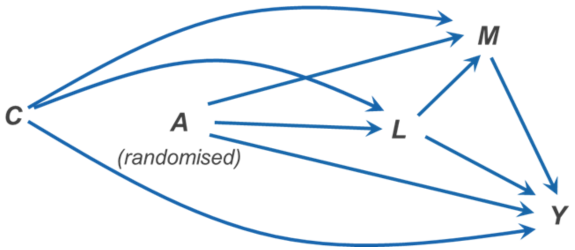
C (set of) baseline covariates

L post-treatment covariates

U unobserved variables

Setting as Causal DAG

Minimal realistic causal DAG under randomisation of A but typically without randomisation of M (some subset of (C, L) may be unmeasured / unobservable:



C baseline confounders of M and Y;
L known as treatment-induced M-Y confounding

Conditional Independence

Random variables X and Y are conditionally independent given Z iff

$$P(X, Y \mid Z) = P(X \mid Z)P(Y \mid Z)$$

We denote this with

$$X \perp\!\!\!\perp Y \mid Z$$

Note: Unlike causation, conditional (in)dependence is symmetric; it can be read off of a causal DAG by 'd-separation' (not covered here)

Association vs. Causation

Potential Outcomes (POs)

Association: observing A helps to predict Y

Causation: manipulating A changes distribution of Y

Popular formalism: potential outcomes $Y(a)$ (cf. Rubin, 1974)

$Y(a) =$ value that Y would take under an intervention setting $A = a$.

$$P(Y \mid \text{intervene to set } A = a) \stackrel{\sim}{=} P(Y(a))$$

also know as counterfactuals, because $\{Y(a), Y(a'); a \neq a'\}$ can logically not be observed together

Identification

Total Causal Effect

$P(Y(a))$ is **identified** with C pre-exposure covariates from observational data on (Y, A, C) under

Assumption of **conditional exchangeability** given C:

- graphically: all backdoor-paths from A to Y blocked by C
- with POs: $Y(a) \perp\!\!\!\perp A \mid C$

Consistency: if $A = a$ then $Y = Y(a)$

Also: Positivity!

Note: Obtain effect of A on mediator by replacing Y with M

G-Formula

Identified by the **g-formula** (standardisation)

$$P(Y(a)) \overset{identif.}{=} \sum_c P(Y \mid A = a, C = c)P(C = c)$$

$\Rightarrow$ can identify e.g. **(average) total causal effect (TCE)** (A binary)

$$E(Y(1)) - E(Y(0)) \overset{\text{identif}}{=} \sum_c [E(Y|A = 1, C = c) - E(Y|A = 0, C = c)]P(C = c)$$

or by IPTW etc. *(g-formula: Robins (1986:MathMod))*

Cross-World Concepts vs SWIGs

Quantities that involve interventions setting A to **different values at the same time** are known as **cross world** because they cannot be realised in a **single world**

Their identification typically relies on assumptions about the joint distribution of $\{Y(a), Y(a'), a \neq a'\}$ (or of other variables), which can never be observed together

$\Rightarrow$ Critical view of this has prompted the development of the **single-world framework**, in particular: single-world intervention graphs: **SWIGs**

(Robins & Richardson, 2013:UAI)

Aside (1)

Target trial

Single-world estimands can typically be described by a (more or less) straightforward experimental design that does not involve ‘turning back time’ or raising the dead, and that would yield the desired estimand with no further structural assumptions

In other words, we can describe a possible, though perhaps unethical or expensive, target trial for our estimand

Aside (2)

Cross-world concepts:

— Individual causal effects

— certain definitions of ‘Harm’

— Principal stratum effects

— so-called *Natural (in)direct effects* (cf. later)

$\Rightarrow$ All involve joint counterfactual concepts; and make cross-world assumptions for identification

Decision Theoretic Approach

Instead of potential outcomes, use regime indicator(s) F_A, F_M indicating either observational $F_A = \emptyset$ or interventional regimes $F_a = a$ *(Dawid, 2021:JCI)*

Then,

$$P(Y|F_A = a, F_M = m) \quad \text{replaces} \quad P(Y(a, m))$$

and, e.g., controlled direct effect (more later) corresponds to

$$CDE = E(Y|F_A = 1, F_M = 0) - E(Y|F_A = 0, F_M = 0)$$

Note: counterparts to assumptions s.a. consistency can be formulated at distributional level

Remark: cross-world quantities cannot be expressed in this framework

Part 2:

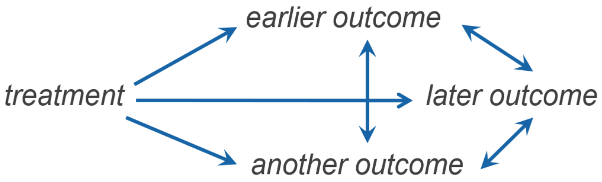
Research questions: Lost in mediation

(based on: Didelez, 2025:chapter:HandbookEpidemiology)

Many Effects of Treatment?

Not a mediation question:

- what are various *different effects* of treatment?



Multiple Total Causal Effects

Idea: In view of various intercurrent events, before we attempt a ‘causal mediation analysis’, we may want to describe the causal effect of A (randomised) on various relevant post-randomisation quantities $\longrightarrow$ **causal exploration**

Identified as per randomisation: effects of treatment initiation on

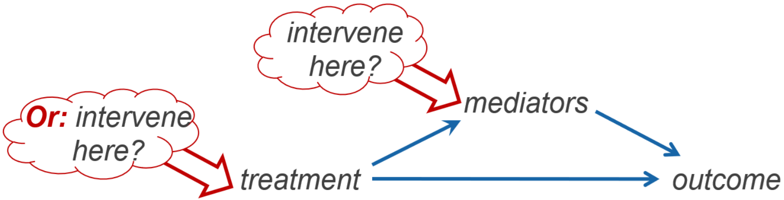
1) adherence and side effects
2) frequency of rescue medication
3) co-interventions or quality of care
4) relevant biomarkers (blood-pressure, ...)
5) death (as competing event) etc.

$\Rightarrow$ Provides first insights / generates hypotheses on workings of treatment

Where Best to Intervene?

Also not a mediation question:

- *where best* to intervene – on treatment or mediator?



- Example: is it better to increase number of check-ups or to improve care after diagnosis (if you can't do both)?

Where Best to Intervene?

Still: total causal effect(s)

$\Rightarrow$ Compare:

— total effect of A on Y (identified by randomisation)

with

— total effect of M on Y;

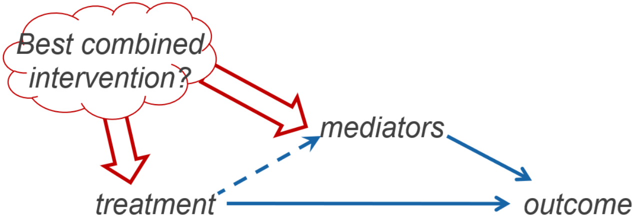
$\rightarrow$ needs adjustment for confounding as M not randomised

$\Rightarrow$ Planning phase should ensure possible confounders are measured

Joint / Sequ. / Adapt. Intervention?

Also not a mediation question:

- what is best *joint* (seq./dyn.) intervention on treatment & mediator



- Example: Treatment followed by a fixed rule for when to administer a co-intervention *(adaptive / dynamic intervention)*

Joint / Sequ. / Adapt. Intervention?

Think of M as second treatment A_2

Potential outcome: $Y(a, m)$ = outcome if we set $A = a$ and $M = m$ by an intervention

1) Fix $M = m$ — controlled direct effect

2) Contrast pairs of treatments — joint effect

3) Make m a function of covariates (C, L) — adaptive/dynamic treatment

Controlled Direct Effect

Intervene on A while keeping fixed M by intervention (e.g., at baseline)

$$CDE = E(Y(A = a, M = 0) - Y(A = a', M = 0))$$

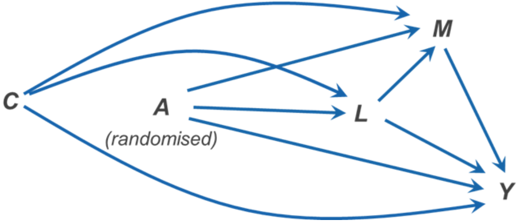
Example:
Effect of primary intervention while not allowing any co-intervention in any arm

Also: ICH-E9 Addendum calls this ‘hypothetical’ estimand

Identification of CDE

(1) No-unobserved-confounding of A and Y (randomised) and

(2) No-unobserved-confounding of M and Y given (A, C, L):



Again: g-formula / IPW etc.

Controlled Direct Effect

Advantage: CDE conceptually simple; identifying conditions straightforward; can be related to parameters of variety of regression models

$\Rightarrow$ **Will suffice in many applications!**

But note: *no corresponding notion of 'indirect' effect*

In fact: M could be prior / post A or both could be non-causes of each other

Joint / Sequential Effects

CDE special case of joint effects:

A_1, A_2 two treatments, possibly in a sequence

$$JCE = E(Y(a_1, a_2)) - E(Y(a'_1, a'_2))$$

or for adaptive treatments, make a_2 function of relevant post-treatment covariates

Identification and estimation well studied

(General identification of joint effects see Shpitser & Pearl (2006); plenty of work on sequential and adaptive treatments)

Remarks

- All the above effects are often meaningful, defined in a single world, explicitly 'actionable', and well studied; they will often tell us much about the desired causal relations
- There is, above, no notion of indirect effect (is that a problem?)
  People seem to expect: 'Total = direct + indirect effect' (is that a RQ?)

  I don't understand why this should be important...
- Further problem: what if we cannot intervene on M?
  Example: $M =$ survival?

  My question back: well what exactly is your research question?

Part 3

Non-manipulable (‘natural’) causal mediation effects

Background

- Traditionally (in many fields): mediation = path analysis, based on **linear structural equation models** (LSEMs).
- **Disadvantage:** LSEMs overly simplistic, no interactions, do not work in non-linear settings etc.
- **More importantly:** Parameter-based notions of (in)direct effects have no 'life' outside particular parametric model

Nested Counterfactuals

Key quantity:

$$Y(a, M(a'))$$

where $a \neq a'$ is possible

In words: The potential outcome for Y if A were set by intervention to a and M were set by intervention to its value had A instead been set by intervention to a'

$\Rightarrow$ Cross-world concept, as $A = a$ and $A = a'$ for the same individual at the same time is impossible

Natural (In)Direct Effects

definition

Definition *(Robins & Greenland, 1992:Epidem; Pearl 2001:UAI)*

$$NDE = E(Y(a', M(a')) - Y(a, M(a')))$$

$$NIE = E(Y(a, M(a')) - Y(a, M(a)))$$

(or switch a, a')

Or: other contrasts, e.g. relative risks (the choice of contrast is not essential)

Robins & Greenland (1992:Epidem) called them ‘pure’ effects
Robins & Richardson (2011:bookChapt) qualified them as *non-manipulable*

Natural (In)Direct Effects

meaning?

Remark

It is usually left unclear / not explained how to intervene on M to make it take the value $M(a')$ while at the same time $A = a$

Note: *Pearl (2001:UAI)* provides a story about 'eliminating adverse side-effects (prompting aspirin use M) of a drug'
$\longrightarrow$ We will return to this later

Natural Effects Target Trial?

Imai et al (2013:JRSSA) describe a crossover design:
set A randomly at first, remember value of M, in 2nd phase set A to the other value but fix the value of M to the first one $\rightarrow$ NIE

They assume

- causal consistency *across time points*
- no carry-over effects

$\Rightarrow$ Cross-world assumptions that allow us to 'turn back time'

They give some (not very convincing) examples

Note: Designs are still interesting, just not convincingly targetting NDE/NIE

Natural (In)Direct Effects

early & ongoing debate

- Robins & Greenland 'dismissed' these concepts of (in)direct effects

  (see also Robins (2003:HSSS-book))

  ...they had (single-world) FFRCISTG (no SWIGs yet) models in mind, not using cross-world assumptions

- Pearl (2001:UAI), instead, had (cross-world) NPSEM-IE in mind with stronger assumptions so that he could 'identify' the natural (in)direct effects

Effect decomposition

properties of NDE/NIE

Assuming only consistency; no particular parametric model

First, recall that $Y(a, M(a)) = Y(a)$ then for total effect:

$$
\begin{aligned}
E(Y(a') - Y(a)) &= E(Y(a', M(a')) - Y(a, M(a))) \\
&= E(Y(a', M(a')) - Y(a, M(a'))) \\
&\quad + E(Y(a, M(a')) - Y(a, M(a))) \\
&= NDE + NIE
\end{aligned}
$$

$\Rightarrow$ proportion mediated $= NIE/(NDE + NIE)$

Note: decomposition depends on reference value a or a' (interactions!)

NDE/NIE: Identifying Assumptions

As before: consistency, positivity

No unmeasured confounding

$Y(a,m) \perp\!\!\!\perp A \mid C, \quad M(a) \perp\!\!\!\perp A \mid C$ $\longrightarrow$ both valid under randomised A ($C = \emptyset$)

$Y(a,m) \perp\!\!\!\perp M \mid (A = a, C)$ $\longrightarrow$ no unmeasured M-Y confounding

Cross-world independence:

$Y(a,m) \perp\!\!\!\perp M(a') \mid C$

(for detailed discussion cf. Andrews & Didelez, 2021:Epidemiology)

Cross-World Independence

$$Y(a, m) \perp\!\!\!\perp M(a') \mid C$$

When valid? How to reason about it? Cannot be read-off of causal DAG

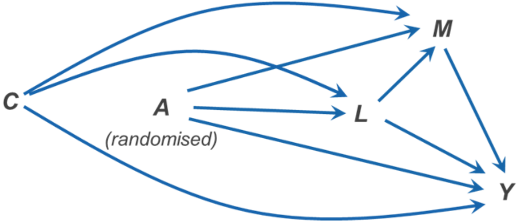
Not even in principle empirically verifiable

Some (even stronger) assumptions imply the above:
— rank preserving structural models
— linear-additive structural equation & constant individ. effects
— Robins et al (2022) give an example where only CWI is violated

(Robins et al, 2022:book-chapter:arXiv)

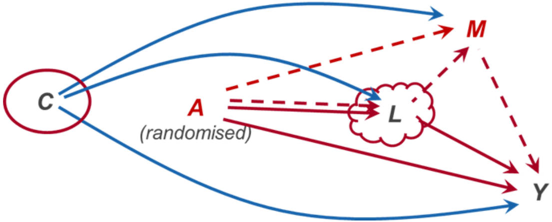
Cross-World Independence

Cross-world independence does not generally hold if there is treatment-induced confounding, i.e. $L \neq \emptyset$



Cross-World Independence

Cross-world independence does not generally hold if there is treatment-induced confounding, i.e. $L \neq \emptyset$



Because: extended nested counterfactual $Y(a, M(a')) = Y(a, L(a), M(a', L(a')))$ — and $\{L(a'), L(a)\}$ are not generally mutual. independent

L also called 'recanting witness' *(Avin et al, 2005:UAI)*

Mediational G-Formula

C baseline covariates, L empty

Under above identifying assumptions we have the identifying functional:

$$\begin{aligned} E(Y(a, M(a'))) &= \sum_{m,c} E(Y \mid A = a, M = m, C = c) \\ &\times P(M = m \mid A = a', c) P(C = c) \end{aligned}$$

⇒ Plenty of methods to estimate this (R packages: ‘mediation’ or ‘medflex’)

Aside

Didelez, Dawid, Geneletti (2006) showed:

Can define a *stochastic* intervention on M, denoted by $F_M = d_{M|CA}$, s.t.

$P(Y|F_A = a, F_M = d_{M|CA=a'})$ and $P(Y(a, M(a')))$

have same identifying functional: mediational g-formula

where C must be chosen so that the total effect equals the sum of the stochastic direct and indirect effects

Restriction: choose C s.t. $\quad Y \perp\!\!\!\perp F_M | (M, C, F_A = a) \quad$ (*)

Aside ctd.

- The above restriction (*) sort of replaces the CWI (without multiple worlds)
- It has been re-derived later by *Lok (2016:SiM)* with potential outcomes instead of decision theoretic framework, in context of ‘organic’ (in)direct effects
- Aka ‘randomized interventional (analogue) mediation effects’ etc.

  (VanderWeele et al, various papers)

- Arguably, the concept is not very useful; not clear what the *practical* counterpart of $F_M = d_{M|CA}$ could be — how does this help decision making?

  see also arguments by *Miles (2023:JRSSB)*

- Careful with alternative estimands — must be wary of ‘identity slippage’

  (Sarvet et al, 2023:arXiv)

Longitudinal & Survival Setting

Violation of CWI by treatment-induced confounding implies:

NDE/NIE not identified with **survival outcomes** and **longitudinal mediator(s)** as prior survival is always required for *existence* of later variables

$\Rightarrow$ recanting witness

(Didelez, 2019:LiDA;
Fulcher, Shpitser, Didelez, Zhou, Scharfstein, 2021:Biom)

Remarks

- Natural (in)direct effects defined in terms of nested counterfactuals $Y(a, M(a'))$ — unclear to what practical setting this could apply
  $\Rightarrow$ how would you 'set' $M(a')$ while $A = a$?
- Perhaps mainly seen as thought experiment
  — but what is the practical use, then?
- When the practical use is addressed, it rarely has to do with the precise interpretation of the nested counterfactual
  ... even Pearl then tells a different story!
- No decision theoretic counterpart of these concepts

Part 4:

Separable (In)Direct Causal Effects

expanding the story

Motivating Example

Example: RCT; new drug does not work as expected *(similar to Pearl, 2001:UAI)*

Investigate:

- new drug has uncomfortable (but not dangerous) side-effect
- adherence is low in treatment arm, presumeably due to side-effect
- with lower adherence, new drug presumeably less effective

What could be done about it?

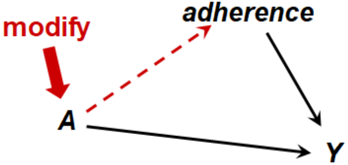
1) Intervene on (enforce better) adherence? $\rightarrow$ CDE

2) Remove side-effect *by changing chem. composition of drug*?

Motivating Example ctd.

.

Modify treatment, specifically targetting undesirable pathway



⇒ leads to **Separable direct effect (SDE)**

A bit more Formally

$A = 1, 0$ (new / old) drug; $M = 1, 0$ (good / bad adherence); $Y = 1, 0$

Assume: treatment components $A = (A^Y, A^M)$

e.g. A^Y = active ingredient; A^M = ingredient yielding side-effect

Potential outcomes:

$Y(A^Y = a, A^M = a')$ under hypoth. chem. modification of drug

Separable direct effect

$$SDE = E(Y(A^Y = 1, A^M = 1)) - E(Y(A^Y = 0, A^M = 1))$$

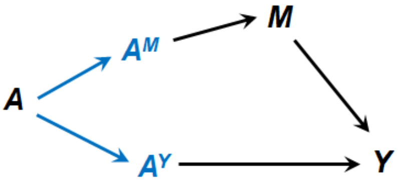
= treatment effect if we remove ingredient responsible for side-effect ($\approx$ old drug)

Formally: Separable Effects

Robins & Richardson (2011:bookChapt)

Assume A has components $A = (A^M, A^Y)$ s.t.

- observationally: $A \equiv A^M \equiv A^Y$
- A^M 'activates' the mediating pathway
- A^Y 'activates' pathway avoiding M_t
- hypothetical intervention:
  can set A^M and A^Y to different values



Target of inference:

$$E(Y(A^Y = a, A^M = a'))$$

Separable direct effect (SDE): vary $a \in \{0, 1\}$, fix a'

Separable indirect effect (SIE): fix a, vary $a' \in \{0, 1\}$

More Examples

Many treatments / exposures consist of (potentially) separable components:

- actual double blind RCTs: $A^M =$ 'pill', $A^Y =$ 'active ingredient'

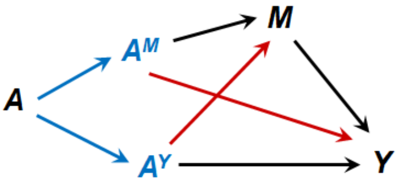
  (not even hypothetical!)
- medical Marijuana: $A^M =$ 'hallucinogenic substances', $A^Y =$ 'other subst.'
- weight-loss programmes: $A^M =$ 'meetings', $A^Y =$ 'change in diet'
- breast feeding: $A^M =$ 'physical experience', $A^Y =$ 'mother's milk'
- treatments with different secondary care / preventive medication

$\Rightarrow$ ... often, when subject matter experts explain in *practical* terms their interest in mediational questions, it seems to be about separable effects

Separability

Well-defined?

- absence of red edges
  = 'dismissible' component
  (or exclusion restriction)
- interventions on (A^Y, A^M) may or may not be well-defined / of substantive interest
- Note that observationally: $A \equiv A^M \equiv A^Y$ always



'Not separable': no covariates will identify effects when data $A \equiv A^M \equiv A^Y$

Separable Treatments Target Trial?

Robins et al (2022) describe hypothetical four-arm trial as the target trial for separable effects:

- With all combinations of $A^Y, A^M \in \{0, 1\}$
- Can be meaningful in itself (outside mediation context)

Example: extended placebo trial

Identification

Property 1:

$$Y(a) = Y(A^M = a, A^Y = a)$$

$\approx$ 'causal consistency' between two experimental settings

– RCT with original $A \Rightarrow$ **two-arm trial**

– RCT on components $A^M, A^Y \Rightarrow$ hypothetical **four-arm trial**

(Robins, Shpitser, Richardson, 2022)

Identification

Under do(A^Y, A^M), i.e. in hypothetical four-arm trial:

(measured, baseline) covariates C s.t.

Assumption 1:

$$M \perp\!\!\!\perp A^Y \mid (C, A^M = a^*)$$

Assumption 2:

$$Y \perp\!\!\!\perp A^M \mid (C, M, A^Y = a)$$

Identification

Under these assumptions (+ add. standard assumptions)

Identifying functional:

$$E(Y(A^Y = a, A^M = a')) = \sum_{m,c} E(Y \mid A = a, M = m, C = c) \times P(M = m \mid A = a', c) P(C = c)$$

$\Rightarrow$ Mediational g-formula, again!

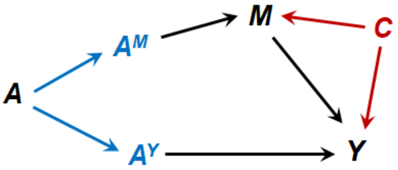
(Robins & Richardson, 2011)
(Robins et al, 2022)

Example ctd. — Side-Effect & Adherence

A = drug (new/old); M = adherence; Y = outcome

Confounding of M and Y by C

- treatment is separable
- separable effects identified from data on (A, M, C, Y)
- but not without data on C (A2 violated)

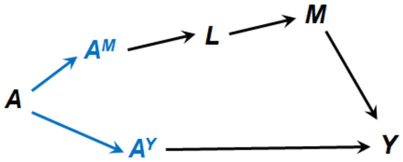


Example ctd. — Side-Effect & Adherence

A = drug (new/old); L = side effect; M = adherence; Y = outcome

So far, simplistic assumptions:

- side effect does not affect outcome other than through adherence
- side effect not affected by active ingredient



$\Rightarrow$ Assumptions hold; sep. effects identified from data on (A, M, Y) or on (A, L, Y)

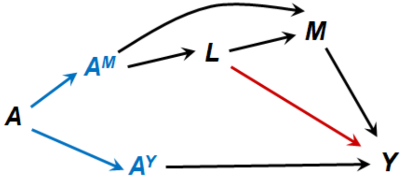
Example ctd. — Side-Effect & Adherence

A = drug (new/old); L = side effect; M = adherence; Y = outcome

What if $L \to Y$?

do not want to assume $Y \perp\!\!\!\perp L \mid M$

$\Rightarrow$ separable effect still identified,
but need data on L (and M)



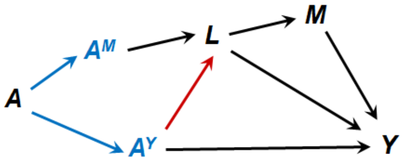
Example ctd. — Side-Effect & Adherence

$A =$ drug (new/old); $L =$ side effect; $M =$ adherence; $Y =$ outcome

What if $A^Y \rightarrow L$?

i.e. active ingredient *does* affect side-effect, perhaps intrinsic part of drug's working

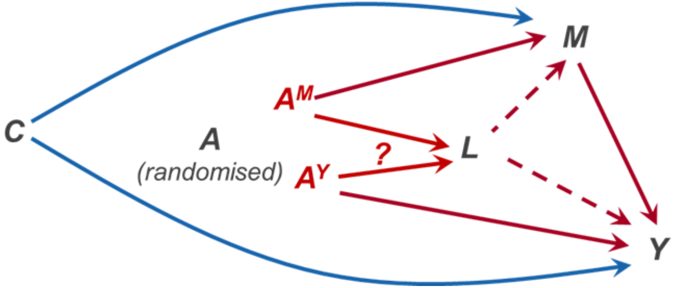
$\Rightarrow$ not separable, not identifed from data where $A = A^Y = A^M$;



But: could still actually carry out the four-arm trial...

Treatment-Induced Confounding?

Separable effects with L: must clarify relation to (A^Y, A^M)



Prompt: prehaps need to re-think our research question!

Remarks

- NDE/NIE and SDE/SIE have the same identifying functional in the absence of treatment-induced confounding

$\Rightarrow$ Statistical estimation can proceed in the same way for both

- BUT: they are different concepts with different assumptions
- Can view separable treatments as single-world reformulation of 'natural' effects concepts (albeit an **expanded** single world)

  — more conducive to elicit actionable / useful research questions
  — see 'Story-led causal inference' *(Young, 2025:Epid)*

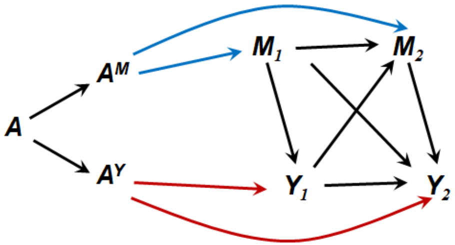
Remarks ctd.

- Treatment-induced confounding by L? Maybe rethink your RQ?
- Survival with longitudinal mediator: NDE/NIE never identified — unlike separable effect *(Didelez, 2019:LiDA)*

Longitudinal / Survival Settings

Separable treatment effects "⇔" certain paths-specific effects:

partitioning of paths by edges starting $A \longrightarrow \cdots$



Note: absence of edges represent key assumptions ('dismissible components')

Remarks ctd.

- Treatment-induced confounding by L? Maybe rethink your RQ?
- Survival with longitudinal mediator: NDE/NIE never identified
  — unlike separable effect *(Didelez, 2019:LiDA)*
- Further extension to competing events leading to new estimand
  (Stensrud et al, 2022:JASA)
- Still, more thought needed with research questions adressing outcomes like 'survival and cognitive function'
  $\longrightarrow$ but see 'conditional separable effects' *(Stensrud et al, 2023:JASA)*

Conclusions & Outlook

Conclusions

Over many technical issues, must not forget most important points:

- What is the research question / target of inference and is it adequately addressed by causal mediation approaches?

  $\rightarrow$ if the question is relevant for decision making, then there must be a 'single-world' estimand reflecting it

- In putative mediation settings: many actionable / useful estimands are **not** about mediation
  – adaptive interventions, effect modification, controlled direct effects
  – using single-world assumptions, thus in principle empirically verifiable

Summary

- Practical usefulness of mediational estimands that rely on cross-world concepts is mysterious
  – why consider them at all?
- Longitudinal / survival settings especially challenging — prior survival always required for later variables to even be defined, but violates CWI;
  multitude of paths: which are relevant to meaningful research questions?
- *Sometimes* useful: expand the story to hypotheticals like imperfect interventions or separable treatments
  $\Rightarrow$ Can help with *eliciting* actionable research questions
  $\Rightarrow$ ... and novel approach for competing events settings

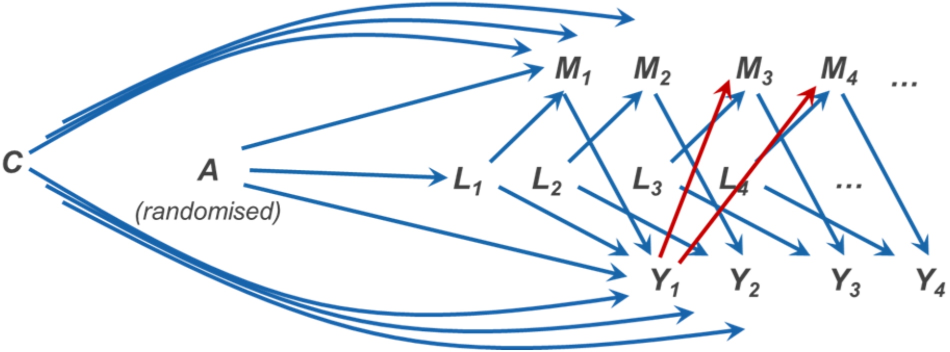
Outlook

Many topics beyond today's talks

- Bias/sensitivity analysis; in particular: bounds of causal effects under violations of various assumptions, in particular of 'cross-world independence'
- Various alternative estimands to get around 'treatment induced' mediator-outcome confounding
  $\rightarrow$ must be wary of 'identity slippage' *(Sarvet et al, 2023:arXiv)*
- Multivariate / high-dim mediators or longitudinal / time-to-event settings
  $\rightarrow$ Be careful with impossible assumptions!
- Lots of cool statistical methods...

More Realistic DAG – with Time

What is the question and what is still identified?



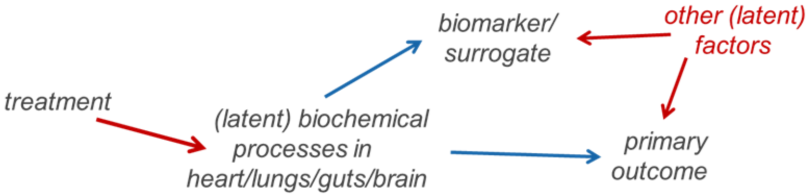
Thank You!

www.leibniz-bips.de/en

Contact
Vanessa Didelez
Leibniz Institute for Prevention Research
and Epidemiology – BIPS
Achterstraße 30
D-28359 Bremen
didelez@leibniz-bips.de

Biomarker / Surrogate Outcome

What defines a suitable biomarker?



With above structure, causal mediation analysis may not be meaningful

Some References

Aalen, Stensrud, Didelez, Daniel, Røysland, Strohmaier, Time-dependent mediators in survival analysis: Modelling direct and indirect effects with the additive hazards model, Biometrical Journal 62, 532-549 (2020)

Andrews, Didelez, Insights into the "cross-world" independence assumption of causal mediation analysis, Epidemiology 32(2), 209-219 (2021)

Didelez, Dawid, Geneletti, Direct and indirect effects of sequential treatments, Proceedings of the 22nd Annual Conference on Uncertainty in Artifical Intelligence (2006)

Didelez, Discussion of "Experimental Designs for Identifying Causal Mechanisms" by Imai, Tingley, Yamamoto – read to the RSS on 14 March 2012, JRSSA 176, 39 (2013)

Didelez, Defining causal mediation with a longitudinal mediator and a survival outcome, Lifetime Data Analysis 25, 593-610 (2019)

Didelez, Causal reasoning and inference in Epidemiology, In: Ahrens, Pigeot (eds) Handbook of Epidemiology, (2025 online)

Di Maria, Didelez, Longitudinal mediation analysis with multilevel and latent growth models: A separable effects causal approach, BMC Medical Research Methodology 24, 248 (2024)

Do, Didelez, et al. The role of psychosocial well-being and emotion-driven impulsiveness in food choices of European adolescents, International Journal of Behavioral Nutrition and Physical Activity 21(1) (2024)

Fulcher, Shpitser, Didelez, Zhou, Scharfstein, Discussion of "Causal mediation of semicompeting risks" by Huang, Biometrics, 1-5 (2021)

Rojas-Saunero, Young, Didelez, Ikram, Swanson, Considering questions before methods in dementia research with competing events and causal goals, American Journal of Epidemiology 192(8), 1415-1423 (2023)

Stensrud, Hernán, Tchetgen Tchetgen, Robins, Didelez, Young, A generalized theory of separable effects in competing event settings, Lifetime Data Analysis 27(4), 588-631 (2021)

Stensrud, Young, Didelez, Robins, Hernán, Separable effects for causal inference in the presence of competing events, Journal of the American Statistical Association 117(537), 175-183 (2022)