FACULTY OF SCIENCES

NON CLINICAL STATISTICS CONFERENCE, LOUVAIN-LA-NEUVE 2022

INFERRING CAUSAL PATHWAYS FROM DATA

Stijn Vansteelandt, Ghent University, Belgium

joint work with Tyler VanderWeele, Rhian Daniel, Johan Steen, Thang Vo

GHENT UNIVERSITY

INTRODUCTION

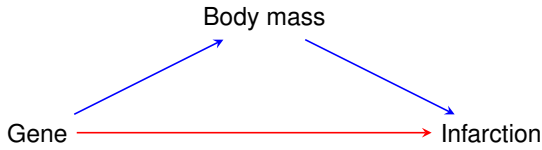
GHENT
UNIVERSITY

INTRODUCTION

- There is longstanding interest in using data
  to gain insight into the mechanism
  that underlies the effect of an exposure on an outcome.

- Mediation analyses are designed for this purpose.

PLEIOTROPY IN GENETIC ASSOCIATION STUDIES



Italian Genetic Study of Early-onset Myocardial Infarction

Does the FTO gene exert an effect on the risk of infarction
that is unmediated by changes in body mass?
How much of the genetic effect is mediated by body mass?

(Ardissino et al., 2011; Berzuini et al., 2012)

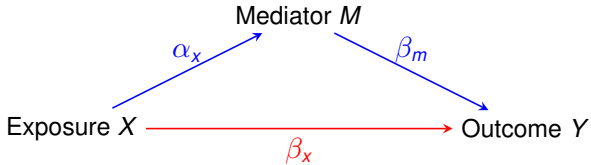
TRADITIONAL MEDIATION ANALYSIS

LOOKING BACK IN TIME...

- Mediation analysis has started to develop
  around the framework of path analysis and structural equation models.

  (Wright, 1934)
- Its development took off in the 80's
  since the seminal publication by Baron and Kenny (1986)
  in the psychological / sociological sciences.
- This framework dominates current practice.

MEDIATION ANALYSIS 1.0



$$Y = \beta_0 + \underbrace{\beta_x}_{\text{direct effect}} X + \underbrace{\beta_m}_{\text{effect of } M \text{ on } Y} M + \epsilon_y$$

$$Y = \beta_0' + \underbrace{\beta_x'}_{\text{total effect}} X + \epsilon_y'$$

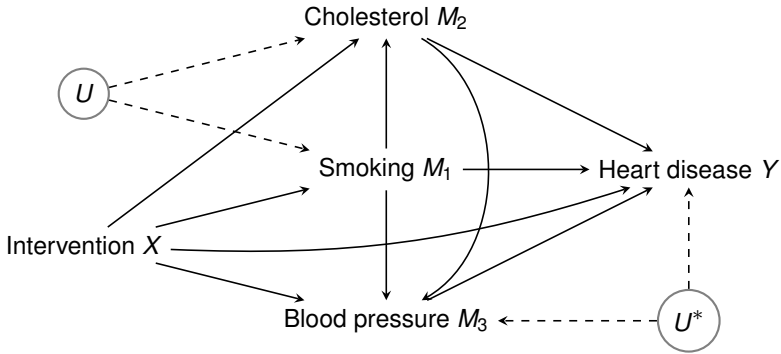
$$M = \alpha_0 + \underbrace{\alpha_x}_{\text{effect of } X \text{ on } M} X + \epsilon_m$$

The indirect effect is commonly calculated as

$$\beta_x' - \beta_x = \alpha_x \beta_m$$

BUT... CONFOUNDING PATTERNS ARE COMPLICATED

Confounding is subtle, and not given due consideration.

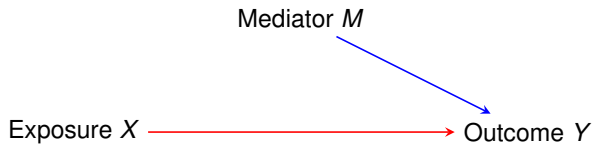


Even in the absence of unmeasured confounding,
many pathways cannot be identified.

(Avin et al., 2005; Daniel et al., 2017)

BUT... NON-LINEAR ASSOCIATIONS ARE NON-COLLAPSIBLE

- Techniques for linear models need not carry over to non-linear models.
- Adding even independent variables to a non-linear model tends to change coefficients systematically.

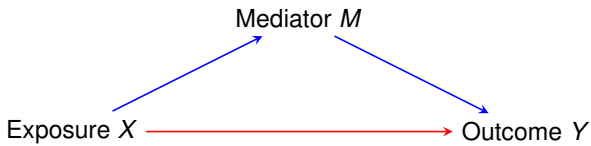


- Such non-collapsibility makes difference-of-coefficient methods problematic.

(Greenland, Robins and Pearl, 1998)

BUT... EFFECTS LACK INTERPRETATION

- Product-of-coefficient methods are also problematic.



- E.g. consider models

$$\begin{aligned} E(M|X) &= \alpha_0 + \alpha_1 X \\ \operatorname{logit} P(Y = 1|X, M) &= \beta_0 + \beta_1 X + \beta_2 M \end{aligned}$$

- *How to interpret the product $\alpha_1\beta_2$ of a mean difference and a log odds ratio?*

MEDIATION ANALYSIS 2.0

- The problem with traditional mediation analysis
  is that there is an abundance of estimation methods,
  but no understanding what they are estimating.
- In a revolutionary paper, Robins and Greenland (1992) identified these concerns
  and came up with direct and indirect effect estimands.
- This has led to a complete re-development on mediation analysis,
  which has taken off rapidly since 2010.

(Robins and Greenland, 1992; Pearl, 2001; Didelez et al., 2006; VanderWeele and Vansteelandt, 2009, 2010; Imai et al., 2010; VanderWeele, 2015)

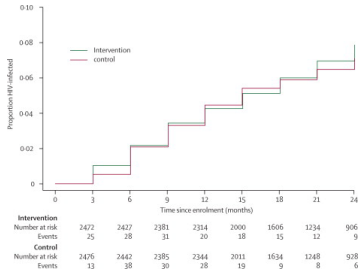
Take home message

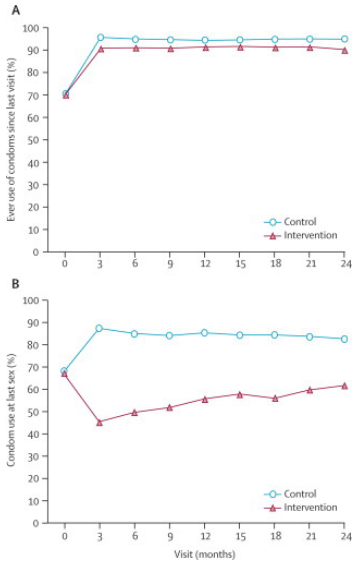
Modern mediation analysis techniques are applicable to non-linear models
and careful about problems of confounding.

NATURAL DIRECT AND INDIRECT EFFECTS

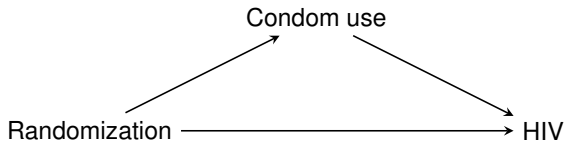
COMPLICATIONS IN THE OPEN-LABEL MIRA TRIAL

Padian et al.,
Lancet 2007





INTEREST LIES IN THE DIRECT EFFECT = 'NET' EFFECT



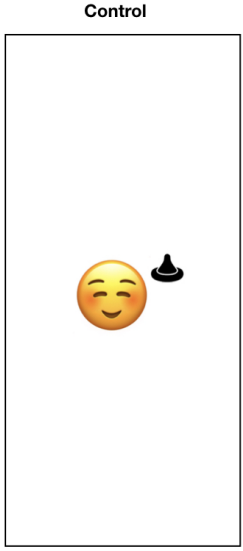
direct effect = 'net' effect

What would have been the ITT effect
had condom use not been affected by the intervention?

(Padian et al., 2007; Rosenblum et al., 2009)

How to formalise the notion of a direct effect?

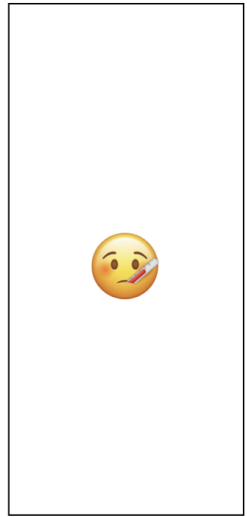
SUPPOSE WE RANDOMISE A GIVEN WOMAN TO CONTROL...



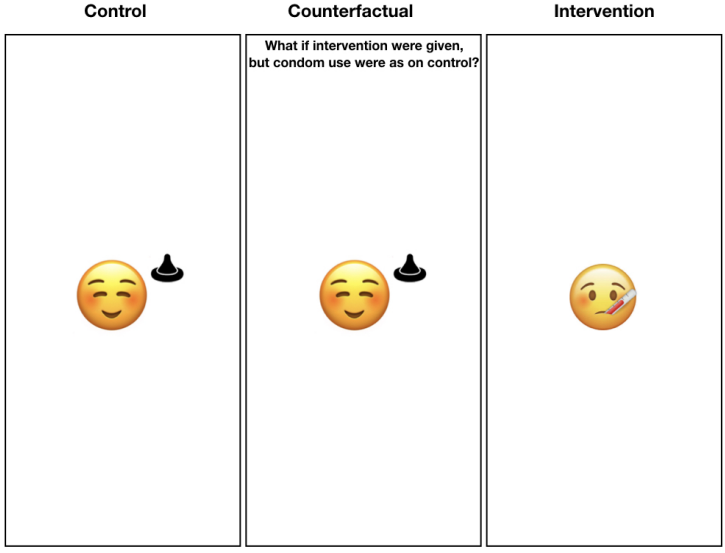
SUPPOSE WE RANDOMISE THAT WOMAN TO INTERVENTION...

Control

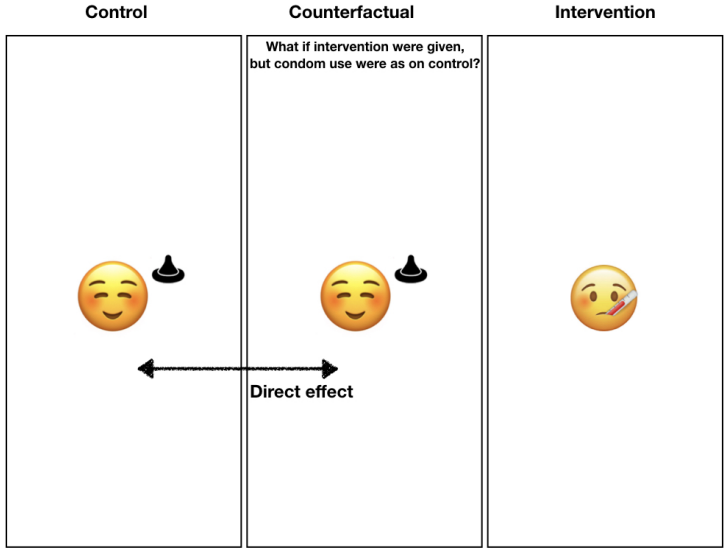
Intervention



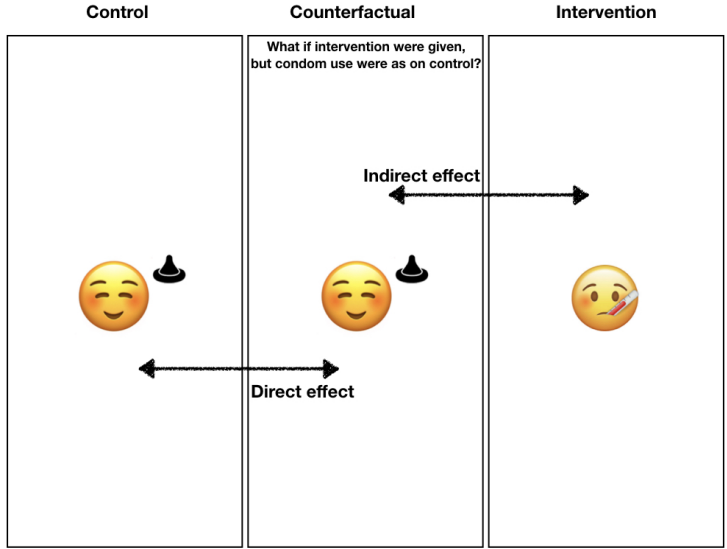
WHAT IF SHE HAD NOT CHANGED CONDOM USE...?



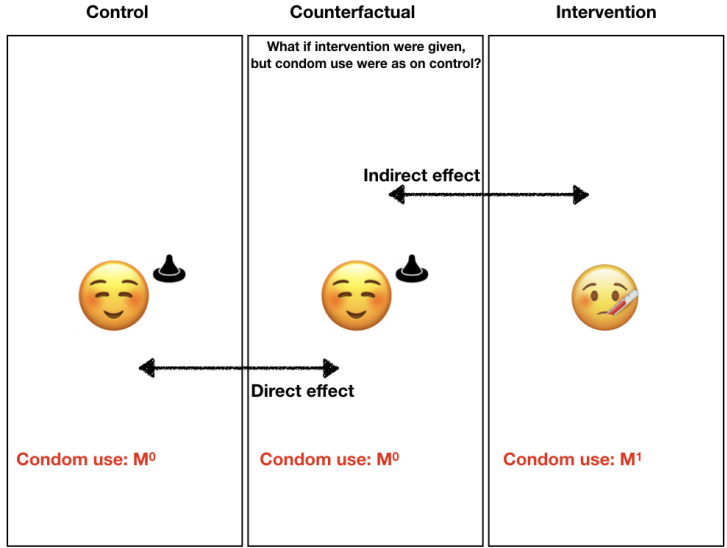
NATURAL DIRECT EFFECT



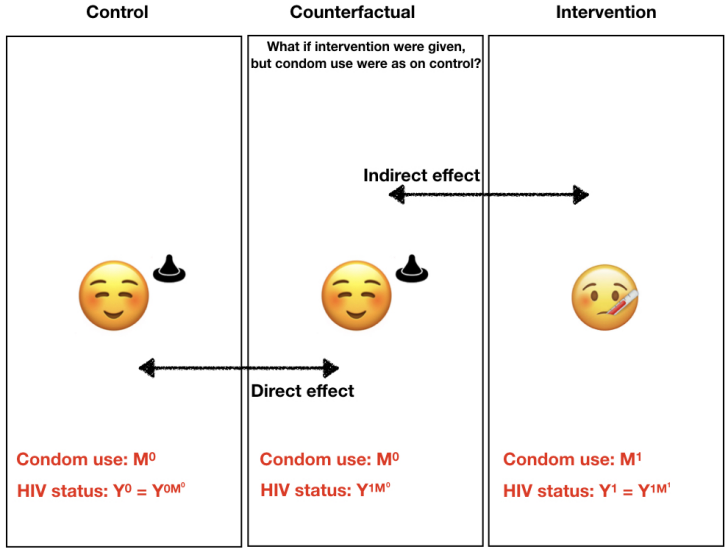
NATURAL INDIRECT EFFECT



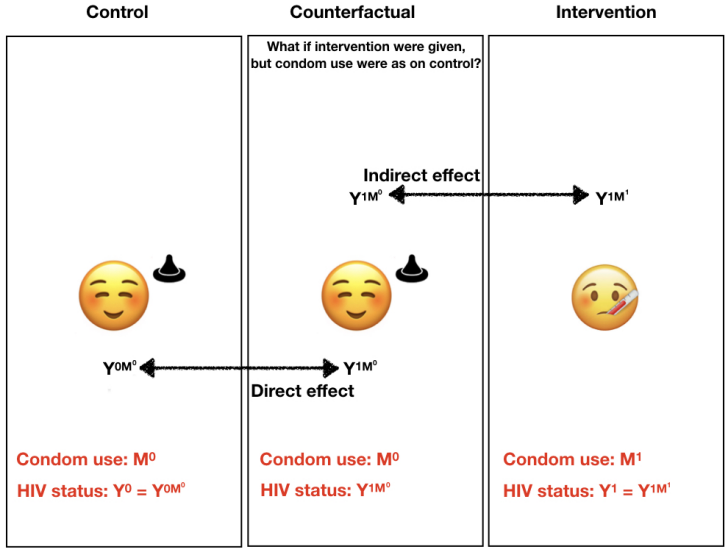
COUNTERFACTUAL DATA (1)



COUNTERFACTUAL DATA (2)



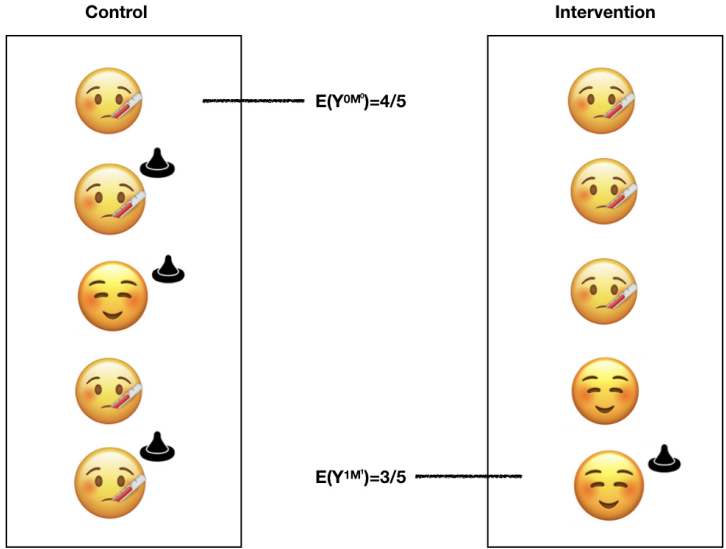
Formal definition of natural (in)direct effect



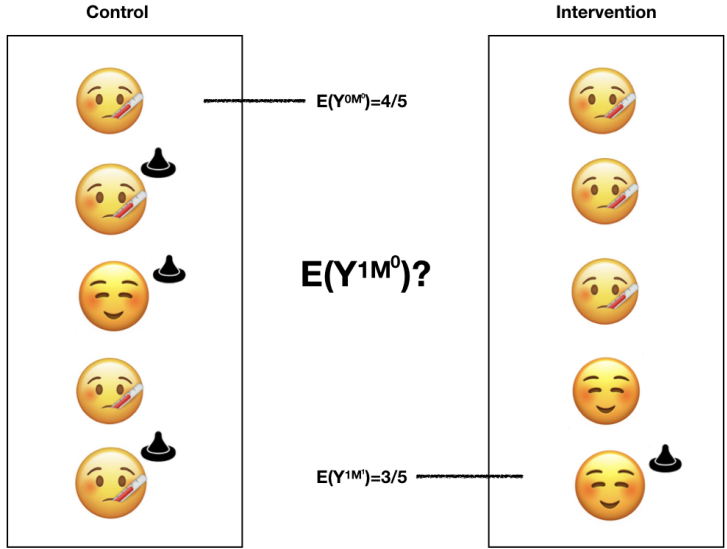
SINGLE MEDIATION ANALYSIS

GHENT
UNIVERSITY

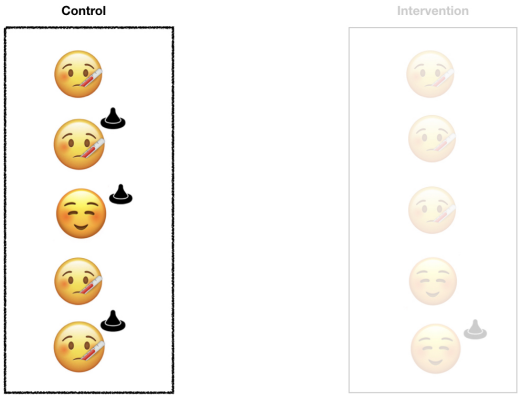
CONSIDER NOW A 'REAL' STUDY



THE FUNDAMENTAL PROBLEM IN MEDIATION ANALYSIS



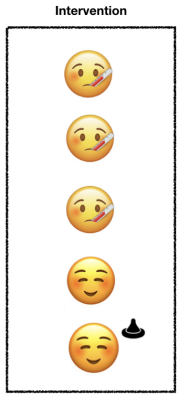
How to infer $E(Y^{1M^0})$?



- Using the mediation formula.

  (Pearl, 2001; VanderWeele and Vansteelandt, 2009, 2010)
- We do not observe M^0 for everyone.
- We will therefore 'stochastically' predict M^0 using a prediction model for the 'untreated'.
- This model should adjust for confounding of the X-M association.

How to infer $E(Y^{1M^0})$?



- We next predict what outcome would be on treatment, at level M^0.
- And then average these predictions.
- This is very flexible!
- Since we evaluate the effect of 2 interventions, *setting treatment to 1 and mediator to* M^0, the prediction model must adjust for confounding of the X-Y and M-Y associations.
- Mediation analyses thus necessitate confounding adjustment, even in randomised experiments!

Mediation analysis using mediation

(Imai, Keele and Tingley, 2010)

```
> install.packages(mediation)
> library(mediation)
> r = mediate(mody, modm, mediator = "cont", treat = "mouldbin", sims =1000)
> summary(r)
```

```
-------------------------------------------------------------------------------
       Effect                 |  Mean           [95% Conf. Interval]
------------------------------+------------------------------------------------
       ACME1                  |  .0047935      .0030167      .0069733
       ACME0                  |  .0038464      .0023583      .0056088
       Direct Effect 1        |  .0237035      .0091926         .0382
       Direct Effect 0        |  .0227564      .0088654      .0368293
       Total Effect           |  .0275499      .0131839      .0417295
       % of Total via ACME1   |  .1750696      .1148706      .3635872
       % of Total via ACME0   |    .14048      .0921748       .291751

       Average Mediation      |    .00432      .0026953      .0062752
       Average Direct Effect  |    .02323      .0090411      .0375581
       % of Tot Eff mediated  |  .1577748      .1035227      .3276691
-------------------------------------------------------------------------------
```

FITTING NATURAL EFFECT MODELS USING MEDFLEX

(Lange, Vansteelandt and Bekaert, 2012)

```
> library(medflex)
> imp <- ne.impute(UPB ~ factor(attbin) + negaff + gender + educ + age,
   family = binomial, data = UPBdata)
> fit.ne <- ne.model(UPB ~ attbin0 + attbin1 + gender + educ + age,
   family = binomial, expData = impData, se = "robust")
> summary(fit.ne)

Natural effect model with robust standard errors based on the sandwich estimator
---
Exposure: attbin
Mediator(s): negaff
---
Parameter estimates:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.807111   0.829331  -2.179 0.029332 *
attbin01     0.907959   0.289594   3.135 0.001717 **
attbin11     0.376392   0.100549   3.743 0.000182 ***
genderM      0.227916   0.286977   0.794 0.427081
educM       -0.212607   0.543467  -0.391 0.695645
educH       -0.277743   0.553364  -0.502 0.615726
age         -0.007281   0.014894  -0.489 0.624928
```

Too good to be true?

- We can never observe Y^{1M^0}.
- So how come we found a way to estimate $E\left(Y^{1M^0}\right)$?
- It is because of implicit reliance on untestable assumptions.

  (Robins and Richardson, 2010)
- In particular, that the M-Y association is unconfounded.
- The required no unmeasured confounding assumptions are somewhat stronger than ordinarily needed.
- Even if one had experimental data to learn the effect of X on M and of M on Y, confounding adjustment remains needed to combine these effects.
- The required assumptions also make extensions to multiple mediators subtle.

  (VanderWeele T, Vansteelandt S. Mediation analysis with multiple mediators. Epidemiologic methods. 2014 Jan 3;2(1):95-115.)

LONGITUDINAL MEDIATION ANALYSIS

GHENT
UNIVERSITY

THE LEADER TRIAL

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 JULY 28, 2016 VOL. 375 NO. 4

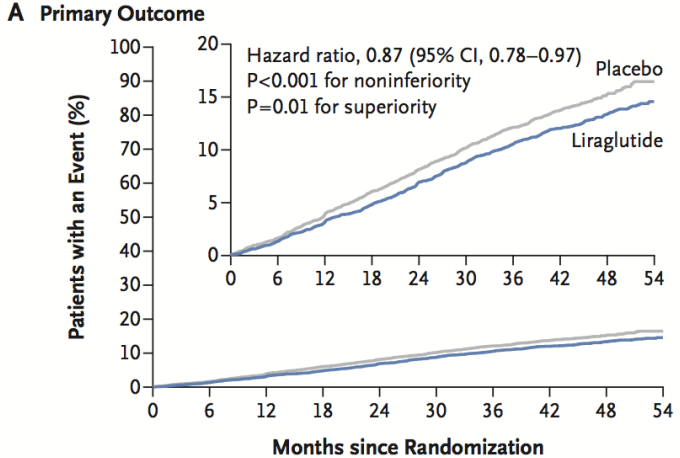
Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes

Steven P. Marso, M.D., Gilbert H. Daniels, M.D., Kirstine Brown-Frandsen, M.D., Peter Kristensen, M.D., E.M.B.A., Johannes F.E. Mann, M.D., Michael A. Nauck, M.D., Steven E. Nissen, M.D., Stuart Pocock, Ph.D., Neil R. Poulter, F.Med.Sci., Lasse S. Ravn, M.D., Ph.D., William M. Steinberg, M.D., Mette Stockner, M.D., Bernard Zinman, M.D., Richard M. Bergenstal, M.D., and John B. Buse, M.D., Ph.D., for the LEADER Steering Committee on behalf of the LEADER Trial Investigators*

DESIGN OF THE LEADER TRIAL

- Population: patients with Type II diabetes and high cardiovascular risk.
- Liraglutide: once-daily injectable drug for the treatment of Type II diabetes, branded as Victoza or Saxenda.
- Primary endpoint: time from randomisation to first MACE (non-fatal stroke, non-fatal myocardial infarction or cardiovascular death).

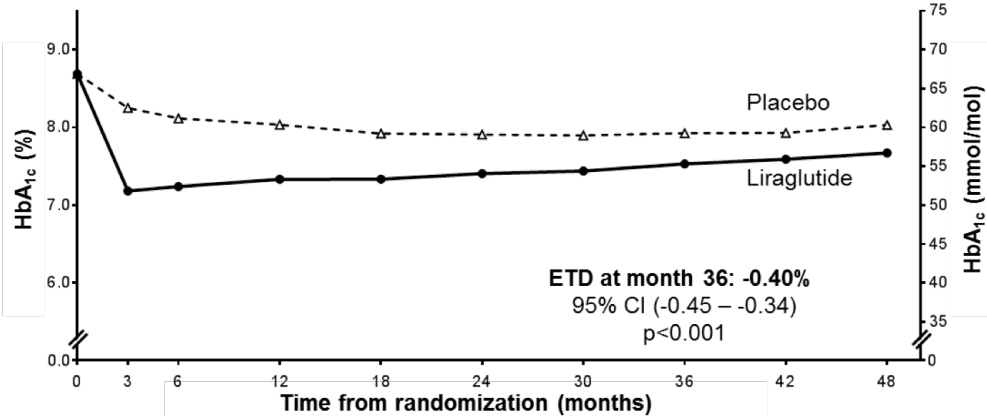
PRIMARY ITT ANALYSIS



Mechanism

- Significant reductions of major cardiovascular events, were also found in SUSTAIN-6 (semaglutide).
- The mechanism is not well understood, however.
- Aim: to develop insight into the precise mechanism whereby liraglutide treatment reduces the risk of MACE.

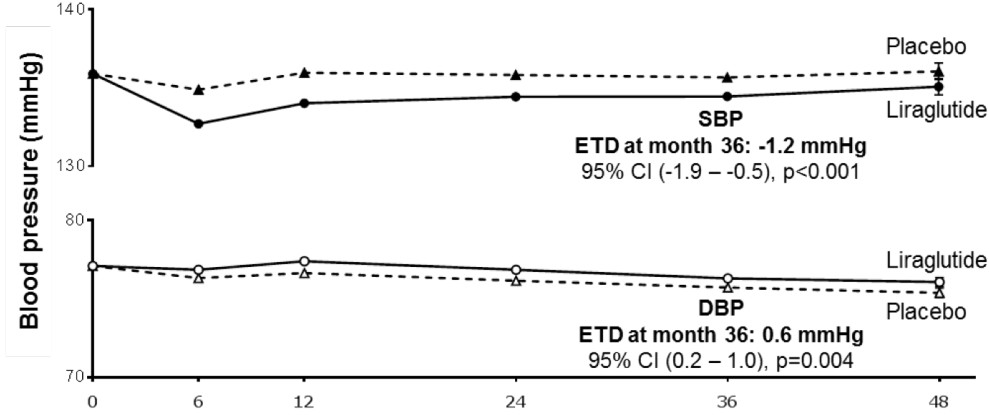
EFFECT ON GLYCATED HEAMOGLOBIN



Number of patients at each visit

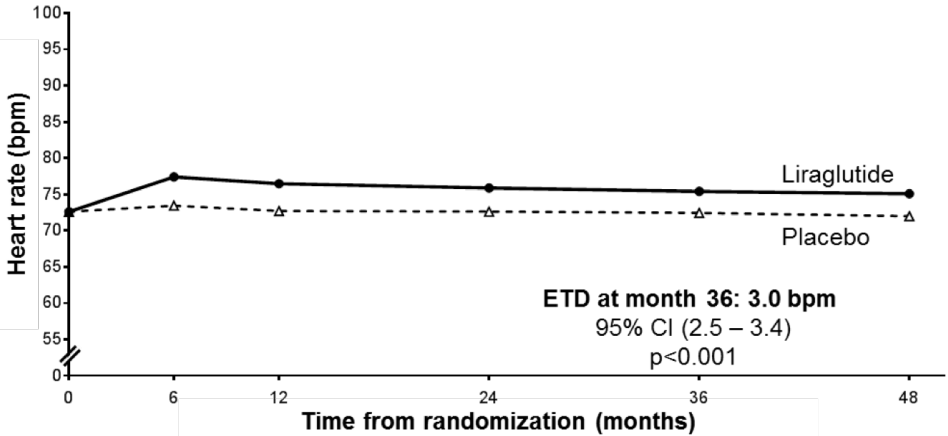
Liraglutide	4668	4402	4355	4295	4135	4034	3877	3810	2349	809
Placebo	4672	4413	4355	4235	4030	3905	3742	3640	2303	756

EFFECT ON BLOOD PRESSURE



Number of patients at each visit						
Liraglutide	4668	4445	4332	4107	3859	823
Placebo	4672	4435	4295	3975	3699	767

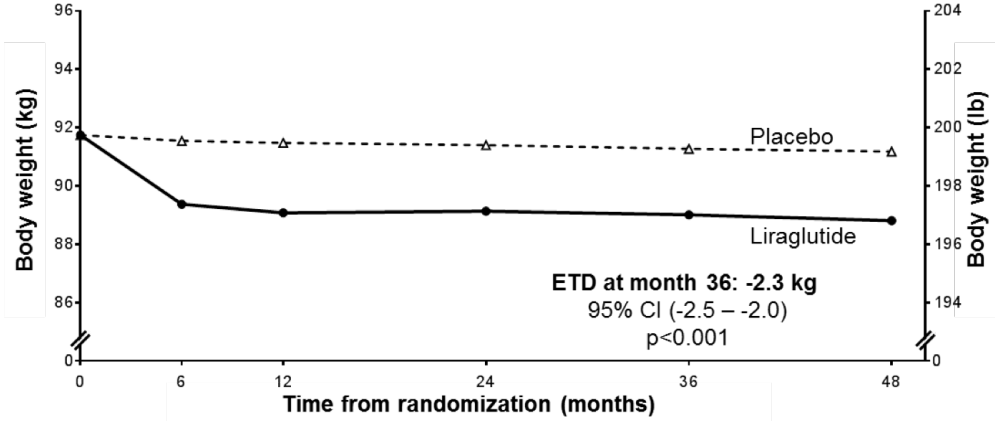
EFFECT ON HEART RATE



Number of patients at each visit

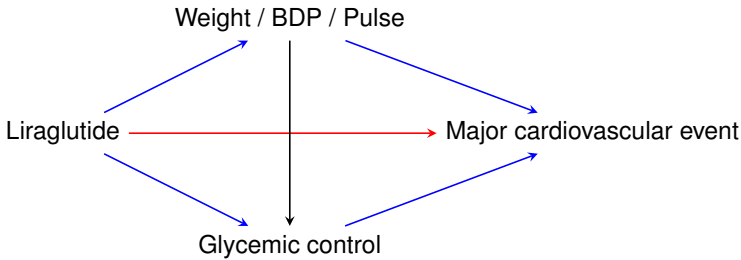
Liraglutide	4668	4442	4330	4099	3853	824
Placebo	4672	4434	4294	3971	3695	767

EFFECT ON BODY WEIGHT



Number of patients at each visit						
Liraglutide	4667	4434	4324	4088	3835	824
Placebo	4671	4423	4285	3970	3680	766

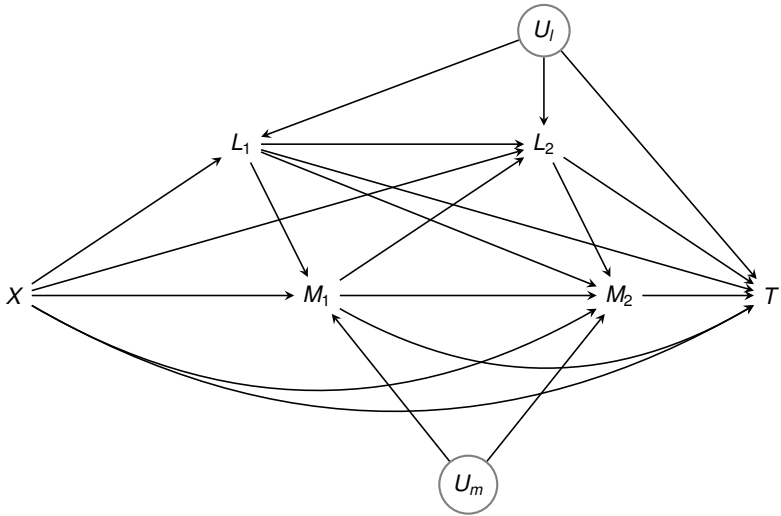
MECHANISM IN CLINICAL TRIALS



Key question

Why does liraglutide reduce the risk of major cardiovascular events?

WE ARE DEALING WITH A COMPLEX STRUCTURE...

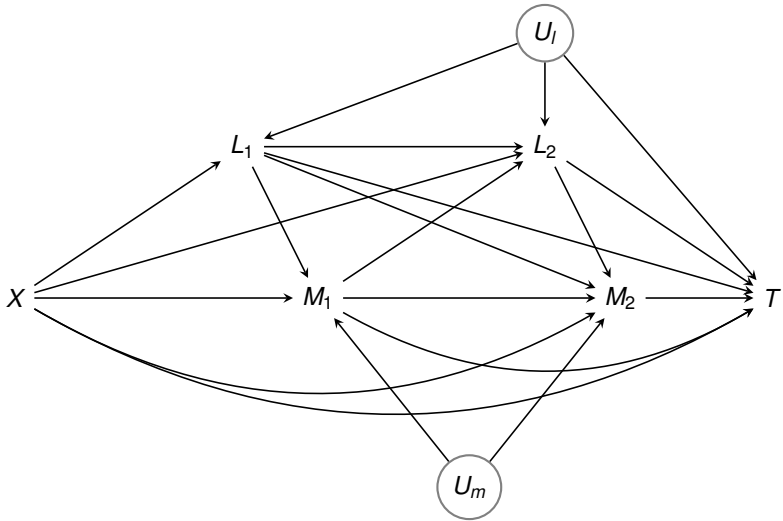


It is common to ignore the longitudinal structure

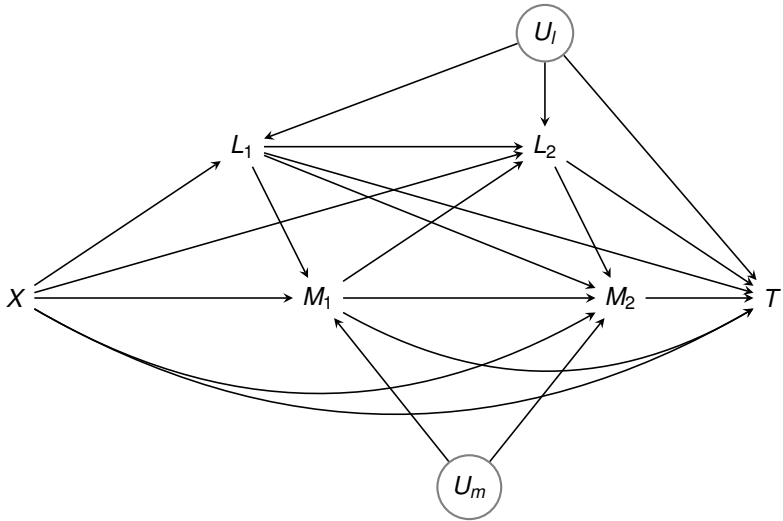
This induces bias.

- When considering the first realisation of the mediator,
  one risks to underestimate the indirect effect, by ignoring later realisations.
- When considering the last realisation or some AUC,
  there is a potential for bias due to reverse causality.
- Some patients experience the event before the mediator is assessed.
- Valid confounding adjustment becomes impossible.

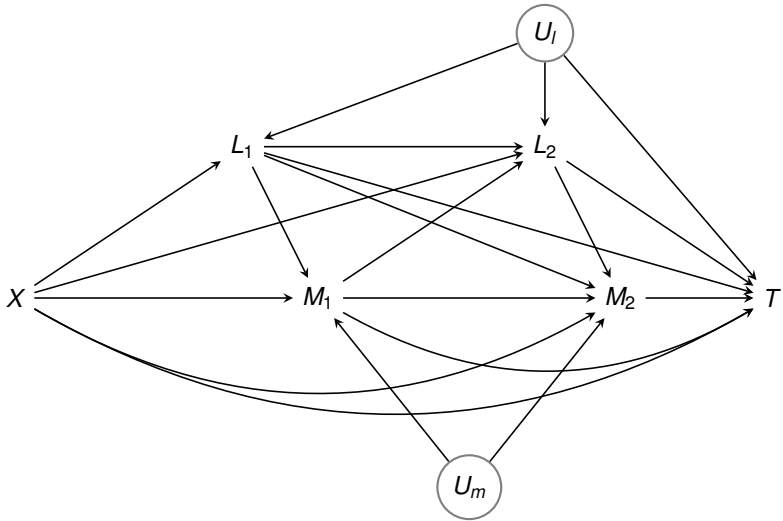
FEATURES: RANDOMISED TREATMENT



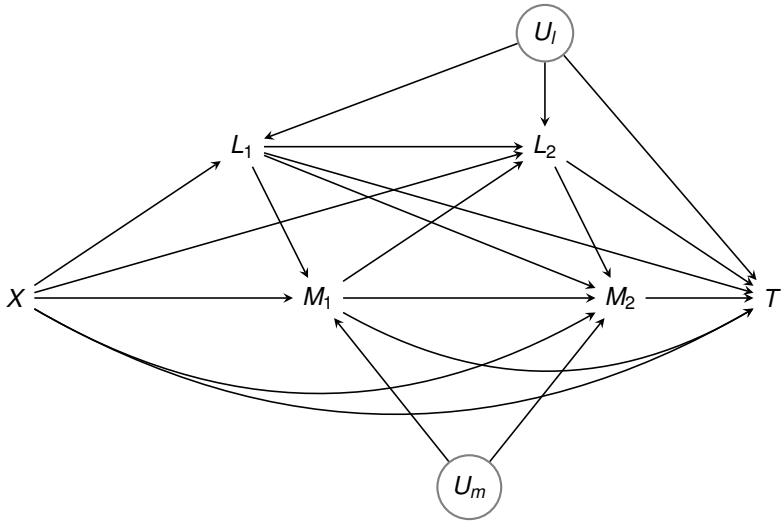
NO CONFOUNDING BY UNMEASURED VARIABLES



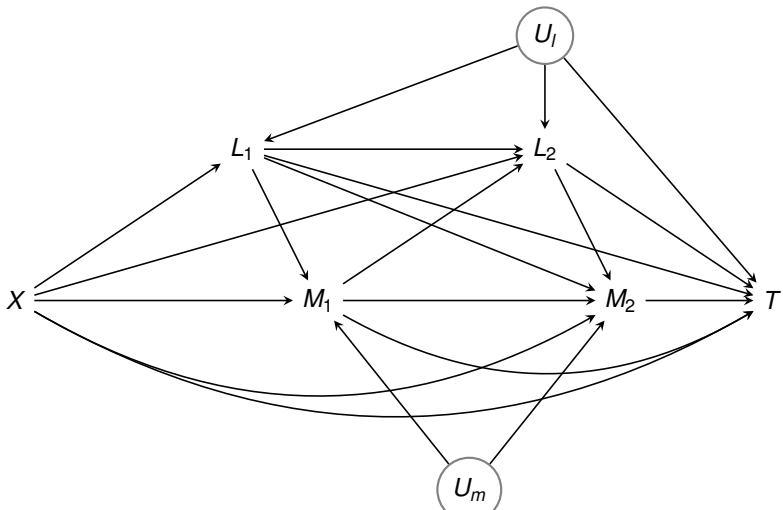
POSSIBILITY OF LAGGED EFFECTS



UNMEASURED COMMON CAUSES OF MEDIATORS/CONFOUNDERS



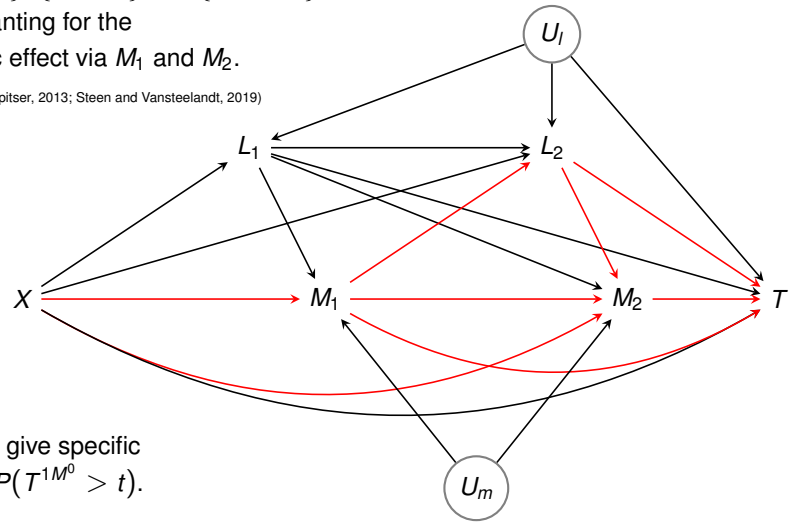
UNMEASURED CONFOUNDING OF COVARIATE - OUTCOME ASSOCIATION



MOST PATHWAYS NOT IDENTIFIED

3 districts $\{X\}$, $\{M_1, M_2\}$ and $\{L_1, L_2, T\}$.
None is recanting for the
path-specific effect via M_1 and M_2.

(Richardson, 2009; Shpitser, 2013; Steen and Vansteelandt, 2019)



These paths give specific
meaning to $P(T^{1M^0} > t)$.

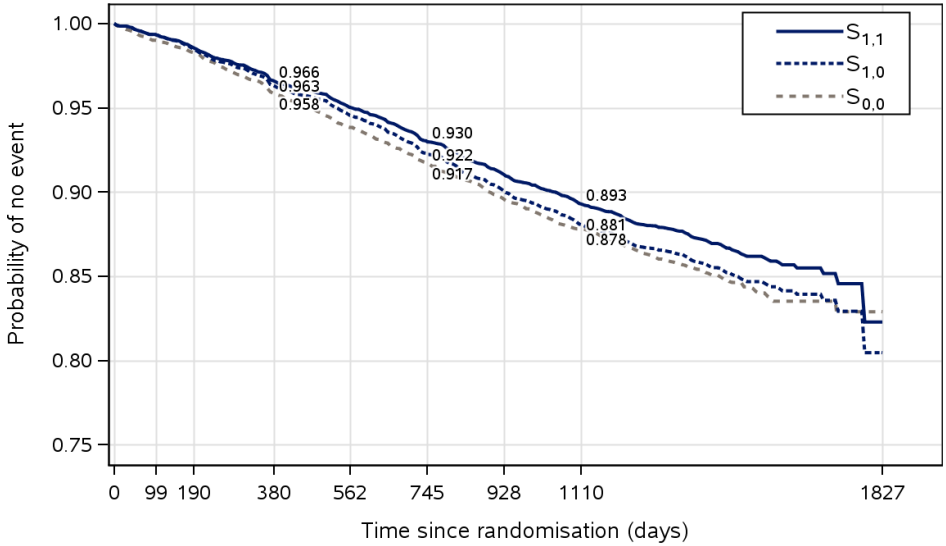
THE MEDIATIONAL G-FORMULA...

... works by simulating how the world would have looked like
under the considered intervention.

Thus to estimate $P\left(T^{1M^0} > t\right)$ for $3 < t \leq 6$:

1. For each patient, based on his baseline covariates L_0,
   predict (randomly) whether he will survive 3 months on treatment.
2. For each patient that was predicted to survive 3 months,
   predict (randomly) his covariate data L_1 on treatment as L_1^1 based on the observed data L_0.
3. For each such patient, predict (randomly) his mediator data M_1 on control as M_1^0
   based on the observed data L_0 and the predicted data L_1^1.
4. For each such patient, predict (randomly) whether he will survive time t on treatment,
   based on the observed data L_0 and the predicted data L_1^1 and M_1^0.
5. We then evaluate the percentage that survived time t.

RESULTS


1.00
0.95
0.90
0.85
0.80
0.75
Probability of no event
0.966
0.963
0.958
0.930
0.922
0.917
0.893
0.881
0.878
$S_{1,1}$
$S_{1,0}$
$S_{0,0}$
0
99
190
380
562
745
928
1110
1827
Time since randomisation (days)

NATURAL EFFECT MODELS

- Alternatively, we can fit natural effect models.

  (Lange, Vansteelandt and Bekaert, 2012; Vansteelandt, Bekaert and Lange, 2012)

- These extend marginal structural models.

  (Hernan, Brumback and Robins, 2001)

- The hazard if 'exposure were set to a'
  and the 'mediators to the level at treatment a^*' can be parameterised using

$$\lambda^{a,a^*}(t) = \lambda_0(t) \exp(\alpha a + \beta a^*) \quad \text{for all } t, a, a^*$$

  (Vo, Davies-Kershaw, Hackett and Vansteelandt, 2020)

- The direct effect is

$$\exp(\alpha) = \frac{\lambda^{1,0}(t)}{\lambda^{0,0}(t)}$$

  and the indirect effect is

$$\exp(\beta) = \frac{\lambda^{1,1}(t)}{\lambda^{1,0}(t)}$$

DUPLICATE - WEIGHT - ESTIMATE

Individual	Start	Stop	Status	A	A^*	M_t	M_{t-1}	L_t	L_{t-1}	L_0
1	0	1	0	1	1	0	0	0	0	l_{01}
1	1	1.5	1	1	1	m_{11}	0	l_{01}	0	l_{01}
2	0	0.9	2	0	0	0	0	0	0	l_{02}
3	0	1	0	1	1	0	0	0	0	l_{03}
3	1	2	0	1	1	m_{13}	0	l_{13}	0	l_{03}
3	2	3	0	1	1	m_{23}	m_{13}	l_{23}	l_{13}	l_{03}
1	0	1	0	1	0	0	0	0	0	l_{01}
1	1	1.5	1	1	0	m_{11}	0	l_{01}	0	l_{01}
2	0	0.9	2	0	1	0	0	0	0	l_{02}
3	0	1	0	1	0	0	0	0	0	l_{03}
3	1	2	0	1	0	m_{13}	0	l_{13}	0	l_{03}
3	2	3	0	1	0	m_{23}	m_{13}	l_{23}	l_{13}	l_{03}

```
> coxph(Surv(Start, Stop, Status) ~ A + A*, weights = w)
```

SUMMARY

GHENT
UNIVERSITY

Summary

- We have gone quite some way in making mediation analyses match the needs that practical applications pose.
- The assumptions are strong, and one must be cautious not to become overly ambitious.
- Currently, there is vigorous research on using machine learning methods to assist mediation analysis.
- Much work remains to be done, both methodologically, as well as on implementation and application.

REFERENCES

Introductory:

Imai K, Keele L and Tingley D. A General Approach to Causal Mediation Analysis. *Psychol Meth* 2010.

Lange, T., Vansteelandt, S. and Bekaert, M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidem* 2012.

Pearl J. Direct and Indirect Effects. In Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence, San Francisco, 2001.

Robins JM and Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992.

VanderWeele TJ (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press.

VanderWeele TJ and Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface* 2009.

VanderWeele TJ and Vansteelandt S. Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *Am J Epidem* 2010.

On multiple mediators:

VanderWeele, T. and Vansteelandt, S. Mediation Analysis with Multiple Mediators. *Epidem Meth* 2013.

Steen, J., Loeys, T., Moerkerke, B. and Vansteelandt, S. Flexible mediation analysis with multiple mediators. *Am J Epidem* 2017.

Vansteelandt, S. and Daniel, R.M. Interventional effects for mediation analysis with multiple mediators. *Epidemiology* 2017.

On longitudinal / survival endpoints:

Vansteelandt, S., Linder, M., Vandenberghe, S., Steen, J. and Madsen, J. Mediation Analysis of Time-to-event Endpoints Accounting for Repeatedly Measured Mediators Subject to Time-varying Confounding. *Statist Med* 2019.

Vo, T. T., Davies-Kershaw, H., Hackett, R., and Vansteelandt, S. Longitudinal mediation analysis of time-to-event endpoints in the presence of competing risks. *Lifetime Data Analysis* 2022.

How to fit natural effect models? (1)

- When $a = a^*$ we have a standard marginal structural Cox model

$$\lambda^{a,a}(t) = \lambda_0(t) \exp\{(\alpha + \beta)a\} \quad \text{for all } t, a$$

- The total effect $\alpha + \beta$ can thus be estimated by fitting this marginal structural Cox model.
- That is, by fitting model

$$\lambda(t) = \lambda_0(t) \exp\{(\alpha + \beta)A\} \quad \text{for all } t$$

to the observed data, inversely weighting the data by

$$W_A = \frac{1}{P(A|L_0)}$$

- Inverse probability of censoring weighting may be needed.

How to fit natural effect models? (2)

- If we weigh subject i's contribution to the risk set at time t by

$$W_M(t) = \prod_{s:t_s \leq t} \frac{P(M_{s,i}|A = 1 - A_i, \overline{M}_{s-1,i}, \overline{L}_{s,i}, T \geq t_s)}{P(M_{s,i}|A = A_i, \overline{M}_{s-1,i}, \overline{L}_{s,i}, T \geq t_s)}$$

  then we effectively change the mediator
  to a draw from the mediator distribution at the opposite exposure level.

- We can thus with the natural effect model with $a^* = 1 - a$,

$$\lambda^{a,1-a}(t) = \lambda_0(t) \exp\{(\alpha - \beta)a + \beta) \quad \text{for all } t, a$$

  by fitting the Cox model

$$\lambda(t) = \lambda_0^*(t) \exp\{(\alpha - \beta)A\} \quad \text{for all } t$$

  to the weighted risk sets, weighting the risk set at time t by

$$W_A \times W_M(t)$$