Direct and indirect effects

Stijn Vansteelandt

*Ghent University, Belgium*

*London School of Hygiene and Tropical Medicine, U.K.*
Mediation analysis

Mediation analysis is one approach towards inferring mechanism: by attempting to disentangle

- **direct effects**: that part of the exposure effect which is not mediated by a given set of potential mediators.
- **indirect / mediated effects**: that part of the exposure effect which is mediated by a given set of potential mediators.

![Mediation diagram]

Mediator M

Exposure X → Outcome Y
The Baron-Kenny approach

- The standard approach, due to Baron and Kenny (1986), focuses on linear models (with independent errors):

\[ Y = \theta_0 + \theta_1 X + \theta_2 M + \epsilon_Y \]
\[ M = \beta_0 + \beta_1 X + \epsilon_M \]

- They interpret \( \theta_1 \) as the direct effect and \( \theta_2 \beta_1 \) as the indirect effect of a unit increase in the exposure w.r.t. mediator \( M \).
Overview

There are broadly two lines of research:

1. **effect decomposition**: *how to decompose a total effect into direct and indirect components?*
   - What exactly do we mean by direct and indirect effect?
   - How to decompose effects in non-linear models?
There are broadly two lines of research:

1. **effect decomposition**: *how to decompose a total effect into direct and indirect components?*
   - What exactly do we mean by direct and indirect effect?
   - How to decompose effects in non-linear models?

2. **confounding**: *how to deal with complex confounding patterns?*
The MIRA trial

Padian et al., Lancet 2007
controlled direct effect (Robins and Greenland, 1992; Pearl, 2001)

The effect of exposure on outcome that would be observed if the mediator were controlled uniformly at a fixed value.
The MIRA trial

controlled direct effect in the absence of condom use

Let $Y(x, m)$ denote the counterfactual HIV status under exposure $X = x$ (1: HIV prevention; 0: control) and frequency of condom use $M = m$.

The difference in HIV risk that we would observe in a randomized microbicide trial if condoms were not available:

$$E\{Y(1, 0) - Y(0, 0)\}$$
Effect decomposition

Controlled direct effects

The MIRA trial

controlled direct effect under a 100% condom use frequency

The difference in HIV risk that we would observe in a randomized microbicide trial if condoms were always used:

$$E\{Y(1, 1) - Y(0, 1)\}$$

This direct effect is likely 0.
Natural direct effects

- In this setting, it is not realistic to think of forcing the mediator to be the same for all subjects.
- **Natural direct effects** (Robins and Greenland, 1992; Pearl, 2001) allow for natural variation in the level of the mediator between subjects.
In this setting, it is not realistic to think of forcing the mediator to be the same for all subjects.

Natural direct effects (Robins and Greenland, 1992; Pearl, 2001) allow for natural variation in the level of the mediator between subjects.

A subject’s natural level of the mediator is taken to be the (counterfactual) value $M(0)$ it would have taken if the exposure were 0.
natural direct effect

The difference in HIV risk that we would observe in a randomized microbicide trial if condom use remained as in the absence of microbicides:

$$E\{ Y(1, M(0)) - Y(0, M(0)) \}$$

It thus roughly expresses what the intention-to-treat effect would have been, had condom use not been affected.
This formalism also enables a meaningful definition of indirect effect.
Natural indirect effects

- This formalism also enables a meaningful definition of indirect effect.
- Robins and Greenland (1992) define the total indirect effect as

$$\text{total effect} - \text{natural direct effect} = E\{Y(1, M(1)) - Y(0, M(0))\} - E\{Y(1, M(0)) - Y(0, M(0))\}$$

No similar result for controlled direct effects.
Natural indirect effects

- This formalism also enables a meaningful definition of indirect effect.
- Robins and Greenland (1992) define the total indirect effect as

\[ \text{total effect} - \text{natural direct effect} = E\{Y(1, M(1)) - Y(0, M(0))\} - E\{Y(1, M(0)) - Y(0, M(0))\} \]

\[ = E\{Y(1, M(1)) - Y(1, M(0))\} \]

- No similar result for controlled direct effects.
Summary: effect decomposition

- Traditional Baron-Kenny approach decomposes total effects into direct and indirect components, but
  - interpretation is vague;
  - there is no natural extension to non-linear models.
Traditional Baron-Kenny approach decomposes total effects into direct and indirect components, but
- interpretation is vague;
- there is no natural extension to non-linear models

Framework of natural direct effects enables effect decomposition **regardless of the data distribution!**


References on effect decomposition and estimation


The standard approach

- From now on: **controlled direct effects**.
- These are commonly inferred by adjusting the association between exposure $X$ and outcome $Y$ for the mediator $M$ (Baron and Kenny, 1986):

$$E(Y|X, M) = \gamma_0 + \gamma_1 X + \gamma_2 M$$
The standard approach

- From now on: **controlled direct effects**.
- These are commonly inferred by adjusting the association between exposure $X$ and outcome $Y$ for the mediator $M$ (Baron and Kenny, 1986):

\[ E(Y|X, M) = \gamma_0 + \gamma_1 X + \gamma_2 M \]

Even when $X$ is randomly assigned, this may introduce a collider-stratification bias.
No unmeasured confounders

We assume that all confounders $L$ for the association between mediator and outcome have been measured.

Additional adjustment for $L$ removes this bias:

$$E(Y|X,M,L) = \gamma_0 + \gamma_1 X + \gamma_2 M + \gamma_3 L$$
It is often realistic to believe that some of those confounders \( L \) are themselves affected by the exposure.

Additional adjustment for \( L \) then continues to introduce bias.
Robins (1999) proposes inverse weighting the data by

\[ \frac{1}{f(M|X, L)} \]
Inverse probability weighted estimation (2)

This removes the association between the mediator and its causes, so that only a direct effect remains.
Confounding

Inverse probability weighted estimation (3)

An estimate of the direct exposure effect $\beta$ may thus be obtained by regressing outcome on exposure and mediator, after weighting each subject by

$$1 \quad \frac{1}{f(M|X, L)}$$

interpretation

Fitting model

$$E(Y|X, M) = \alpha + \beta X + \gamma M$$

after inverse weighting by $1/f(M|X, L)$ yields estimates of the parameters in the marginal structural model (Robins et al., 2000)

$$E\{Y(x, m)\} = \alpha + \beta x + \gamma m$$
Limitations of inverse probability weighting

- IPW estimators may behave *erratically* in finite samples
  - when the mediator $M$ is quantitative;
  - or has strong predictors $X$ and $L$.
- This is because small densities $f(M|X, L)$ can make subjects with weight
  \[
  \frac{1}{f(M|X, L)}
  \]
  highly *influential*.
- In view of this, **G-estimators** have been proposed (Robins, 1994; Goetgeluk, Vansteelandt and Goetghebeur, 2008; Vansteelandt, 2009).
First, remove the indirect effect from the outcome, $Y^* \equiv Y - \hat{\gamma} M$, where $\hat{\gamma}$ is estimate from a regression model

$$E(Y|X, M, L) = \delta_1 + \delta_2 X + \delta_3 L + \gamma M$$
Now only a direct effect remains.
We thus estimate the direct effect parameter $\beta$ by fitting

$$E(Y - \hat{\gamma}M|X) = \alpha + \beta X$$

The resulting parameter $\beta$ can be interpreted as a controlled direct effect:

$$E\{Y(1, m) - Y(0, m)\} = \beta$$
Exposure-mediator interactions

When the model

\[ E \{ Y(x, m) - Y(0, m) | C \} = \beta_1 x + \beta_2 x m \]

is of interest, then we first fit the standard regression model

\[ E(Y | X, M, L, C) = \delta_1 + \delta_2 X + \delta_3 L + \gamma M + \beta_2 XM + \lambda C \]

and next

\[ E(Y - \gamma M - \beta_2 XM | X, C) = \alpha + \beta_1 X \]
Direct effects on the additive hazard scale

- Martinussen et al. (2011) extend G-estimation to additive hazard models.
- Their initial focus is on the difference in hazard functions

\[ \gamma_{X,m}(t) dt = E \left\{ dN_{(1,m)}(t)|F_{(1,m),t} \right\} - E \left\{ dN_{(0,m)}(t)|F_{(0,m),t} \right\}. \]

- This cannot be interpreted as a direct (causal) effect because the two subgroups may not be exchangeable.
Confounding
Survival analysis

Direct effects on the additive hazard scale

- Martinussen et al. (2011) extend G-estimation to additive hazard models.
- Their initial focus is on the difference in hazard functions

\[ \gamma_{X,m}(t) dt = E \left\{ dN_{(1,m)}(t) | F_{(1,m),t} \right\} - E \left\{ dN_{(0,m)}(t) | F_{(0,m),t} \right\}. \]

- This cannot be interpreted as a direct (causal) effect because the two subgroups may not be exchangeable.
- We will therefore define the **controlled (cumulative) direct effect** (Robins and Greenland, 1992; Pearl, 2001) of \( X \) on survival time \( T \) other than through \( M \) as

\[ \Gamma_{X,m}(t) = \int_0^t \gamma_{X,m}(s) \, ds \]
Cumulative direct effect

This encodes a controlled direct effect because

$$\exp \left\{ -\Gamma_{X,m}(t) \right\} = \frac{P\{T(1, m) > t\}}{P\{T(0, m) > t\}}$$

example: MIRA trial

This is the relative risk of avoiding HIV by time $t$ on intervention versus control in the hypothetical situation where male condom use was uniformly kept at level $m$. 
First stage: assess mediator effect

The mediator’s effect on the survival time can be obtained from a standard Aalen additive hazards analysis (Aalen, 1989)

\[ E \{ dN(t) | \mathcal{F}_t, X, M, L \} = \{ \psi_0(t) + \psi_X(t)X + \psi_M(t)M + \psi_L(t)L \} R(t) dt \]
Second stage: remove mediator effect

- We will now correct the event time by removing the mediator effect.
- This requires correcting the increment $dN(t)$ as well as the risk set $R(t)$ at each time $t$. 
Second stage: remove mediator effect

We will now correct the event time by removing the mediator effect.
This requires correcting the increment $dN(t)$ as well as the risk set $R(t)$ at each time $t$.
The correction in the increment is achieved by substituting $dN(t)$ with

$$dN(t) - \psi_M(t)Mdt$$
Second stage: remove mediator effect

- We will now correct the event time by removing the mediator effect.
- This requires correcting the increment $dN(t)$ as well as the risk set $R(t)$ at each time $t$.
- The correction in the increment is achieved by substituting $dN(t)$ with
  
  $$dN(t) - \psi_M(t) M dt$$

- The correction in the risk set is achieved by substituting $R(t)$ with
  
  $$R(t) \exp \left\{ M \int_0^t \psi_M(s) ds \right\}$$
Third stage: estimate total effect on corrected counting process

From this, it can be shown that

\[
\left( \frac{1}{X} \right) R(t) \exp \left\{ M \int_0^t \psi_M(s) ds \right\}
\]

modification of risk set

\[
\times \left\{ dN(t) - M \psi_M(t) dt - \gamma_0(t) dt - X \gamma_X(t) dt \right\}
\]

residual

is an unbiased estimating function.

From this, a closed form estimator is obtained.
Application to Danish 1905 cohort

- Goal: direct effect of carrying apoe4 mutation on survival other than through activity of daily living.
- Intermediate confounding by cognitive functioning.
Direct cumulative effect of carrying apoe4 mutation on survival other than through activity of daily living
Confounding

- Traditional Baron-Kenny approach ignores confounding of the association between mediator and outcome.
- When, as often, such confounders are themselves affected by the exposure, standard regression methods are no longer applicable.
- Other ‘manipulations’ of causal diagrams needed.
- We have discussed 2 ‘generic’ approaches for controlled direct effects.
Confounding

- Traditional Baron-Kenny approach ignores confounding of the association between mediator and outcome.
- When, as often, such confounders are themselves affected by the exposure, standard regression methods are no longer applicable.
- Other ‘manipulations’ of causal diagrams needed.
- We have discussed 2 ‘generic’ approaches for controlled direct effects.
- **Inverse probability weighting** works by removing the association between mediator and exposure, and thus removing the indirect effect from the data.
- This approach works for any outcome type, but is essentially limited to discrete mediators.
G-estimation works by removing the effect of mediator on outcome, and thus also removing the indirect effect from the data.

This approach works for any mediator type and is much more powerful than inverse probability weighting, but cannot handle any type of outcome:

- linear models for continuous outcomes (Vansteelandt, 2009);
- log-linear models for positive-constrained outcomes (Vansteelandt, 2009);
- logistic models for dichotomous outcomes (Vansteelandt, 2010);
- additive hazard models for survival times (Martinussen et al., 2011).
References on intermediate confounding


Vansteelandt S. Estimation of controlled direct effects on a dichotomous outcome using logistic structural direct effect models. *Biometrika* 2010; 97, 921-934.