Mediation analysis with SEM or causal inference: Where is the difference?

Bianca De Stavola and Rhian Daniel

London School of Hygiene and Tropical Medicine, UK

3rd UK Mplus Conference
May 24, 2012
Mediation concerns the extent to which the effect of one variable on another is mediated by some intermediate variable.
Mediation concerns the extent to which the effect of one variable on another is mediated by some intermediate variable.

The focus is therefore on understanding mechanisms, assuming causality from exposure to outcome via the mediator.
The study of mediation

Two main strands for the study of mediation:

- **Social sciences / psychometrics** (Baron and Kenny, 1986):
  - direct/indirect effects defined for a given model
  - mostly estimated within the SEM framework but also with a sequence of regression models

- **Causal inference literature** (Robins and Greenland, 1992; Pearl, 2001):
  - based on *counterfactuals* and leading to general definitions of direct/indirect effects
  - estimated using ‘novel’ parametric and semi-parametric methods (G-computation, *etc.*)
Outline

1. Introduction
2. Mediation within the SEM framework
3. Causal inference framework
4. Mediation in causal inference
5. Bridging the two frameworks
6. Summary
Consider this general set-up, with exposure $X$, mediator $M$, outcome $Y$, and confounders $C$ and $L$:

\[ \text{If } Y \text{ and } M \text{ are continuous variables...} \]
Consider this general set-up, with exposure $X$, mediator $M$, outcome $Y$, and confounders $C$ and $L$:

If $Y$ and $M$ are continuous variables . . .
A Linear Structural Equation Model (LSEM) (1)

...we could specify a LSEM corresponding to this diagram:

\[
\begin{align*}
  m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
  y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 c_i + \beta_4 l_i + \epsilon_{2i}
\end{align*}
\]

\[\epsilon_{1i} \sim \text{IIN}(0, \sigma_1^2) \text{ and } \epsilon_{2i} \sim \text{IIN}(0, \sigma_2^2), \text{ } \epsilon_{1i} \text{ and } \epsilon_{2i} \text{ uncorrelated with each other and the variables in their equation,} \]

\[i = 1, \ldots, N.\]

If the model correctly specified ...
A Linear Structural Equation Model (LSEM) (1)

...we could specify a LSEM corresponding to this diagram:

\[
\begin{align*}
   m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
   y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 c_i + \beta_4 l_i + \epsilon_{2i}
\end{align*}
\]

\(\epsilon_{1i} \sim \text{II}(0, \sigma_1^2)\) and \(\epsilon_{2i} \sim \text{II}(0, \sigma_2^2)\), \(\epsilon_{1i}\) and \(\epsilon_{2i}\) uncorrelated with each other and the variables in their equation, 

\(i = 1, \ldots, N\).

If the model correctly specified...
A Linear Structural Equation Model (LSEM) (2)

If the model correctly specified:

\[
\begin{align*}
    m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
    y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 c_i + \beta_4 l_i + \epsilon_{2i}
\end{align*}
\]

- **direct effect of** $X$ **on** $Y$: $\beta_1$
- **indirect effect of** $X$ **on** $Y$: $\alpha_1 \times \beta_2$.

(‘Product of coefficients method’)

Bianca De Stavola & Rhian Daniel/Mediation · 24 May 2012
a) These definitions are specific to this model

b) For a meaningful interpretation they depend on the assumption that:
   - there are no unaccounted confounders
   - the parametric model is correctly specified (e.g. effect of $M$ is linear, no $X - M$ interaction, etc.)

   *These assumptions however are rarely discussed.*

c) There are no equivalent results for binary/categorical outcomes and/or mediators [although see Muthén 2011]

d) The linear nature of the model is key and very restrictive.
The causal inference framework

- In this framework, definitions of direct and indirect effects are general: they do not depend on the specification of a particular statistical model.
- There are several alternative definitions.
- Today focus on:
  - controlled direct effect (CDE)
  - pure natural direct effect (PNDE)
- **However**, they involve quantities that are not all observable: *potential outcomes* and the *potential mediators*. 
Potential outcomes

- $Y(x)$:
  the potential values of $Y$ that would have occurred had $X$ been set, possibly counter to fact, to the value $x$.

- $M(x)$:
  the potential values of $M$ that would have occurred had $X$ been set, possibly counter to fact, to the value $x$.

- $Y(x, m)$:
  the potential values of $Y$ that would have occurred had $X$ been set, possibly counter to fact, to the value $x$ and $M$ to $m$.

For simplicity consider the case where $X$ is binary.
It also helps to start with the definition of total causal effect.
The average **total causal effect** of $X$, comparing exposure level $X = 1$ to $X = 0$, can be defined as the linear contrast

$$TCE = E[Y(1)] - E[Y(0)]$$

This is a comparison of two hypothetical worlds: in the first, $X$ is set to 1, and in the second $X$ is set to 0.

Note that $TCE = E[Y(1)] - E[Y(0)] \neq E[Y|X = 1] - E[Y|X = 0]$. 
To identify $TCE$ we need to be able to infer $E[Y(1)]$ and $E[Y(0)]$ from the observed data. This is possible if these assumptions are satisfied:

- **consistency**: $Y(x) = Y$ when $X = x$
- **conditional exchangeability**: $Y(x)$ can be inferred from $Y(x)$ of comparable others when $X \neq x$, i.e. there is no unmeasured confounding between $X$ and $Y$.

If so, then

$$TCE = E[Y(1) - Y(0)] = E \left\{ E[Y(1) - Y(0)|C] \right\}$$

$$= \int \{ E[Y|X = 1, C = c] - E[Y|X = 0, C = c] \} f_C(c) \, dc$$

where $C$ is the vector of confounders.
The average **controlled direct effect** of $X$ on $Y$ when $M$ is controlled at $m$ is

\[
 CDE(m) = E[Y(1, m)] - E[Y(0, m)]
\]

This is a comparison of two hypothetical worlds:

- In the first, $X$ is set to 1, and in the second $X$ is set to 0.
- In both worlds, $M$ is set to $m$.
- By keeping $M$ fixed at $m$, we are getting at the direct effect of $X$, unmediated by $M$.
- In general $CDE(m)$ varies with $m$. 

**Intro Mediation: SEMs Causal inference framework**
**Mediation: causal inference**
**Bridging the two**
**Summary**

Mediation in Causal inference (1)
Controlled Direct Effect (CDE)
The average **controlled direct effect** of $X$ on $Y$ when $M$ is controlled at $m$ is

$$CDE(m) = E[Y(1, m)] - E[Y(0, m)]$$

This is a comparison of two hypothetical worlds:

- In the first, $X$ is set to 1, and in the second $X$ is set to 0.
- In both worlds, $M$ is set to $m$.
- By keeping $M$ fixed at $m$, we are getting at the direct effect of $X$, unmediated by $M$.
- In general $CDE(m)$ varies with $m$. 

---

Bianca De Stavola & Rhian Daniel/Mediation · 24 May 2012
The average pure natural direct effect of $X$ on $Y$ is

$$PNDE = E[Y(1, M(0))] - E[Y(0, M(0))]$$

This is a comparison of two hypothetical worlds:

- In the first, $X$ is set to 1, and in the second $X$ is set to 0.
- In both worlds, $M$ is set to the natural value $M(0)$, i.e. the value it would take if $X$ were set to 0 (0 being pure).
- Since $M$ is the same (within individual) in both worlds, we are still getting at the direct effect of $X$, unmediated by $M$. 
Controlled or natural?

- The controlled direct effect:
  - requires fewer assumptions to be identified (see next slides)
  - is said to be prescriptive,
  - useful for assessing interventions

- The pure natural direct effect:
  - is said to be descriptive (since $M$ takes naturally occurring levels under $X = 0$)
  - leads to defining the average total natural indirect effect of $X$ on $Y$:

$$TNIE = TCE - PNDE = E[Y(1, M(1))] - E[Y(1, M(0))]$$
Identification of $CDE$, $PNDE$ and $TNIE$ requires certain assumptions. These include as before appropriate specifications of:

(i) **Consistency** assumption
(ii) **Conditional exchangeability** assumption

For the *natural* effects, $PNDE$ and $TNIE$, we need an additional assumption.
Identifying assumptions for \( CDE \)

\[
CDE(m) = E[Y(1,m)] - E[Y(0,m)]
\]

The two assumptions are specified as:

(i) **Consistency:** \( Y(x, m) = Y \) when \( X = x \) and \( M = m \)

(ii) **Conditional exchangeability**, loosely:

- no unmeasured confounding of \( X - Y \)
- no unmeasured confounding of \( M - Y \)

(In this context: **Sequential Ignorability** assumption)

If these are satisfied then

\[
CDE(m) = E[Y(1,m) - Y(0,m)]
= \int E[Y|X = 1, M = m, C = c, L = l] f_L|C,X (l|c, 1) f_C(c) \, dl \, dc
- \int E[Y|X = 0, M = m, C = c, L = l] f_L|C,X (l|c, 0) f_C(c) \, dl \, dc
\]
Identifying assumptions for **PNDE** and **TNIE** (1)

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

The two assumptions are specified as:

(i) **Strong consistency:**
- \( Y(x, m) = Y \) when \( X = x \) and \( M = m \) and \( M(x) = M \) when \( M = m \)
- \( Y(x, M(x)) = Y \) when \( X = x \)

(ii) **Conditional exchangeability:**
- no unmeasured confounding of \( X - Y \)
- no unmeasured confounding of \( M - Y \)
- no unmeasured confounding of \( X - M \)

*required because calculations involve \( M(0) \)*

Additionally, either of two additional assumptions is required . . .
Additionally, either of two additional assumptions is required:

(iii) No intermediate confounders, i.e., no $M - Y$ confounders $L$ are affected by $X$

(iv) Some restrictions on $X - M$ interactions in effect on $Y$. ¹

If these are satisfied then,

$$PNDE = E[Y(1, M(0)) - Y(0, M(0))]$$

$$= \int \{E[Y|X = 1, M = m, C = c, L = l] - E[Y|X = 0, M = m, C = c, L = l]\}$$

$$\times f_{M|C,X,L}(m|c,0,l)f_{L|C}(l|c)f_C(c) \, dm \, dl \, dc$$

¹ This has alternative definitions in the literature
There is a wide range of options for estimation of $CDE$, $PNDE$ and $TNIE$, for most combinations of $M$ and $Y$:

- **G-computation**—very flexible and efficient but heavy on parametric modelling assumptions (not robust).

- **Inverse probability of treatment weighting (IPTW)**—fewer modelling assumptions so more robust, less efficient. Not practical when $M$ is continuous.

- **G-estimation**—a half-way house in terms of robustness-efficiency trade-off, but more complex to understand.
Muthén (2011) shows that, if we apply these definitions to our LSEM:

\[
\begin{align*}
  m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
  y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 c_i + \beta_4 l_i + \epsilon_{2i}
\end{align*}
\]

We find that (if this model is correct):

- If consistency and conditional exchangeability and no intermediate confounders hold, then

  \[ CDE(m) = \beta_1 \]

- If strong consistency, conditional exchangeability and no intermediate confounders hold, then

  \[ PNDE = \beta_1 \quad TNIE = \alpha_1 \beta_2 \]

- This can be extended to more general settings.
Bridging the two frameworks

Muthén (2011) shows that, if we apply these definitions to our LSEM:

\[
\begin{align*}
    m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
    y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 c_i + \beta_4 l_i + \epsilon_{2i}
\end{align*}
\]

We find that (if this model is correct):

- If consistency and conditional exchangeability and no intermediate confounders hold, then
  \[CDE(m) = \beta_1\]

- If strong consistency, conditional exchangeability and no intermediate confounders hold, then
  \[PNDE = \beta_1 \quad TNIE = \alpha_1 \beta_2\]

- This can be extended to more general settings.
If the model is:

\[
\begin{align*}
    m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
    y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 x_i m_i + \beta_4 c_i + \beta_5 l_i + \epsilon_{2i}
\end{align*}
\]

Then, applying the formal definitions, under the appropriate assumptions,

\[
\begin{align*}
    CDE(m) &= \beta_1 + \beta_3 m \\
    PNDE &= \beta_1 + \beta_3 \alpha_0 \\
    TNIE &= \beta_2 \alpha_1 + \beta_3 \alpha_1
\end{align*}
\]
If the model is:

\[
\begin{align*}
    m_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 l_i + \epsilon_{1i} \\
    y_i &= \beta_0 + \beta_1 x_i + \beta_2 m_i + \beta_3 c_i + \beta_4 l_i + \epsilon_{2i} \\
    l_i &= \delta_0 + \delta_1 x_i + \epsilon_{3i}
\end{align*}
\]

Then, applying the formal definitions, under the appropriate assumptions,

\[
\begin{align*}
    CDE(m) &= \beta_1 + \delta_1 \beta_4 \\
    PNDE &= \beta_1 + \delta_1 \beta_4 \\
    TNIE &= \beta_2 \alpha_1 + \delta_1 \beta_2 \alpha_2
\end{align*}
\]
‘Estimation by combination’

These results give us another approach to estimate the causally-defined direct/indirect effects.

As shown by Muthén (2011)

ML (or otherwise) estimates of the LSEM parameters can be combined to obtain estimates of the relevant mediation estimands (with SEs derived using the delta method).

However:

- approach is restricted to settings with continuous $Y$, $M$ (and $L$),
- although Muthén (2011) extended derivation for non-continuous $Y$ and $M$ the approach is still in its infancy.
These results give us another approach to estimate the causally-defined direct/indirect effects.

As shown by Muthén (2011)

ML (or otherwise) estimates of the LSEM parameters can be combined to obtain estimates of the relevant mediation estimands (with SEs derived using the delta method).

However:

- approach is restricted to settings with continuous $Y$, $M$ (and $L$),
- although Muthén (2011) extended derivation for non-continuous $Y$ and $M$ the approach is still in its infancy
Comparison by simulation: $X - M$ interaction

$M$ and $Y$ Normal, $X$ binary, $X - M$ interaction, N=10,000:

<table>
<thead>
<tr>
<th>Method</th>
<th>CDE(0)</th>
<th>PNDE</th>
<th>TNIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>0.40</td>
<td>0.80</td>
<td>0.35</td>
</tr>
<tr>
<td>Muthén</td>
<td>0.378</td>
<td>0.804</td>
<td>0.358</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.022)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>G-comp</td>
<td>0.378</td>
<td>0.805</td>
<td>0.360</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.023)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>G-estim</td>
<td>0.378</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
### Comparison by simulation: intermediate confounder

$L, M$ and $Y$ Normal, $X$ binary, intermediate confounder, $N=10,000$:

<table>
<thead>
<tr>
<th>Method</th>
<th>CDE(0)</th>
<th>PNDE</th>
<th>NTIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>0.425</td>
<td>0.425</td>
<td>0.281</td>
</tr>
<tr>
<td>Muthén</td>
<td>0.415</td>
<td>0.415</td>
<td>0.301</td>
</tr>
<tr>
<td>G-comp</td>
<td>0.412</td>
<td>0.412</td>
<td>0.305</td>
</tr>
<tr>
<td>G-estim</td>
<td>0.415</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.012)</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.023)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>
Two main strands for the study of mediation:

- **Social sciences / psychometrics:**
  - direct/indirect effects defined for a given model
  - estimated within the SEM framework

- **Causal inference literature:**
  - based on *counterfactuals* and leading to general definitions of direct/indirect effects
  - estimated using ‘novel’ parametric and semi-parametric methods (G-computation, *etc.*)
  - with continuous $Y$ and $M$ can also be estimated within the SEM framework (even in the presence of $X - M$ interaction, intermediate confounding, *etc.*).

- Other estimation approaches are more flexible
Final comments

- The choice of the causally-defined effect of interest is crucial: controlled or natural?

- As Muthén (2011) said
  
  "To claim that effects are causal, it is not sufficient to use causally defined effects."

- Their identification requires stringent, unverifiable, assumptions

- Hence need for sensitivity analyses (Imai et al, 2010).
References


Mplus code

MODEL:
[L*0] (delta0);
L*1;
L ON T*.5 (delta1);
m*1;
[m*2] (gamma0);
m ON T*.5 (gamma1);
m ON L*.25 (gamma2);
[Y*1] (beta0);
Y*.5;
Y ON M*.5 (beta1);
Y ON T*.4 (beta2);
Y ON L*.2 (beta4);
MODEL CONSTRAINT:
NEW(pie*.25 pde*.8);
pie=beta1*gamma1+beta1*delta1*gamma2;
pde=beta2+beta4*delta1;