Estimating Causal Mediation Effects from a Single Regression Model

Jeffrey D. Blume, PhD

Joint work with, and slides courtesy of,

Christina T. Saunders, PhD
Berry Consultants

VANDERBILT UNIVERSITY

Noevember, 2019

Slide Appendix

System congruency

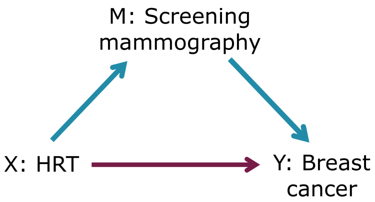
Suppose one believes M depends on X quadratically and writes the following system:

$$\mathbb{E}[Y|X,M] = \beta_0 + \beta_X X + \beta_M M$$
$$\mathbb{E}[M|X] = \alpha_0 + \alpha_X X + \alpha_{X^2} X^2$$
$$\mathbb{E}[Y|X] = \beta_0^* + \beta_X^* X$$

- This system is not congruous because marginalizing over the Full outcome model $\mathbb{E}_{M|X}[Y|X,M]$ does not yield the specified marginal model $\mathbb{E}[Y|X]$
- A properly specified system of equations would include an X^2 term in the third equation
- The difference and product approaches are not comparable because they are actually using *different systems* to estimate mediation effects

Usefulness of natural direct and indirect effects

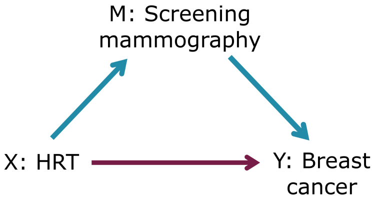
Suppose we are interested in estimating the mediating role of screening mammography in the relation between HRT and breast cancer.



- More breast cancer is observed among women taking hormone replacement therapy (HRT).
- Some portion of the excess cases is due to increased detection.

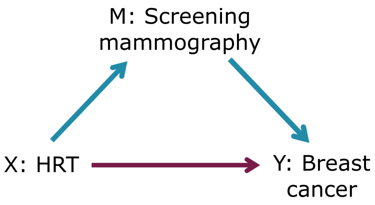
  But how much?

Usefulness of natural direct and indirect effects



The NIE tells us the breast cancer risk for a woman receiving HRT under the mammographic screening she received under HRT versus the risk under the screening *she would have had under no HRT*.

The controlled direct effect (CDE) and portion eliminated (PE)



- The CDE $= Y(x,m) - Y(x_\circ, m)$ corresponds to the effect of HRT on breast cancer that would remain after an intervention that sets mammography screening to a specific level m.
- The PE corresponds to the effect of HRT on breast cancer that would be eliminated after an intervention that sets mammography screening to a specific level.

The single-model approach

From the Full model,

$$\begin{bmatrix} \hat{\boldsymbol{\beta}}_{\boldsymbol{X}} \\ \hat{\boldsymbol{\beta}}_{\boldsymbol{M}} \end{bmatrix} \sim \mathsf{MVN}_k \left(\begin{bmatrix} \boldsymbol{\beta}_{\boldsymbol{X}} \\ \boldsymbol{\beta}_{\boldsymbol{M}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{V}_{\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{XM}} \\ \boldsymbol{V}_{\boldsymbol{MX}} & \boldsymbol{V}_{\boldsymbol{M}} \end{bmatrix} \right)$$

The single-model approach

From the Full model,

$$\begin{bmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{X}}\\ \hat{\boldsymbol{\beta}}_{\boldsymbol{M}}\end{bmatrix} \sim \mathsf{MVN}_k\left(\begin{bmatrix}\boldsymbol{\beta}_{\boldsymbol{X}}\\ \boldsymbol{\beta}_{\boldsymbol{M}}\end{bmatrix}, \begin{bmatrix}\boldsymbol{V}_{\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{X}\boldsymbol{M}}\\ \boldsymbol{V}_{\boldsymbol{M}\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{M}}\end{bmatrix}\right)$$

From multivariate normal theory,

$$\begin{aligned}(\hat{\boldsymbol{\beta}}_{\boldsymbol{X}}|\hat{\boldsymbol{\beta}}_{\boldsymbol{M}} = 0) \sim \mathsf{MVN}_p(&\boldsymbol{\beta}^*_{\boldsymbol{X}}, \boldsymbol{V}^*_{\boldsymbol{X}})\\ \boldsymbol{\beta}^*_{\boldsymbol{X}} &= \boldsymbol{\beta}_{\boldsymbol{X}} - \boldsymbol{V}_{\boldsymbol{X}\boldsymbol{M}}\boldsymbol{V}_{\boldsymbol{M}}^{-\mathbf{1}}\boldsymbol{\beta}_{\boldsymbol{M}}\end{aligned}$$

The single-model approach

From the Full model,

$$\begin{bmatrix} \hat{\boldsymbol{\beta}}_{\boldsymbol{X}} \\ \hat{\boldsymbol{\beta}}_{\boldsymbol{M}} \end{bmatrix} \sim \mathsf{MVN}_k \left(\begin{bmatrix} \boldsymbol{\beta}_{\boldsymbol{X}} \\ \boldsymbol{\beta}_{\boldsymbol{M}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{V}_{\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{X}M} \\ \boldsymbol{V}_{M\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{M}} \end{bmatrix} \right)$$

From multivariate normal theory,

$$\begin{aligned} (\hat{\boldsymbol{\beta}}_{\boldsymbol{X}} | \hat{\boldsymbol{\beta}}_{\boldsymbol{M}} = 0) &\sim \mathsf{MVN}_p(\boldsymbol{\beta}^*_{\boldsymbol{X}}, \boldsymbol{V}^*_{\boldsymbol{X}}) \\ \boldsymbol{\beta}^*_{\boldsymbol{X}} &= \boldsymbol{\beta}_{\boldsymbol{X}} - \boldsymbol{V}_{\boldsymbol{X}M} \boldsymbol{V}_M^{-\mathbf{1}} \boldsymbol{\beta}_{\boldsymbol{M}} \\ \Delta = \boldsymbol{\beta}^*_{\boldsymbol{X}} - \boldsymbol{\beta}_{\boldsymbol{X}} &= -\boldsymbol{V}_{\boldsymbol{X}M} \boldsymbol{V}_M^{-\mathbf{1}} \boldsymbol{\beta}_{\boldsymbol{M}} \end{aligned}$$

The single-model approach

From the Full model,

$$\begin{bmatrix} \hat{\boldsymbol{\beta}}_{\boldsymbol{X}} \\ \hat{\boldsymbol{\beta}}_{\boldsymbol{M}} \end{bmatrix} \sim \mathsf{MVN}_k\left(\begin{bmatrix} \boldsymbol{\beta}_{\boldsymbol{X}} \\ \boldsymbol{\beta}_{\boldsymbol{M}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{V}_{\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{X}\boldsymbol{M}} \\ \boldsymbol{V}_{\boldsymbol{M}\boldsymbol{X}} & \boldsymbol{V}_{\boldsymbol{M}} \end{bmatrix} \right)$$

From multivariate normal theory,

$$\begin{aligned} (\hat{\boldsymbol{\beta}}_{\boldsymbol{X}} | \hat{\boldsymbol{\beta}}_{\boldsymbol{M}} = 0) &\sim \mathsf{MVN}_p(\boldsymbol{\beta}^*_{\boldsymbol{X}}, \boldsymbol{V}^*_{\boldsymbol{X}}) \\ \boldsymbol{\beta}^*_{\boldsymbol{X}} &= \boldsymbol{\beta}_{\boldsymbol{X}} - \boldsymbol{V}_{\boldsymbol{X}\boldsymbol{M}} \boldsymbol{V}_{\boldsymbol{M}}^{-1} \boldsymbol{\beta}_{\boldsymbol{M}} \\ \Delta = \boldsymbol{\beta}^*_{\boldsymbol{X}} - \boldsymbol{\beta}_{\boldsymbol{X}} &= -\boldsymbol{V}_{\boldsymbol{X}\boldsymbol{M}} \boldsymbol{V}_{\boldsymbol{M}}^{-1} \boldsymbol{\beta}_{\boldsymbol{M}} \end{aligned}$$

Using properties of the sampling distribution, we see that the difference of coefficients Δ equals some product of coefficients.

Empiric vs implied total effect

Once the Full outcome model and the Mediator model are specified, there is one theoretical Reduced model; however, there can be several mediator models that imply *the same form* for the Reduced model.

(F)	$\mathbb{E}[Y\|X,M,W]$	$= \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW$
(M)	(a) $\mathbb{E}[M\|X,W]$	$= \alpha_0 + \alpha_X X + \alpha_W W$
	(b) $\mathbb{E}[M\|X,W]$	$= \delta_0 + \delta_X X + \delta_W W + \delta_{XW} XW$
(R)	$\mathbb{E}[Y\|X,W]$	$= \gamma_0 + \gamma_X X + \gamma_W W + \gamma_{XW} XW$

Both marginal models have the same form. As a result, the *fitted* estimates of the marginal model effects $\hat{\gamma}$ will be the same.

The implied coefficients from the two marginal models are different:

(a): $(\beta_0 + \alpha_0\beta_M) + (\beta_X + \alpha_X\beta_M)X + (\beta_W + \alpha_W\beta_M)W + \beta_{XW}XW$
(b): $(\beta_0 + \delta_0\beta_M) + (\beta_X + \delta_X\beta_M)X + (\beta_W + \delta_W\beta_M)W + (\beta_{XW} + \delta_{XW}\beta_M)XW$

The mediation formula

$$\begin{aligned}\mathsf{NDE} &= \mathbb{E}[Y(x, M(x_\circ)) - Y(x_\circ, M(x_\circ))] \\ &= \Sigma_{c,m} \left(\mathbb{E}[Y|x, m, c] - \mathbb{E}[Y|x_\circ, m, c]\right) \mathsf{P}[m|x_\circ, c]\mathsf{P}[c]\end{aligned}$$

$$\begin{aligned}\mathsf{NIE} &= \mathbb{E}[Y(x, M(x)) - Y(x, M(x_\circ)] \\ &= \Sigma_{c,m}\mathbb{E}[Y|x, m, c] \left(\mathsf{P}[m|x, c] - \mathsf{P}[m|x_\circ, c]\right) \mathsf{P}[c]\end{aligned}$$

$$\begin{aligned}\mathsf{TE} &= \mathbb{E}[Y(x) - Y(x_\circ)] \\ &= \Sigma_c \left(\mathbb{E}[Y|x, c] - \mathbb{E}[Y|x_\circ, c]\right) \mathsf{P}[c]\end{aligned}$$

Nomenclature

SEM Name	PO Name(s)	PO Definition
Direct Effect of X	$(\beta_X + \beta_{XM} M)(x - x_\circ)$	
a) set $M = m$	Controlled direct effect	$Y(x, m) - Y(x_\circ, m)$
b) set $M = E[M \mid x_\circ, c]$	Natural direct effect †	$Y(x, M(x_\circ)) - Y(x_\circ, M(x_\circ))$
	Pure direct effect §	
	Average direct effect (control) ⋆	
c) set $M = E[M \mid x, c]$	Total direct effect §	$Y(x, M(x)) - Y(x_\circ, M(x))$
	Average direct effect (treatment) ⋆	
Indirect Effect of X		
a) set $X = x_\circ$	Pure indirect effect §	$Y(x_\circ, M(x)) - Y(x_\circ, M(x_\circ))$
	Average causal mediation effect (control) ⋆	
b) set $X = x$	Natural indirect effect †	$Y(x, M(x)) - Y(x, M(x_\circ))$
	Total indirect effect §	
	Average causal mediation effect (treat) ⋆	

†VanderWeele

§Robins and Greenland

⋆Imai et al.

Causal Mediation Models*
$\mathbf{E}[\boldsymbol{Y}\|\boldsymbol{X},\boldsymbol{M}] = \boldsymbol{\beta_0} + \boldsymbol{\beta_X X} + \boldsymbol{\beta_M M}$ $\mathrm{E}[M\|X] = \alpha_0 + \alpha_X X$
$\mathbf{E}[\boldsymbol{Y}\|\boldsymbol{X},\boldsymbol{M},\boldsymbol{C}] = \boldsymbol{\beta_0} + \boldsymbol{\beta_X X} + \boldsymbol{\beta_M M} + \boldsymbol{\beta_C C}$ $\mathrm{E}[M\|X,C] = \alpha_0 + \alpha_X X + \alpha_C C$
$\mathbf{E}[\boldsymbol{Y}\|\boldsymbol{X},\boldsymbol{M},\boldsymbol{C}] = \boldsymbol{\beta_0} + \boldsymbol{\beta_X X} + \boldsymbol{\beta_M M} + \boldsymbol{\beta_C C} + \boldsymbol{\beta_{XC} XC}$ $\mathrm{E}[M\|X,C] = \alpha_0 + \alpha_X X + \alpha_C C + \alpha_{XC} XC$
$\mathbf{E}[\boldsymbol{Y}\|\boldsymbol{X},\boldsymbol{M_1},\boldsymbol{M_2}] = \boldsymbol{\beta_0} + \boldsymbol{\beta_X X} + \boldsymbol{\beta_{M_1} M_1} + \boldsymbol{\beta_{M_2} M_2}$ $\mathrm{E}[M_1\|X] = \alpha_{01} + \alpha_1 X$ $\mathrm{E}[M_2\|X] = \alpha_{02} + \alpha_2 X$
$\mathbf{E}[\boldsymbol{Y}\|\boldsymbol{X},\boldsymbol{M_1},\boldsymbol{M_2},\boldsymbol{C}] = \boldsymbol{\beta_0} + \boldsymbol{\beta_X X} + \boldsymbol{\beta_{M_1} M_1} + \boldsymbol{\beta_{M_2} M_2} + \boldsymbol{\beta_C C}$ $\mathrm{E}[M_1\|X,C] = \alpha_{01} + \alpha_1 X + \alpha_{C1} C$ $\mathrm{E}[M_2\|X,C] = \alpha_{02} + \alpha_2 X + \alpha_{C2} C$
$\mathbf{E}[\boldsymbol{Y}\|\boldsymbol{X},\boldsymbol{M_1},\boldsymbol{M_2}] = \boldsymbol{\beta_0} + \boldsymbol{\beta_X X} + \boldsymbol{\beta_{M_1} M_1} + \boldsymbol{\beta_{M_2} M_2} + \boldsymbol{\beta_{M_1 M_2} M_1 M_2}$ $\mathrm{E}[M_1\|X] = \alpha_{01} + \alpha_1 X$ $\mathrm{E}[M_2\|X] = \alpha_{02} + \alpha_2 X$ $\mathrm{E}[M_1 M_2\|X] = \alpha_{03} + \alpha_3 X$

* Bolded equation represents the fitted model used in the proposed framework

Illustration: $h(X) = [X, X^2]$

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_M M + \beta_C C$$
$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_1^* X + \beta_2^* X^2 + \beta_C^* C$$

$$\text{TE} = \beta_1^*(x - x_\circ) + \beta_2^*(x^2 - x_\circ^2) = [\boldsymbol{h}(x) - \boldsymbol{h}(x_\circ)]\boldsymbol{\beta}^*_{\boldsymbol{X}}$$

$$\text{CDE} = \beta_1(x - x_\circ) + \beta_2(x^2 - x_\circ^2) = [\boldsymbol{h}(x) - \boldsymbol{h}(x_\circ)]\boldsymbol{\beta}_{\boldsymbol{X}}$$

$$\text{PE} = [x - x_\circ, x^2 - x_\circ^2]\begin{bmatrix}\beta_1^* - \beta_1 \\ \beta_2^* - \beta_2\end{bmatrix} = [\boldsymbol{h}(x) - \boldsymbol{h}(x_\circ)]\Delta$$

$$\text{EMCs } \Delta = -V_{\boldsymbol{X}M}V_M^{-1}\beta_M = -\begin{bmatrix}\text{cov}(\beta_1, \beta_M) \\ \text{cov}(\beta_2, \beta_M)\end{bmatrix}\text{var}(\beta_M)^{-1}\beta_M$$

What if there is moderated mediation?

Exposure-moderator interaction

$$\mathbb{E}[Y|X,M,W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$

$$\mathbb{E}[Y|X,W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

Exposure-moderator interaction

$$\mathbb{E}[Y|X,M,W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$
$$\mathbb{E}[Y|X,W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

- $h(X) = [X, XW]$

Exposure-moderator interaction

$$\mathbb{E}[Y|X, M, W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$
$$\mathbb{E}[Y|X, W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

- $h(X) = [X, XW]$
- $\Delta = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{XW}^* - \beta_{XW} \end{bmatrix}$

Exposure-moderator interaction

$$\mathbb{E}[Y|X, M, W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$
$$\mathbb{E}[Y|X, W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

- $h(X) = [X, XW]$
- $\Delta = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{XW}^* - \beta_{XW} \end{bmatrix} = -V_{XM} V_M^{-1} \beta_M$

$$= -\begin{bmatrix} \mathsf{cov}(\beta_X, \beta_M) \\ \mathsf{cov}(\beta_{XW}, \beta_M) \end{bmatrix} \mathsf{var}(\beta_M)^{-1} \beta_M$$

Exposure-moderator interaction

$$\mathbb{E}[Y|X,M,W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$

$$\mathbb{E}[Y|X,W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

- $h(X) = [X, XW]$
- $\Delta = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{XW}^* - \beta_{XW} \end{bmatrix} = -V_{XM} V_M^{-1} \beta_M$
- PE $= [h(x) - h(x_\circ)]\Delta$

Exposure-moderator interaction

$$\mathbb{E}[Y|X,M,W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$

$$\mathbb{E}[Y|X,W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

- $h(X) = [X, XW]$
- $\Delta = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{XW}^* - \beta_{XW} \end{bmatrix} = -V_{XM} V_M^{-1} \beta_M$
- $\mathsf{PE} = [h(x) - h(x_\circ)]\Delta = [1, w]\Delta$

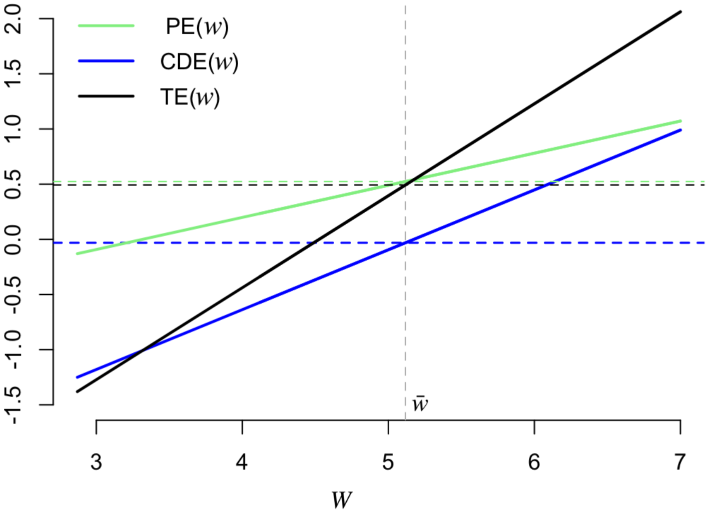
Exposure-moderator interaction

$$\mathbb{E}[Y|X,M,W] = \beta_0 + \beta_X X + \beta_M M + \beta_W W + \beta_{XW} XW \qquad \text{(F)}$$

$$\mathbb{E}[Y|X,W] = \beta_0^* + \beta_X^* X + \beta_W^* W + \beta_{XW}^* XW \qquad \text{(R)}$$

- $h(X) = [X, XW]$
- $\Delta = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{XW}^* - \beta_{XW} \end{bmatrix} = -V_{XM} V_M^{-1} \beta_M$
- PE $= [h(x) - h(x_\circ)]\Delta = [1, w]\Delta$
- The TE, CDE, and PE are *functions* of w

Exposure-moderator interaction: example



Multiple mediators

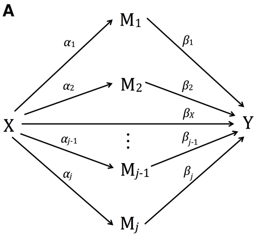
Consider the Full model that contains j mediators

$$\mathbb{E}[Y|X, M, C] = \beta_0 + h(X)\beta_X + \Sigma_{i=1}^{j}\beta_{M_i}M_i + \beta_C C$$

The total PE through $\boldsymbol{M} = [h(x) - h(x_\circ)][-V_{X\boldsymbol{M}}V_{\boldsymbol{M}}^{-1}\boldsymbol{\beta}_{\boldsymbol{M}}]$

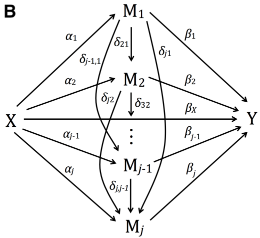
The mediator-specific PE $= [h(x) - h(x_\circ)][-V_{XM_i}V_{M_i}^{-1}\beta_{M_i}]$

Multiple mediators: parallel model



$$\mathrm{E}[Y|X, \boldsymbol{M}] = \beta_0 + \beta_X X + \beta_1 M_1 + \beta_2 M_2 + \cdots + \beta_j M_j$$
$$\mathrm{E}[M_i|X] = \alpha_{0i} + \alpha_i X, i = 1, \ldots, j$$
$$\mathrm{E}[Y|X] = \beta_0^* + \beta_X^* X$$

Multiple mediators: serial model



$$E[Y|X,\boldsymbol{M}] = \beta_0 + \beta_X X + \beta_1 M_1 + \beta_2 M_2 + \cdots + \beta_j M_j$$

$$E[M_i|X] = \alpha_{0i} + \alpha_i X + \Sigma_{k=1}^{i-1}\delta_{ik}M_k, i = 2, \ldots, j$$

$$E[M_1|X] = \alpha_{01} + \alpha_1 X$$

$$E[Y|X] = \beta_0^* + \beta_X^* X$$

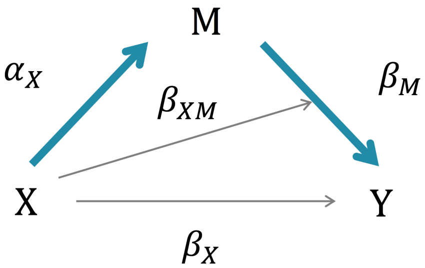
Table of total effect decompositions 😱

<table>
<tr><td colspan="4">Total Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td>Reference Interaction</td><td>Mediated Interaction</td><td>Pure Indirect Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td colspan="2">Portion Attributed to Interaction</td><td>Pure Indirect Effect</td></tr>
<tr><td colspan="2">Pure Direct Effect</td><td>Mediated Interaction</td><td>Pure Indirect Effect</td></tr>
<tr><td colspan="3">Total Direct Effect</td><td>Pure Indirect Effect</td></tr>
<tr><td colspan="2">Pure Direct Effect</td><td colspan="2">Total Indirect Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td>Reference Interaction</td><td colspan="2">Total Indirect Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td colspan="3">Portion Eliminated</td></tr>
</table>

Table of total effect decompositions 😱

<table>
<tr><td colspan="4">Total Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td>Reference Interaction</td><td>Mediated Interaction</td><td>Pure Indirect Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td colspan="2">Portion Attributed to Interaction</td><td>Pure Indirect Effect</td></tr>
<tr><td colspan="2">Pure Direct Effect</td><td>Mediated Interaction</td><td>Pure Indirect Effect</td></tr>
<tr><td colspan="3">Total Direct Effect</td><td>Pure Indirect Effect</td></tr>
<tr><td colspan="2">Pure Direct Effect</td><td colspan="2">Total Indirect Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td>Reference Interaction</td><td colspan="2">Total Indirect Effect</td></tr>
<tr><td>Controlled Direct Effect</td><td colspan="3">Portion Eliminated</td></tr>
</table>

Exposure-mediator interaction



The **pure indirect effect** attributes the interaction effect to the direct effect of the exposure.

Non-nested mediation systems

1. Specify a **Global model** under which the Full and Reduced models are nested

Non-nested mediation systems

1. Specify a **Global model** under which the Full and Reduced models are nested

2. Use formula for $\hat{\Delta}$ to estimate Δ_{RG} between the Reduced model and the Global model

Non-nested mediation systems

1. Specify a **Global model** under which the Full and Reduced models are nested
2. Use formula for $\hat{\Delta}$ to estimate Δ_{RG} between the Reduced model and the Global model
3. Use formula for $\hat{\Delta}$ to estimate Δ_{FG} between the Full model and the Global model

1. Specify a **Global model** under which the Full and Reduced models are nested
2. Use formula for $\hat{\Delta}$ to estimate Δ_{RG} between the Reduced model and the Global model
3. Use formula for $\hat{\Delta}$ to estimate Δ_{FG} between the Full model and the Global model
4. Subtract the corresponding Δs: $\Delta_{\mathsf{PE}} = \Delta_{\mathsf{RG}} - \Delta_{\mathsf{FG}}$.

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \tag{F}$$

$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \tag{R}$$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \quad \text{(G)}$$
$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \quad \text{(F)}$$
$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \quad \text{(R)}$$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \quad \text{(G)}$$

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \quad \text{(F)}$$

$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \quad \text{(R)}$$

$$\Delta_{\mathsf{RG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix}$$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \qquad \text{(G)}$$

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \qquad \text{(F)}$$

$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \qquad \text{(R)}$$

$$\Delta_{\mathsf{RG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} \text{ and } \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix}$$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \qquad \text{(G)}$$
$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \qquad \text{(F)}$$
$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \qquad \text{(R)}$$

$$\Delta_{\mathsf{RG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} \text{ and } \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix} \longleftarrow$$

Use formula for $\hat{\Delta}$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \qquad \text{(G)}$$

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \qquad \text{(F)}$$

$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \qquad \text{(R)}$$

$$\Delta_{\mathsf{RG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} \text{ and } \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix}$$

$$\Delta_{\mathsf{RG}} - \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} - \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix}$$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \quad \text{(G)}$$

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \quad \text{(F)}$$

$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \quad \text{(R)}$$

$$\Delta_{\mathsf{RG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} \text{ and } \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix}$$

$$\Delta_{\mathsf{RG}} - \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} - \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix} = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{X^2}^* - 0 \end{bmatrix} \longleftarrow \Delta_{\mathsf{PE}}$$

Non-nested mediation systems

$$\mathbb{E}[Y|X,M,C] = \gamma_0 + \gamma_X X + \gamma_M M + \gamma_{XM} XM + \gamma_C C + \gamma_{X^2} X^2 \quad \text{(G)}$$

$$\mathbb{E}[Y|X,M,C] = \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_C C \quad \text{(F)}$$

$$\mathbb{E}[Y|X,C] = \beta_0^* + \beta_X^* X + \beta_{X^2}^* X^2 + \beta_C^* C \quad \text{(R)}$$

$$\Delta_{\mathsf{RG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} \text{ and } \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix}$$

$$\Delta_{\mathsf{RG}} - \Delta_{\mathsf{FG}} = \begin{bmatrix} \beta_X^* - \gamma_X \\ \beta_{X^2}^* - \gamma_{X^2} \end{bmatrix} - \begin{bmatrix} \beta_X - \gamma_X \\ 0 - \gamma_{X^2} \end{bmatrix} = \begin{bmatrix} \beta_X^* - \beta_X \\ \beta_{X^2}^* - 0 \end{bmatrix} \longleftarrow \Delta_{\mathsf{PE}}$$

Even for non-nested mediation systems, we can estimate the PE from the fit of one (Global) outcome model.

PE on link scale

Let g=logit link.

$$
\begin{aligned}
(\beta_X^* - \beta_X)(x - x_\circ) &= \text{logit } \mathsf{E}[Y|x,c] - \text{logit } \mathsf{E}[Y|x_\circ,c] \\
&\quad - (\text{logit } \mathsf{E}[Y|x,m,c] - \text{logit } \mathsf{E}[Y|x_\circ,m,c]) \\
&= \log \frac{\text{odds } \mathsf{E}[Y|x,c]}{\text{odds } \mathsf{E}[Y|x_\circ,c]} - \log \frac{\text{odds } \mathsf{E}[Y|x,m,c]}{\text{odds } \mathsf{E}[Y|x_\circ,m,c]} \\
&= \log \frac{\text{odds } \mathsf{E}[Y|x,c]/\text{odds } \mathsf{E}[Y|x_\circ,c]}{\text{odds } \mathsf{E}[Y|x,m,c]/\text{odds } \mathsf{E}[Y|x_\circ,m,c]} \\
&= \log \frac{\text{OR}^{\text{TE}}}{\text{OR}^{\text{CDE}}}
\end{aligned}
$$

$$
\Rightarrow e^{(\beta_X^* - \beta_X)(x - x_\circ)} = \frac{\text{OR}^{\text{TE}}}{\text{OR}^{\text{CDE}}} = \text{OR}^{\text{PE}}
$$

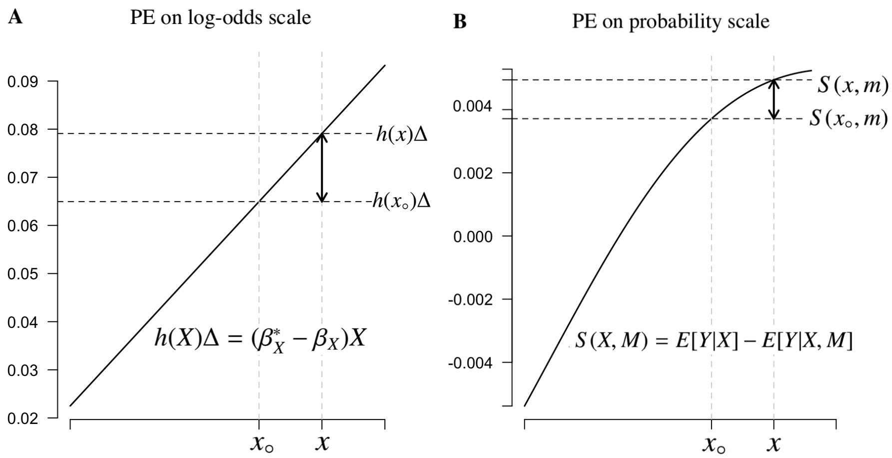
PE on expected value scale

$$g(\mathsf{E}[Y|X,M,C]) = \beta_0 + h(X)\beta_X + \beta_M M + \beta_C C$$
$$g(\mathsf{E}[Y|X,C]) = \beta_0^* + h(X)\beta_X^* + \beta_C^* C$$

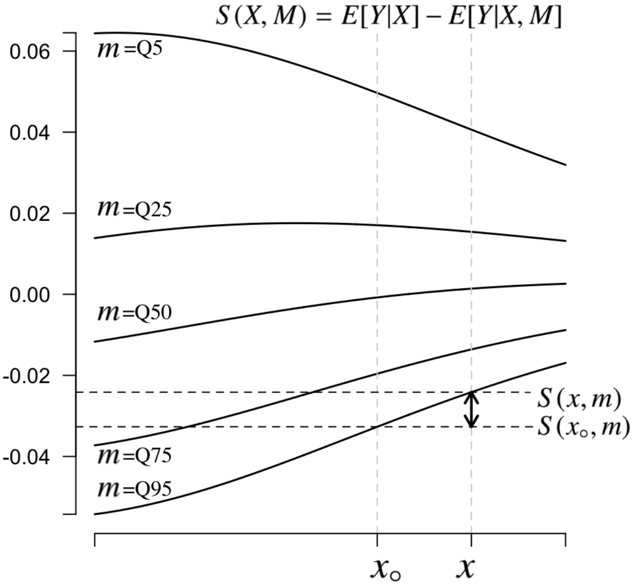
$$\mathsf{E}[Y|x,m,c] = g^{-1}(\eta_F(x,m,c))$$
$$\mathsf{E}[Y|x,c] = g^{-1}(\eta_R(x,c))$$

$$\begin{aligned}\mathsf{PE} &= \mathsf{E}[Y|x,c] - \mathsf{E}[Y|x_\circ,c] - (\mathsf{E}[Y|x,m,c] - \mathsf{E}[Y|x_\circ,m,c]) \\ &= \frac{e^{\eta_R(x,c)}}{1+e^{\eta_R(x,c)}} - \frac{e^{\eta_R(x_\circ,c)}}{1+e^{\eta_R(x_\circ,c)}} - \left(\frac{e^{\eta_F(x,m,c)}}{1+e^{\eta_F(x,m,c)}} - \frac{e^{\eta_F(x_\circ,m,c)}}{1+e^{\eta_F(x_\circ,m,c)}}\right)\end{aligned}$$

Visualizations for GLMs



PE on probability scale



Interpretation at a particular SNP

The exposure is increasing copies of the minor allele (i.e., $x - x_\circ = 1$).

The **total effect** $\text{OR}^{\text{TE}} = \frac{\text{odds E}[Y(1)|c]}{\text{odds E}[Y(1)|c]}$ is the odds ratio for lung cancer comparing those with and without an additional copy of the minor allele.

The **controlled direct effect** $\text{OR}^{\text{CDE}} = \frac{\text{odds E}[Y(1,m)|c]}{\text{odds E}[Y(0,m)|c]}$ is the odds ratio for lung cancer comparing those with and without an additional copy of the minor allele if pack-years were fixed at a certain level m.

The **portion eliminated** $\text{OR}^{\text{PE}} = \text{OR}^{\text{TE}}/\text{OR}^{\text{CDE}}$ is the portion of the total effect of increasing copy of the minor allele on lung cancer that would be eliminated if we were to fix pack-years to the same value m for all persons.

Interpretation at a particular SNP

The **natural direct effect** $\mathrm{OR}^{\mathsf{NDE}} = \frac{\text{odds } \mathsf{E}[Y(1,M(0))|c]}{\text{odds } \mathsf{E}[Y(0,M(0))|c]}$ is the odds ratio for lung cancer comparing those with and without an additional copy of the minor allele *if pack-years were what it would have been without an additional copy of the minor allele*.

The **natural indirect effect** $\mathrm{OR}^{\mathsf{NIE}} = \frac{\text{odds } \mathsf{E}[Y(1,M(1))|c]}{\text{odds } \mathsf{E}[Y(1,M(0))|c]}$ is the odds ratio for lung cancer for those with an additional copy of the minor allele *comparing what would happen if pack-years were set to what it would have been with and without the additional copy of the minor allele*.

Comparison of approaches for logistic regression

	Fitted Models	PE=TE-CDE
Diff. of coefs	$\text{logit } \mathrm{E}[Y\|X,M] = \beta_0 + \beta_X X + \beta_M M$ $\text{logit } \mathrm{E}[Y\|X] = \beta_0^* + \beta_X^* X$	$\beta_X^* - \beta_X$
KHB method	$\text{logit } \mathrm{E}[Y\|X,M] = \beta_0 + \beta_X X + \beta_M M$ $\mathrm{E}[M\|X] = \alpha_0 + \alpha_X X \to \varepsilon_{M\|X}$ $\text{logit } \mathrm{E}[Y\|X,M] = \tilde\beta_0 + \tilde\beta_X X + \tilde\beta_M \varepsilon_{M\|X}$	$\tilde\beta_X - \beta_X$ $= \alpha_X \beta_M$
Mediation formula	$\text{logit } \mathrm{E}[Y\|X,M] = \beta_0 + \beta_X X + \beta_M M$ $\mathrm{E}[M\|X] = \alpha_0 + \alpha_X X$	$\alpha_X \beta_M$
Our formula	$\text{logit } \mathrm{E}[Y\|X,M] = \beta_0 + \beta_X X + \beta_M M$ $\mathrm{E}[M\|X] = \alpha_0 + \alpha_X X, w = \hat p(1-\hat p)$	$-V_{XM} V_M^{-1} \beta_M$ $= \alpha_{X.WLS} \beta_M$

KHB and product approach

$$\tilde{\beta}_x - \beta_X = \frac{\tilde{\gamma}_x}{\tilde{\sigma}_F} - \frac{\gamma_X}{\sigma_F} = \frac{\gamma_X^*}{\sigma_F} - \frac{\gamma_X}{\sigma_F} = \frac{\gamma_X^* - \gamma_X}{\sigma_F}$$

$$\tilde{\beta}_x - \beta_X = \frac{\tilde{\gamma}_x - \gamma_X}{\sigma_F} = \frac{\alpha_X \gamma_M}{\sigma_F} = \alpha_X \beta_M$$

Extension to GLMs

$$g(\mathsf{E}[Y|X, M, C]) = \beta_0 + h(X)\beta_X + \beta_M M + \beta_C C$$
$$g(\mathsf{E}[Y|X, C]) = \beta_0^* + h(X)\beta_X^* + \beta_C^* C$$

- Interpretation of Δ changes
- $-\hat{V}_{XM}\hat{V}_M^{-1}\hat{\beta}_M \approx \hat{\beta}_X^* - \hat{\beta}_X$
- $\mathsf{E}[M|X, C] = \alpha_0 + h(X)\alpha_X + \alpha_C C$, $w = \hat{p}(1 - \hat{p})$

Extension to GLMs

$$g(\mathsf{E}[Y|X, M, C]) = \beta_0 + h(X)\beta_X + \beta_M M + \beta_C C$$

Things to consider:

- Mediation effects defined on link scale or expected value scale
- What is the true marginal model? Marginalizing over logistic doesn't necessarily yield another logistic
- Collapsibility

Working solution:

- Hold the error variance constant by splitting M into the amount that depends on X and the residuals $\varepsilon_{M|X}$ (generalization of the "KHB approach")
- Use Global model idea