Statistical mediation analysis

Martijn Heymans & Judith Rijnhart

mw.heymans@vumc.nl

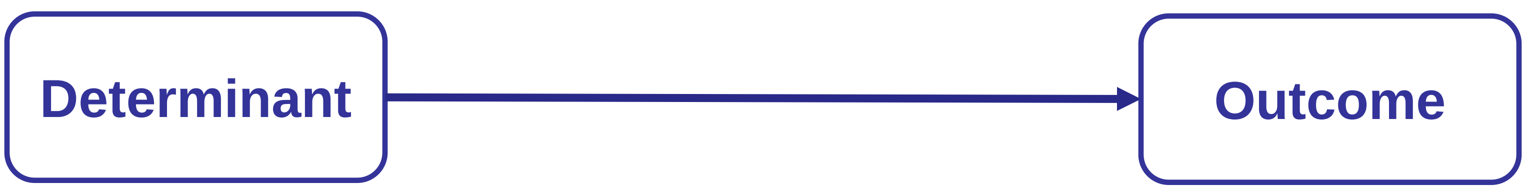
j.rijnhart@vumc.nl

VU medisch centrum

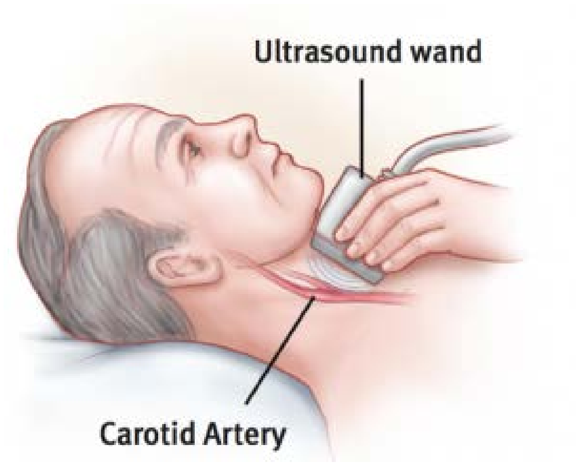
Outline

- Introduction
- Background
- Examples to determine mediation
- Structural Equation Modeling
- Potential outcomes framework
- Future topics

Epidemiology

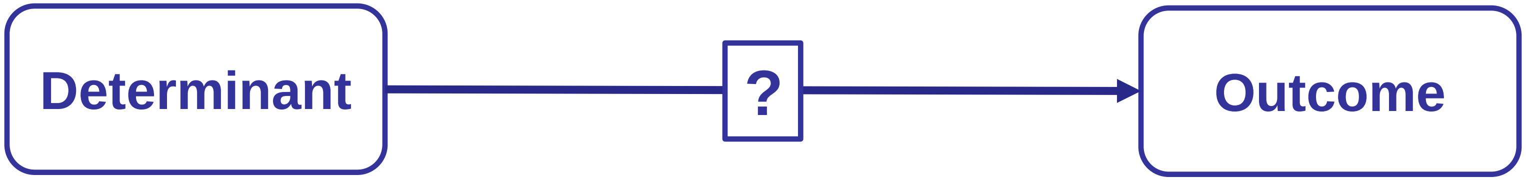


BMI —Decrease→ **Carotid distensibility**

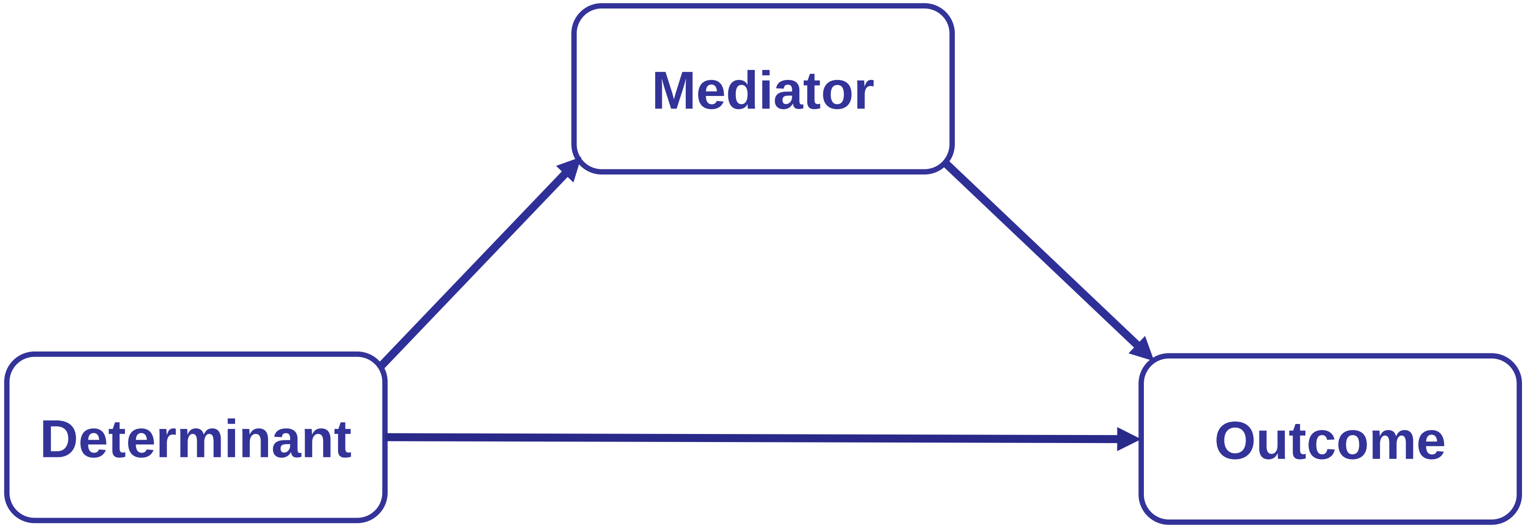


Ferreira et al. Hypertension 2012 (AGGO study)

Questions of why and how?



Mediation analysis

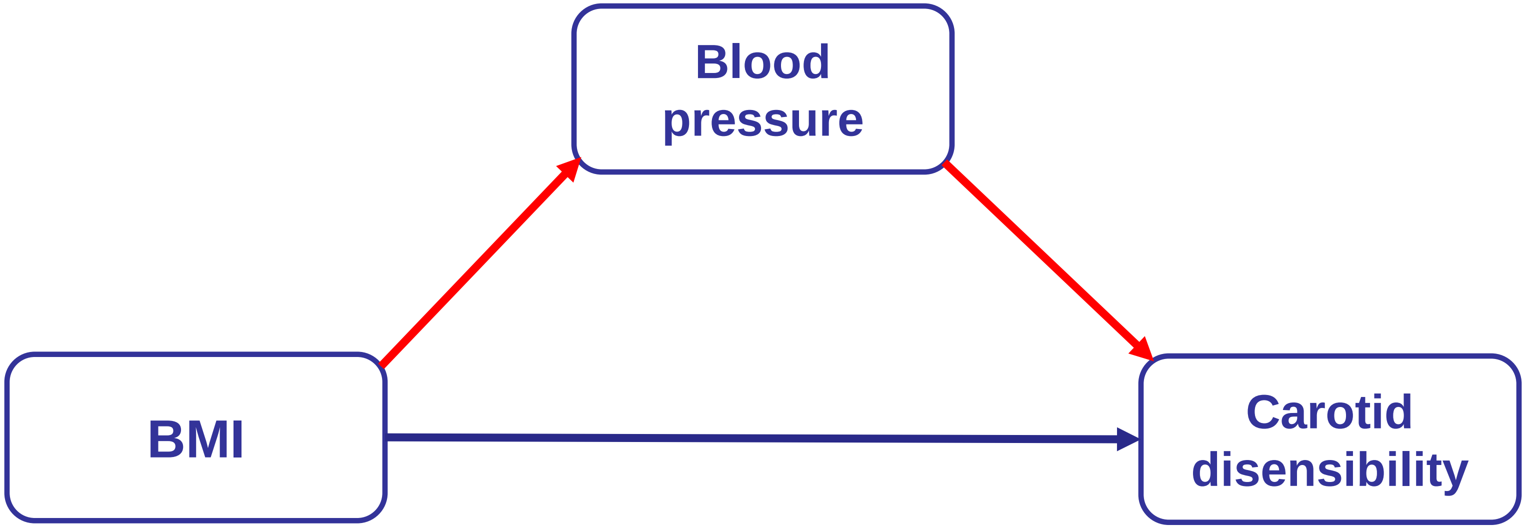


What is a mediator?

Mediator

- Affected by the determinant
- Affects the outcome
- In the causal pathway

Example of a mediator



Hasan M. et al. Circulation 2012.
Benetos A et al. American Journal of Hypertension, 2002.
Ferreira et al. J Hypertens, 2004.

Mediator vs. confounder

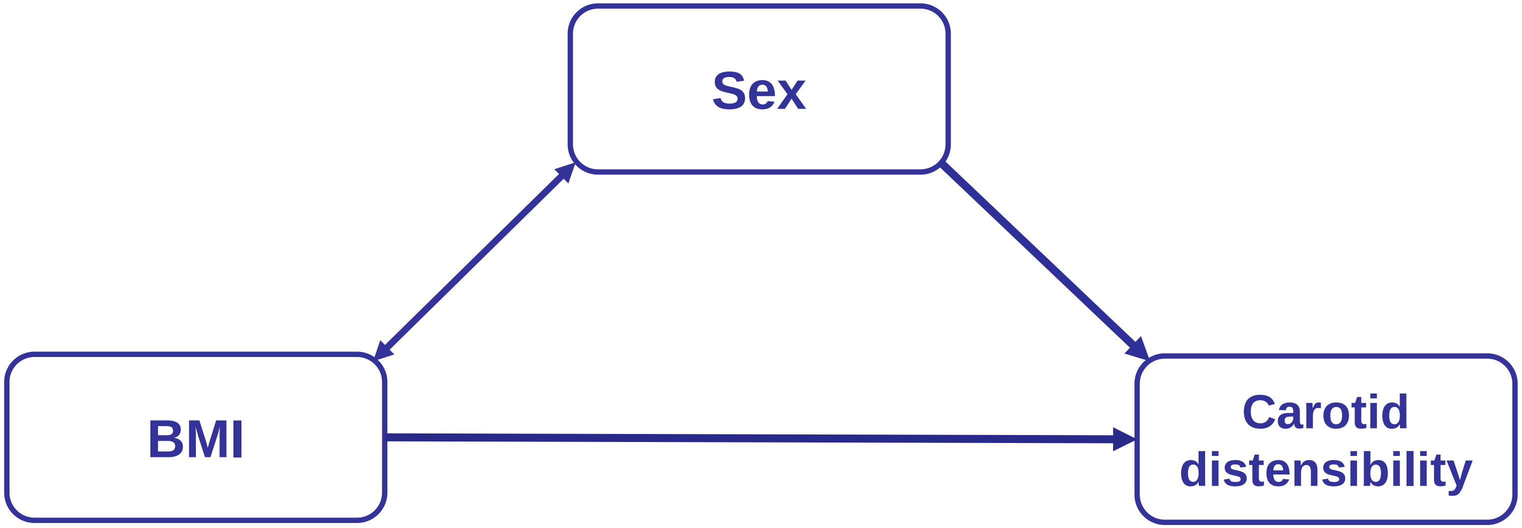
Mediator

- Affected by the determinant
- Affects the outcome
- In the causal pathway

Confounder

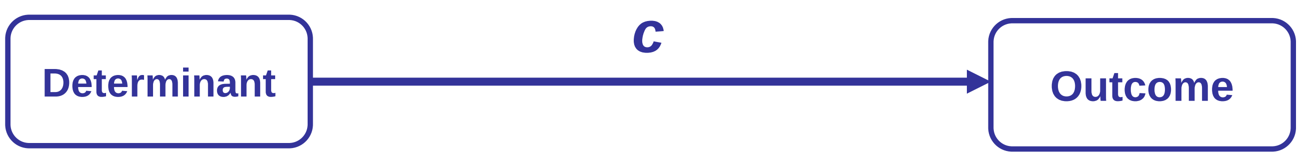
- Association with the determinant
- Affects the outcome
- **Not** in the causal pathway

Example of a confounder



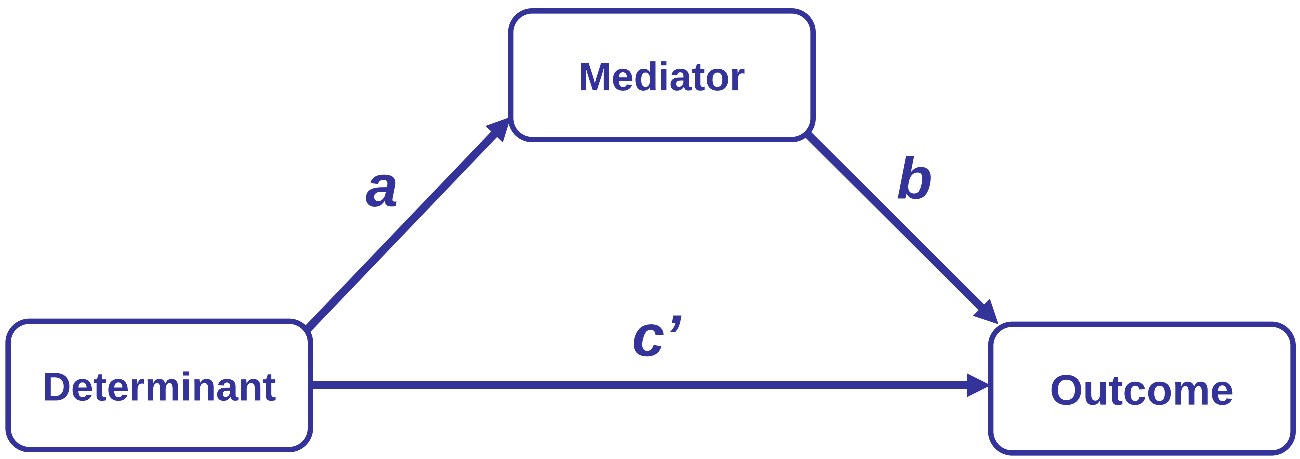
Terminology

Total effect:



Indirect effect:

Direct effect:



Outline

- Introduction
- Background
- Examples to determine mediation
- Structural Equation Modeling
- Potential outcomes framework
- Future topics

1986: Causal steps method

Journal of Personality and Social Psychology
1986, Vol. 51, No. 6, 1173–1182

0022-3514/86/$00.75

The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations

Reuben M. Baron and David A. Kenny
University of Connecticut

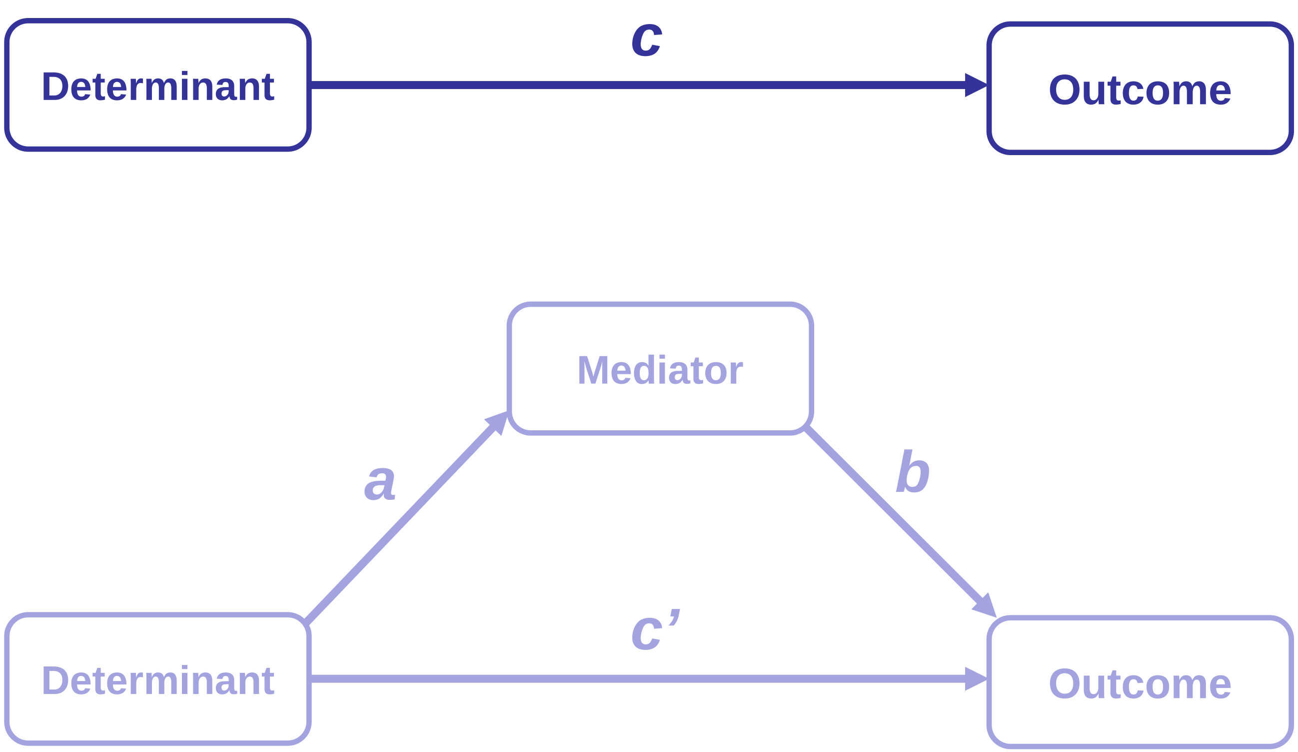
Causal steps method

1. Assess significance of the total effect (*c* path)

2. Assess significance of the *a* path

3. Assess significance of the *b* path

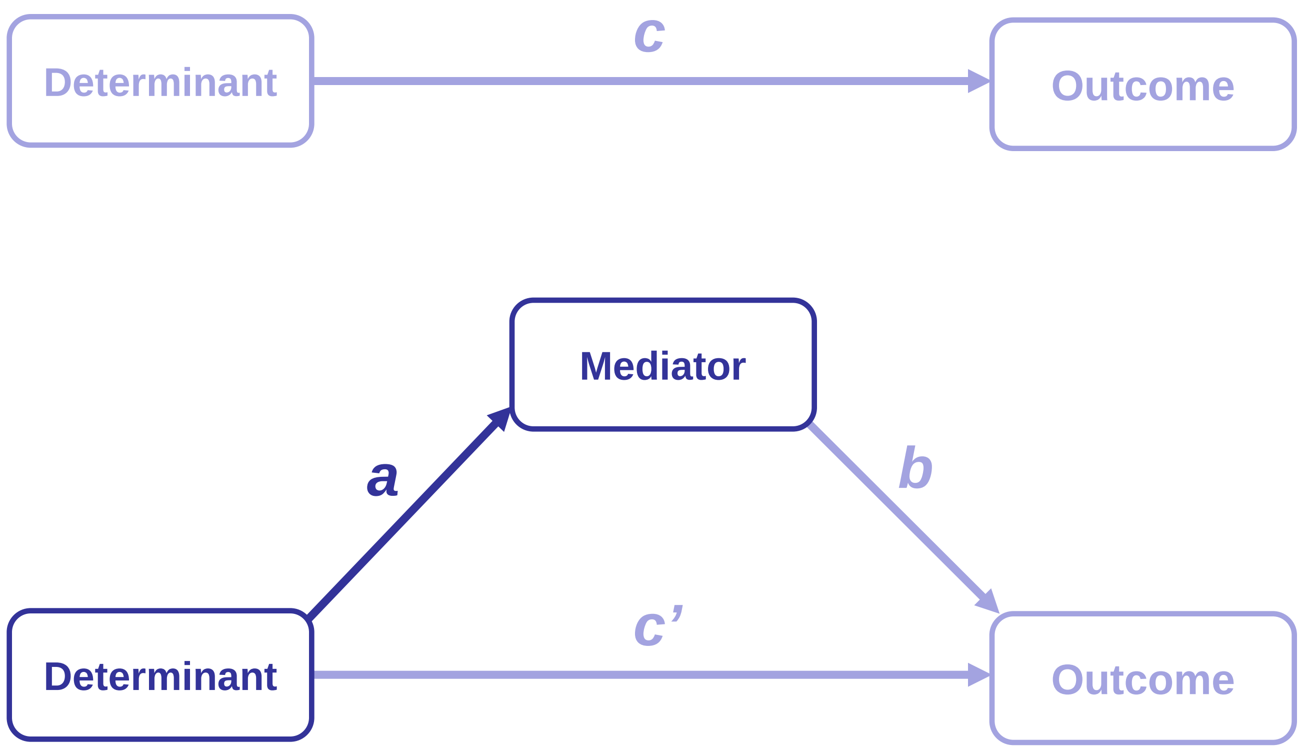
Step 1: significance of the total effect (*c*)



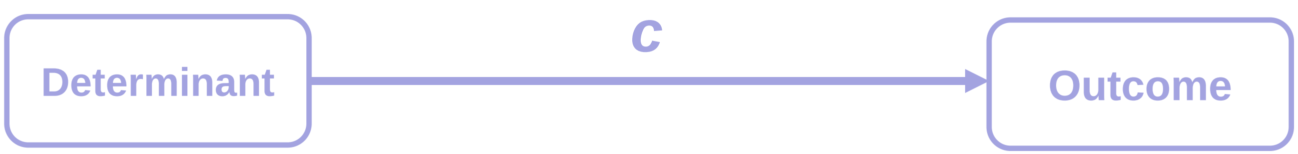
$$Y = i_1 + cX + \varepsilon_1$$

Step 2: significance of the *a* path

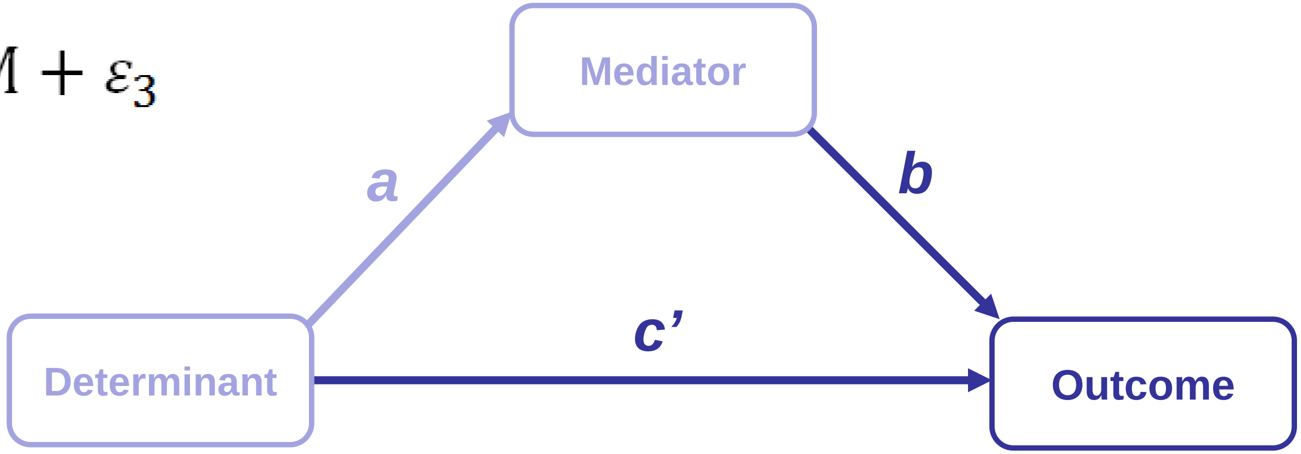
$$\mathrm{M} = i_2 + a\mathrm{X} + \varepsilon_2$$



Step 3: significance of the *b* path



$$Y = i_3 + c'X + bM + \varepsilon_3$$



Outline

- Introduction
- Background
- Examples to determine mediation
- Structural Equation Modeling
- Potential outcomes framework
- Future topics

Example

Does blood pressure mediate the relationship between BMI and carotid distensibility?

Example

Total effect of BMI on carotid distensibility (*c* path)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	36,492	2,383		15,315	,000
	Body mass index in 2000 (kg/m-2)	-,409	,098	-,212	-4,174	,000

a. Dependent Variable: carotid artery distensibility coefficient (10-3/kPa)-(2*diameter*distension+distension^2)/local pulse pressure*diameter^2

Example

Effect of BMI on blood pressure (*a* path)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	57,970	3,202		18,105	,000
	Body mass index in 2000 (kg/m-2)	1,001	,132	,366	7,594	,000

a. Dependent Variable: averaged (brachial) mean blood pressure (mmHg)

Example

Effect of blood pressure on carotid distensibility (*b* path)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	52,637	3,035		17,346	,000
	averaged (brachial) mean blood pressure (mmHg)	-,279	,036	-,393	-7,774	,000
	Body mass index in 2000 (kg/m-2)	-,131	,098	-,068	-1,335	,183

a. Dependent Variable: carotid artery distensibility coefficient (10-3/kPa)-(2*diameter*distension+distension^2)/local pulse pressure*diameter^2

Conclusion

Blood pressure mediates the relationship between BMI and carotid distensibility!

Critique on the causal steps method

- Relies heavily on significance
- No estimate of the mediated effect
- Does not account for inconsistent mediation

Nowadays

Communication Monographs
Vol. 76, No. 4, December 2009, pp. 408–420

Routledge
Taylor & Francis Group

Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium

Andrew F. Hayes

CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE

Current Directions in Mediation Analysis

David P. MacKinnon[1] and Amanda J. Fairchild[2]

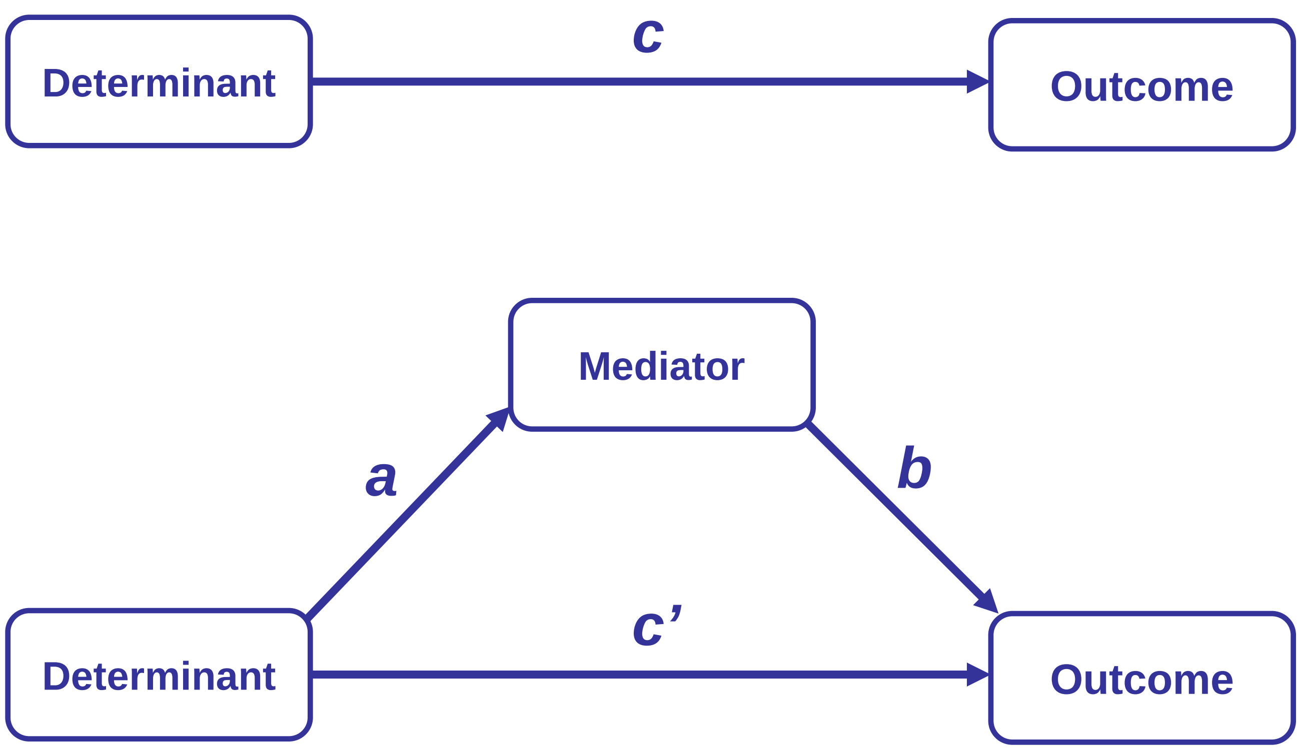
[1]Arizona State University and [2]University of South Carolina

Current practice

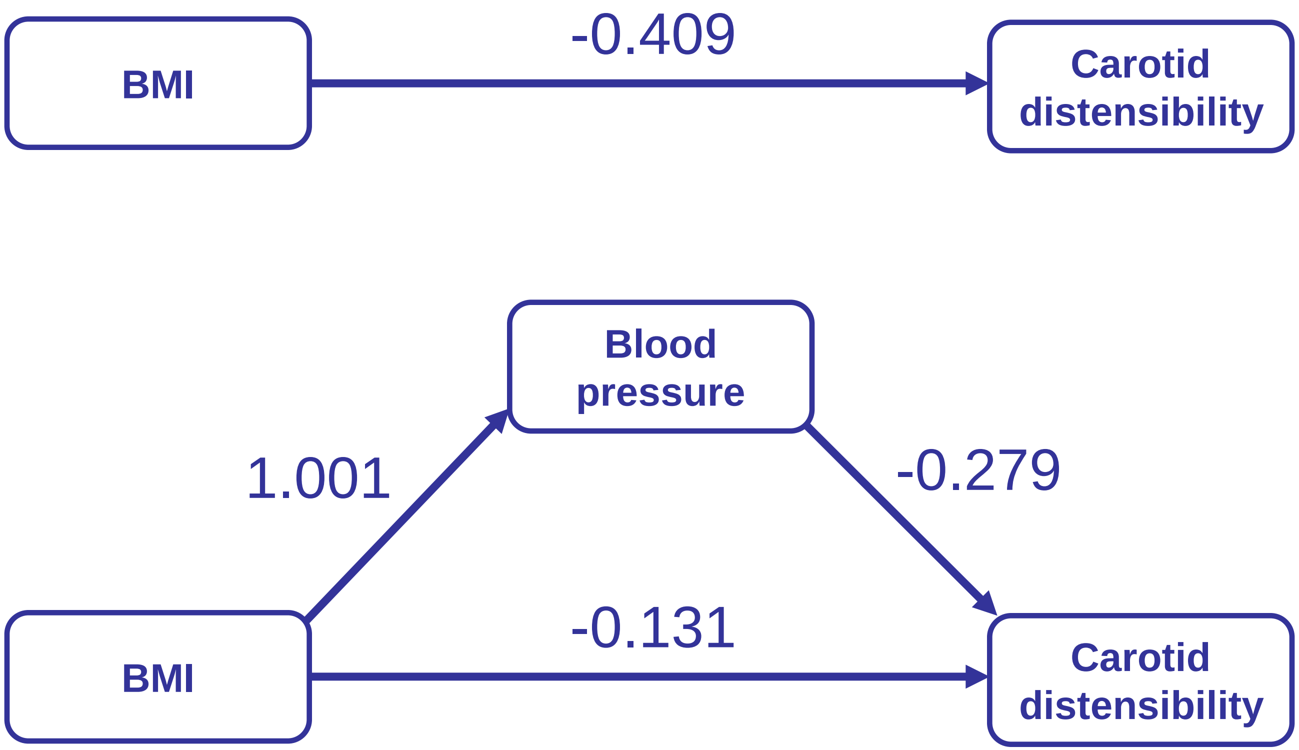
- Calculation of the indirect effect

Indirect effect

- $c - c'$
- $a * b$



Example



$\boldsymbol{c - c'}$ = -0.409 – -0.131 = **-0.278**

$\boldsymbol{a * b}$ = 1.001 * -0.279 = **-0.279**

Example

Summary

Total effect	-0.409
***a* path**	1.001
***b* path**	-0.279
Direct effect	-0.131
Indirect effect	-0.279

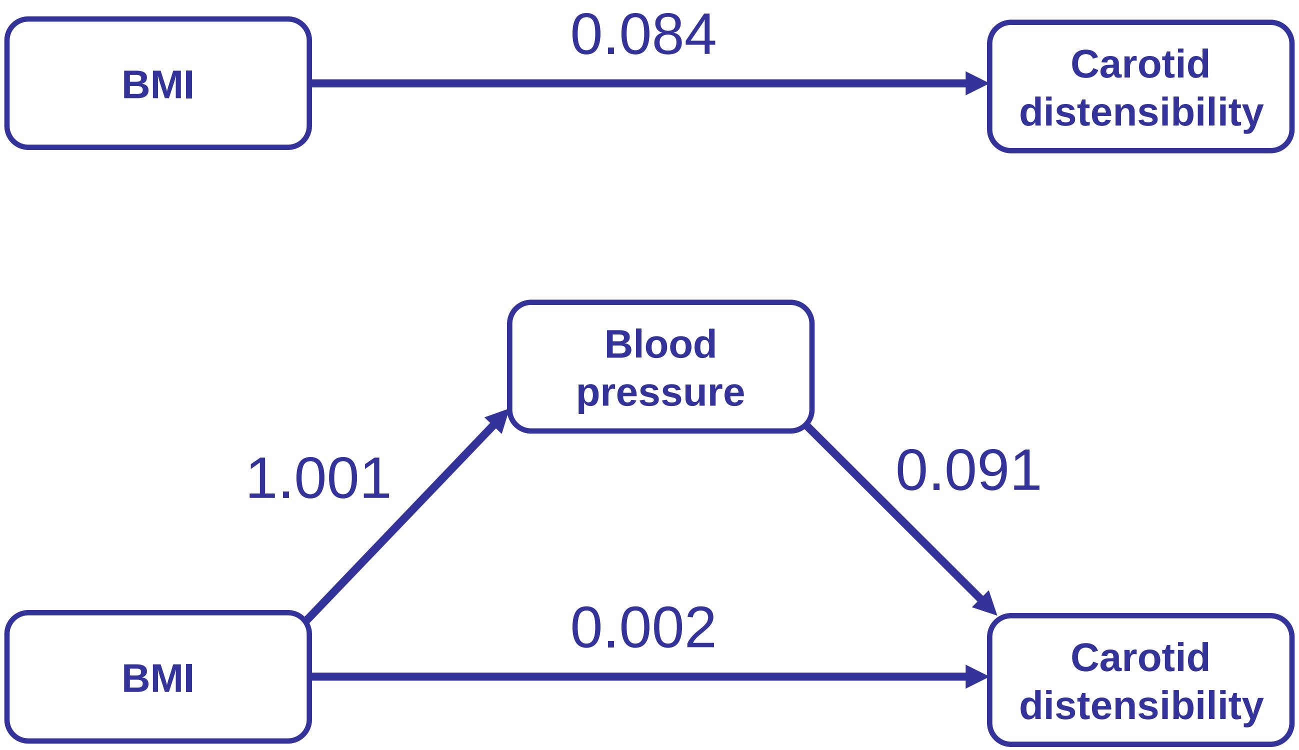
Example – Dichotomous outcome

Does blood pressure mediate the relationship between BMI and carotid distensibility?

Dichotomous carotid distensibility variable, with lowest quartile as cut-off

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Normal CD	281	75,1	75,1	75,1
	Low CD	93	24,9	24,9	100,0
	Total	374	100,0	100,0	

Example – Dichotomous outcome



$$\boldsymbol{c - c'} = 0.084 - 0.002 = \mathbf{0.082}$$

$$\boldsymbol{a * b} = 1.001 * 0.091 = \mathbf{0.091}$$

Example – Dichotomous outcome

Indirect effect	
$c - c'$	0.082
$a * b$	0.091

Which estimate to believe?

Variance in regression models

- The scale of the coefficients is influenced by the total variance in the model

- Change in the total variance → change in the coefficients

Linear regression models

Total variance =

explained variance + unexplained variance

- When the explained variance ↑, unexplained variance ↓
- Total variance always remains the same
- Scale of the coefficients will not be influenced

Logistic regression models

Total variance =

explained variance + 3.14

- When the explained variance ↑, unexplained variance remains 3.14
- Total variance changes as the explained variance changes
- Scale of the coefficients will change

Logistic regression models

$$\mathrm{Y} = i_1 + c\mathrm{X}$$

vs.

$$\mathrm{Y} = i_3 + c'\mathrm{X} + b\mathrm{M}$$

↑ in explained variance

↑ in total variance

↑ of scale of *c'*

***c* – *c'* will be an underestimation**

Example – dichotomous outcome

Indirect effect	
$c - c'$	0.082
$\boldsymbol{a * b}$	**0.091**

Example – dichotomous outcome

Summary

	Ln(OR)	OR
Total effect *ab*+*c'*	0.093	1.10
Direct effect *c'*	0.002	1.00
Indirect effect *ab*	0.091	1.10

Outline

- Introduction
- Background
- Examples to determine mediation
- Structural Equation Modeling
- Potential outcomes framework
- Future topics

Structural Equation Modeling (SEM)

- **Regression like procedure**
- **Multiple regression equations can be calculated simultaneously**
- **The same variable can be both dependent and independent in the same model (in contrast to regression analyses)**
- **Both observed variables and unobserved 'latent' variables (construct as stress or the slope variable in longitudinal models) can be included in the model**
- **SEM Model can be visualized**

Mediation analysis with MPlus

RCT, n=546 schoolchildren
Independent = Group variable
Mediator = Change in sweetened beverages consumption
Outcome = Change in Body Mass Index

```
TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS


                                                         Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Effects from GROUP to ZRE_BMI

  Total               -0.170      0.086     -1.971      0.049
  Total indirect      -0.025      0.020     -1.244      0.214

  Specific indirect

    ZRE_BMI
    ZRE_SSB
    GROUP             -0.025      0.020     -1.244      0.214

  Direct
    ZRE_BMI
    GROUP             -0.145      0.089     -1.639      0.101
```

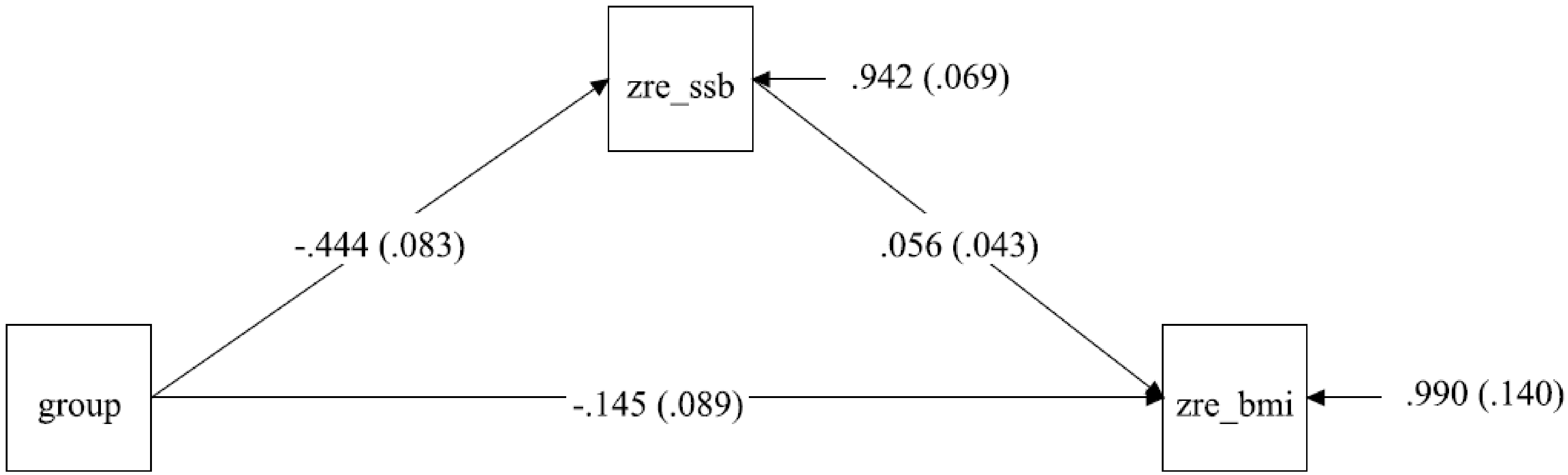
Want to use R: Lavaan package

Mediation analysis with MPlus

```
CONFIDENCE INTERVALS OF TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS


                  Lower .5%  Lower 2.5%    Lower 5%     Estimate     Upper 5%  Upper 2.5%    Upper .5%

Effects from GROUP to ZRE_BMI

  Total             -0.392      -0.339      -0.312      -0.170      -0.028      -0.001       0.052
  Total indirect    -0.076      -0.064      -0.057      -0.025       0.008       0.014       0.026

  Specific indirect

    ZRE_BMI
    ZRE_SSB
    GROUP           -0.076      -0.064      -0.057      -0.025       0.008       0.014       0.026

  Direct
    ZRE_BMI
    GROUP           -0.373      -0.319      -0.291      -0.145       0.001       0.028       0.083
```

Want to use R: Lavaan package

Mplus - Diagram



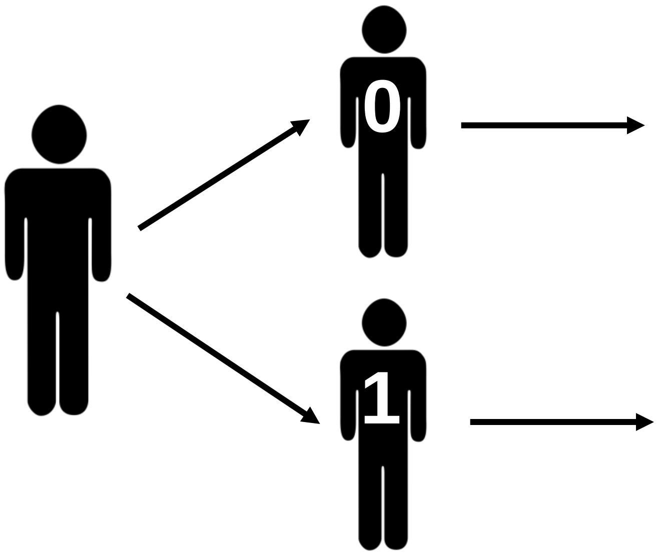
Outline

- Introduction
- Background
- Examples to determine mediation
- Structural Equation Modeling
- Potential outcomes framework
- Future topics

Potential outcomes framework (POF)

- **Estimate Average Causal (Indirect) Effect**
- **Ideal: Potential outcome every person under both determinant (treatment) groups**
- **Observe: Potential outcome of only one determinant (treatment) group**
- **Assumption under POF: Observed = ideal outcome**

Estimating Causal effect



Y(0) **= Estimated outcome under control**

Y(1) **= Observed outcome under treatment**

Average Causal effect = Y(1) – Y(0)

Causal Mediation

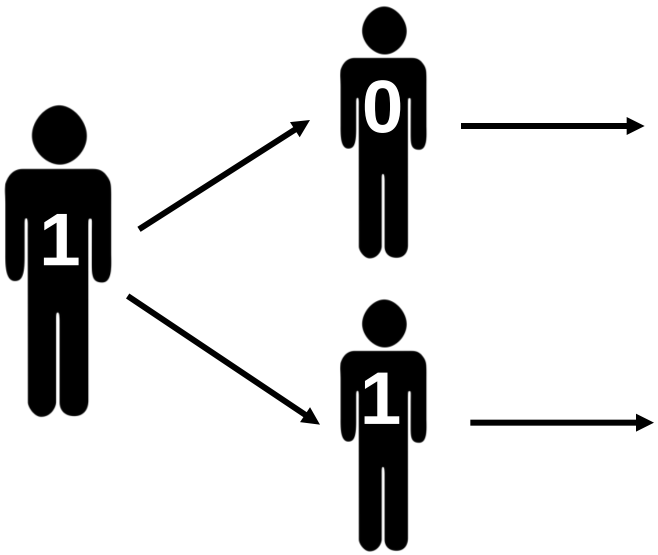
Mediation, defined under the POF framework

Potential outcome depend on Determinant group (e.g. Treatment) and Mediator variable

Causal effects defined in terms of Indirect, Direct and Total effects

R package mediation (Imai et al. 2010)

Causal Mediation effect



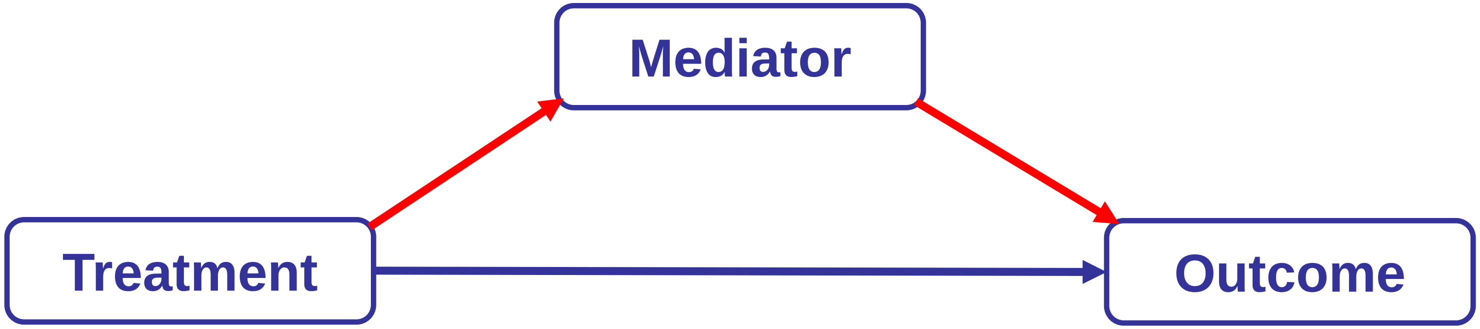
Y(1, M(0)) = **Estimated outcome via mediator, under control**

Y(1, M(1)) = **Observed outcome via mediator, under treatment**

Average Causal Mediation effect = Y(t, M(1)) – Y(t, M(0))

for t=0 and 1

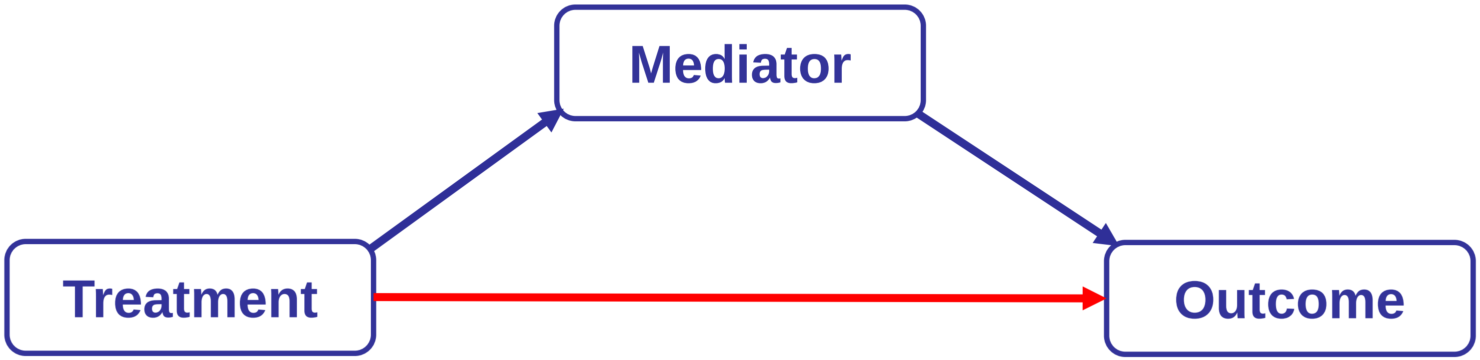
Indirect effect



Indirect effect:

- Part of the effect of the treatment on the outcome via the mediator.
- Difference in the outcome if the treatment categories are kept constant (i.e. direct effect suppressed) and the mediator changes.
- POF: Indirect effect = Y(T=1)M(1) – Y(T=1)M(0)

Direct effect



Direct effect:

- Part of the effect of the treatment on the outcome not via the mediator (i.e. all remaining causal mechanisms).

- Difference in the outcome if treatment changes and the mediator is kept constant.

- POF: Direct effect = Y(1)M(1) – Y(0)M(1)

POF approach

- Indirect effects separately estimated in Treatment (determinant) groups and averaged

- Assumptions (sequential ignorability):
  1. Treatment randomly assigned or wel controlled (adjusted for confounding)
  2. No confounding in Mediator-outcome relationship.

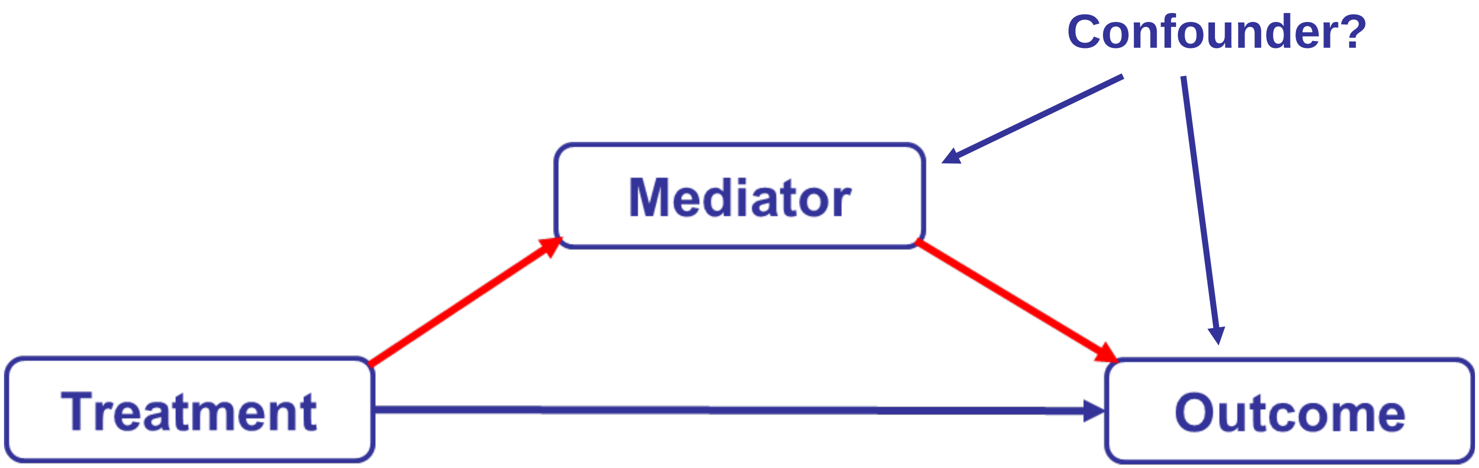
- Advantage POF method:
  - Generalizes easy to models with different types of outcome and mediators as continuous, dichotomous, categorical, non-linear relationships (splines)

Example of Output

```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile
Method

               Estimate 95% CI Lower 95% CI Upper p-value
ACME            -0.0247      -0.0623         0.01   0.180
ADE             -0.1453      -0.3283         0.03   0.096 .
Total Effect    -0.1700      -0.3480        -0.01   0.048 *
Prop. Mediated   0.1452      -0.1914         0.94   0.212
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 546


Simulations: 500

ACME = Average causal mediation effect
ADE = Average direct effect
```

Assumptions (sequential ignorability): No unmeasured confounding

Without adjustment: Also the effect of confounder is estimated in e.g. path b



There is overlap (correlation) between the unexplained variance of model

$\mathrm{M} = i_2 + a\mathrm{X} + \varepsilon_2$ and $\mathrm{Y} = i_3 + c'\mathrm{X} + b\mathrm{M} + \varepsilon_3$

Sensitivity analysis, how?

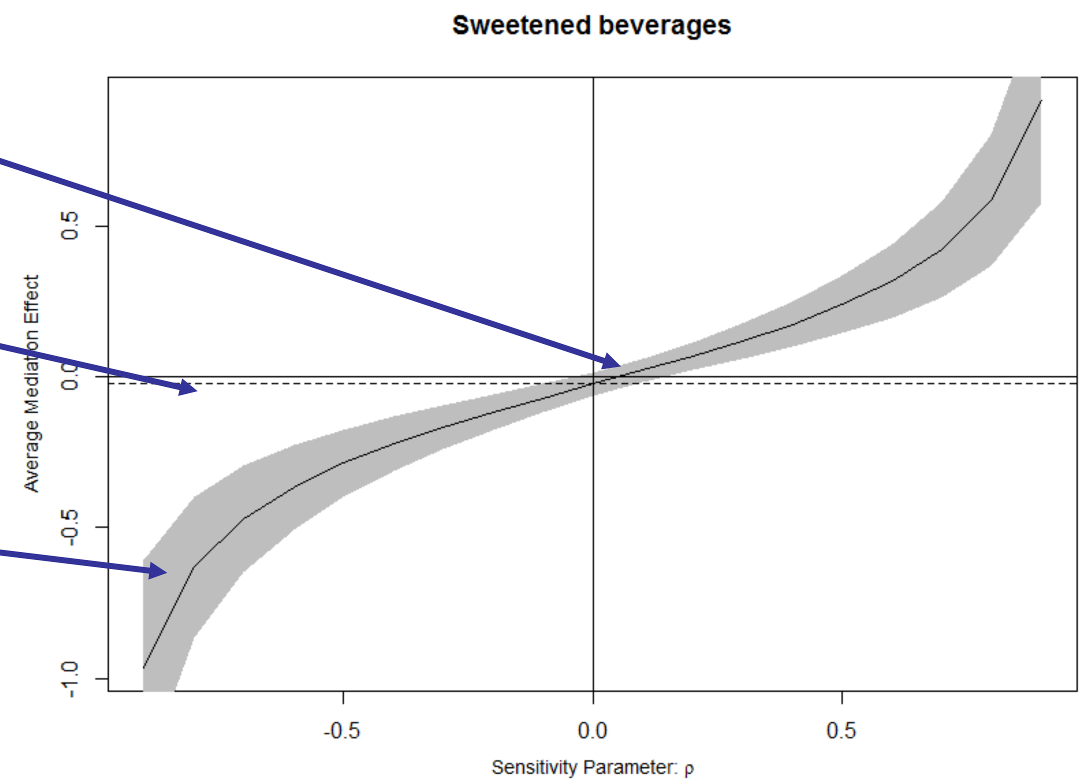
Plot the ACME at different values for the correlation

```
Mediation Sensitivity Analysis for Average Causal
Mediation Effect

Sensitivity Region

     Rho    ACME 95% CI Lower 95% CI Upper
[1,] 0.0 -0.0247      -0.0640       0.0146
[2,] 0.1  0.0211      -0.0179       0.0601
```

ACME < 0 if correlation < 0.1

ACME assumption is met

ACME when correlation varies



Contemporary Clinical Trials Communications 7 (2017) 130–135

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Comparison of methods for the analysis of relatively simple mediation models

Judith J.M. Rijnhart[a,*], Jos W.R. Twisk[a], Mai J.M. Chinapaw[b], Michiel R. de Boer[c], Martijn W. Heymans[a]

[a] Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands
[b] Department of Public and Occupational Health, Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands
[c] Department of Methodology and Applied Biostatistics, Faculty of Earth and Life Sciences, Institute for Health Sciences, Amsterdam Public Health Research Institute, VU University, Amsterdam, The Netherlands

Summary of SEM vs POF

- **Continuous outcome and mediator, SEM = POF**

- **Dichotomous mediator and outcome variable:**

Preliminary results of a simulation study

- **SEM: product of coefficient method valid**

- **Potential outcomes works best at low incidence rates**

Rijnhart, Twisk, Eekhout, Heymans, 2017 (Submitted)

Outline

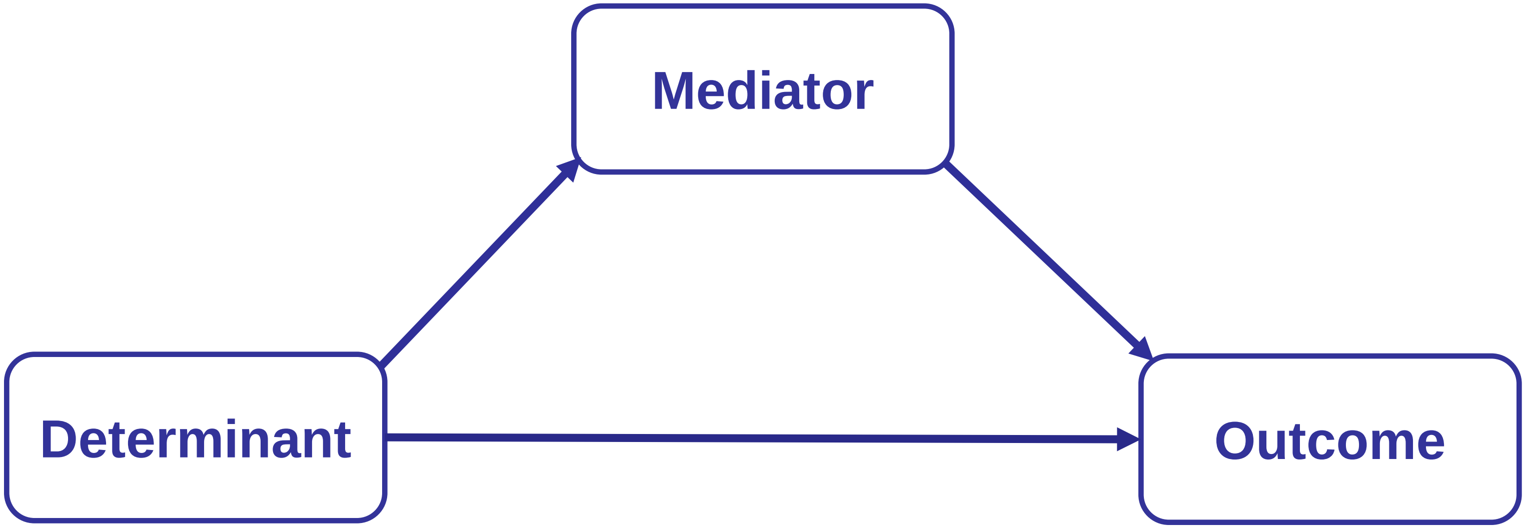
- Introduction
- Background
- Examples to determine mediation
- Structural Equation Modeling
- Potential outcomes framework
- Future topics

Future topics

- Longitudinal mediation analysis

- Bayesian mediation analysis

Longitudinal mediation analysis



Assumption of temporal precedence

- Determinant precedes mediator
- Mediator precedes outcome

Longitudinal mediation analysis

- MacArthur approach
- Cross-lagged panel model
- Latent growth curve model
- Latent difference score model
- Multilevel model
- Multilevel structural equation model

Longitudinal mediation analysis

- **Current project:** comparing methods using real-life data

- **Future project:** enhancing causality in longitudinal mediation models

(Collaboration with David MacKinnon (Arizona State University))

Bayesian mediation analysis

- Combining data with prior information
- No distributional assumptions (advantageous in small samples)

Bayesian mediation analysis

- **Current project:** comparing frequentist and Bayesian MSEM models

  (Collaboration with Emmanuel Lesaffre (KU Leuven))

EpidM course on mediation analysis

Three-day course on mediation analysis (in Dutch)

- 7, 8 & 9 March 2018
- Lectures & computer practicals
- More information on: www.epidm.nl