THE DIRECT EFFECT OF CHILDHOOD SOCIOECONOMIC CIRCUMSTANCES ON ADULT HEALTH: *MODERN MEDIATION METHODS*

Arijit Nandi, arijit.nandi@mcgill.ca

McGill
Department of Epidemiology, Biostatistics and Occupational Health
Département d'épidémiologie, biostatistique et santé au travail

OUTLINE

I. MEDIATION ANALYSIS IN SOCIAL EPIDEMIOLOGY

II. TRADITIONAL METHODS FOR MEDIATION ANALYSIS

III. MODERN MEDIATION METHODS

IV. AN APPLICATION: SES AND HEALTH OVER THE LIFE-COURSE

V. FUTURE DIRECTIONS AND CONCLUDING REMARKS

(1) MEDIATION ANALYSIS IN SOCIAL EPIDEMIOLOGY

MEDIATION ANALYSIS IN SOCIAL EPIDEMIOLOGY

- The goal of mediation analysis is to assess the mechanisms that explain an observed relation between an exposure variable and an outcome and how they relate to a third variable, the mediator

- Mediation analysis is fundamental to understanding how social factors contribute to population health

Figure 1. *The "direct effect" of a social factor on health*

- However, the relevance of mediation to the study of social determinants has not always been appreciated

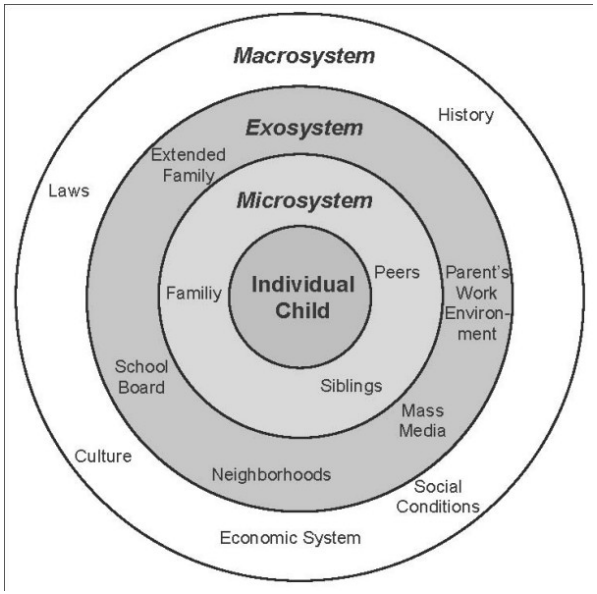
A PARADIGM SHIFT

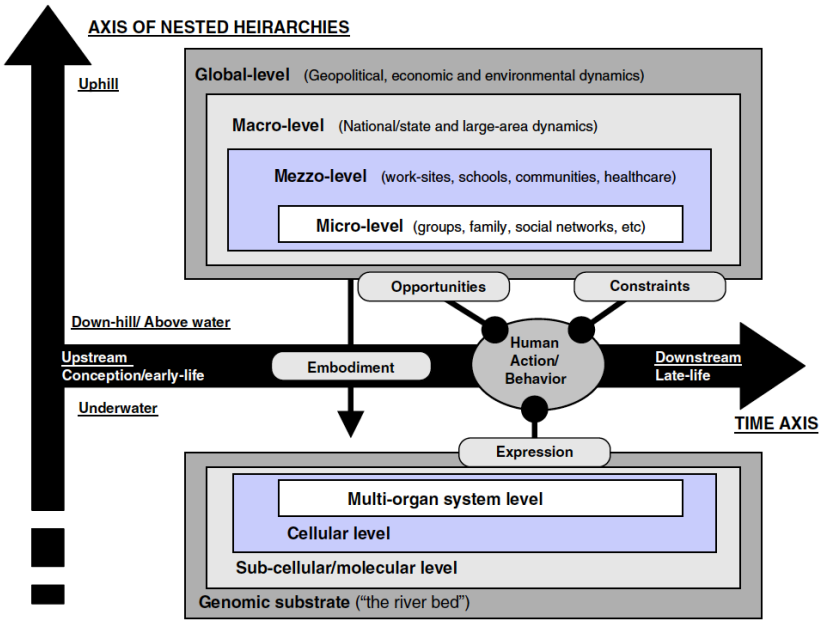
- In the mid to late 1990s, there was substantial debate in epidemiology over what unit of analysis (i.e., micro/genetic, individual/behavioral, or macro/social) is most relevant health

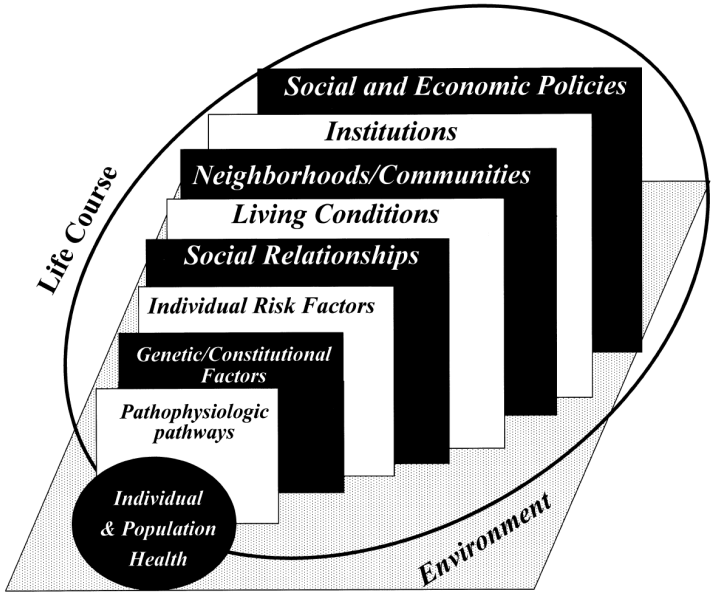
- From this debate emerged the recognition that *"We need to be equally concerned with causal pathways at the societal level and with pathogenesis and causality at the molecular level."* (Susser[2])

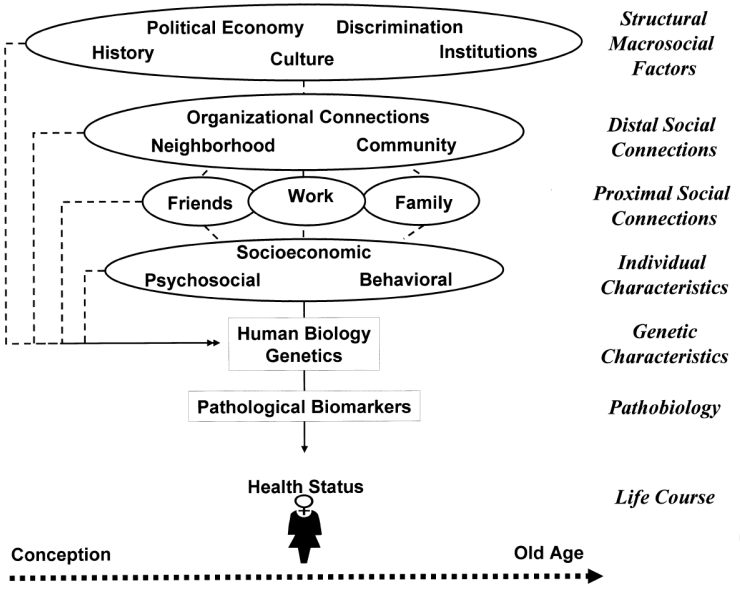
- A multilevel paradigm developed, and with it came a profusion of hierarchical conceptual frameworks—investigating mediation of social by biological factors, as well as earlier by later exposures, has become a prolific research area in social epidemiology

SOME MULTILEVEL FRAMEWORKS









RELEVANCE OF MEDIATION ANALYSIS TO SDH

1. Identifying critical periods over the life course; e.g., does adult SES mediate the effects of childhood SES on health?

2. Assessing factors that explain social/racial/gender inequalities; e.g., do health behaviors mediate social inequalities in health?

3. What are the mechanisms that explain the effects of social interventions; e.g., what explains the effects of CCT programs?

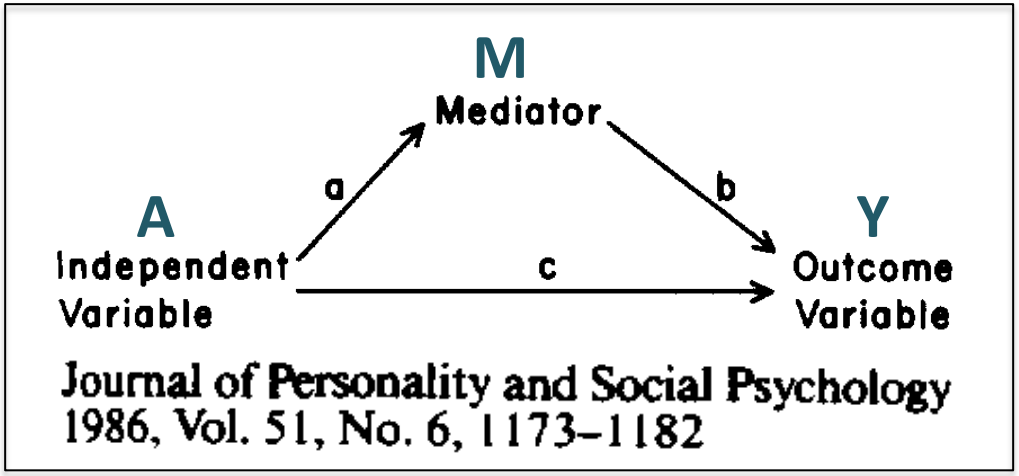
(2) TRADITIONAL METHODS FOR MEDIATION ANALYSIS

BARON AND KENNY'S "PRODUCT METHOD"

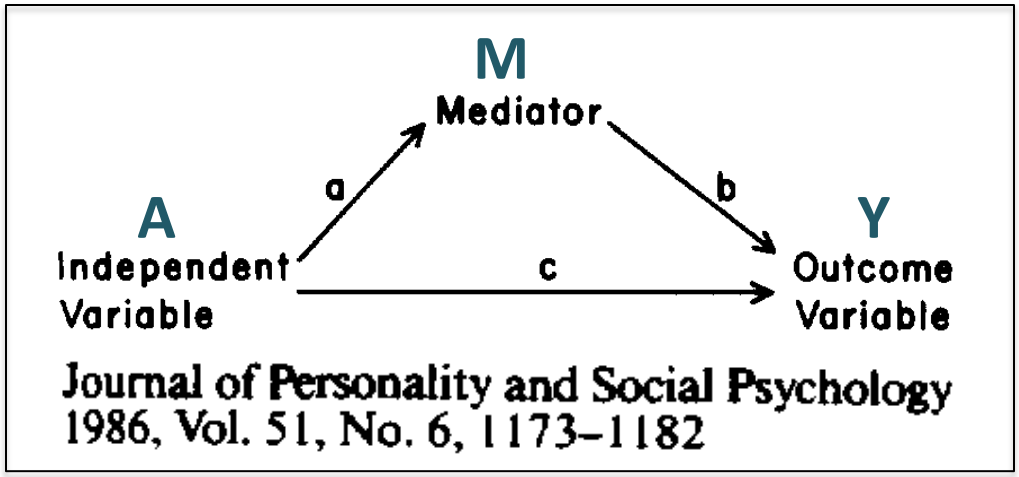
- In parallel to the development of theoretical frameworks has been the evolution of empirical methods for mediation analysis

- One method for mediation analysis, known as the *product method*, was popularized by the work of Baron and Kenny (B&K)

- B&K provided a simple approach for decomposing the overall effect of an exposure on an outcome into direct and indirect parts



ESTIMATING DIRECT AND INDIRECT EFFECTS VIA B&K



With a continuous mediator and outcome, the direct and indirect effects can be estimated by fitting two regression models:

[1] $E[M|a] = \beta_0 + \beta_1 a$

[2] $E[Y|a,m] = \theta_0 + \theta_1 a + \theta_2 m$

The direct effect is given by θ_1 and the indirect effect by $\beta_1\theta_2$

Poll: How common is this approach?

FiveThirtyEight

Nate Silver's Political Calculus

IS IT TIME TO ADMIT THAT WE HAVE A PROBLEM?

- The B&K method has been cited ~50,000 times (Google scholar)
- However, it gives unbiased results under very specific circumstances
- Specifically, the product method cannot be used to estimate direct and indirect effects in models with interactions or nonlinearities
- Further, the assumptions needed for the direct and indirect effects to be interpreted as causal (i.e., unconfounded) effects were unclear

(3) MODERN MEDIATION METHODS

RECENT DEVELOPMENTS IN MEDIATION ANALYSIS

- Over the past decade, there has been a "*veritable explosion*" of methodological developments for causal mediation analysis

- The counterfactual framework has underlined the assumptions needed to identify direct and indirect effects, including the *controlled direct effect*, and *natural direct and indirect effects*

- Newer approaches also accommodate settings with an interaction between the exposure and mediator in the outcome model

THE TOTAL EFFECT

- When performing a mediation analysis, we often start by estimating the *total effect* of an exposure *A* on outcome *Y*

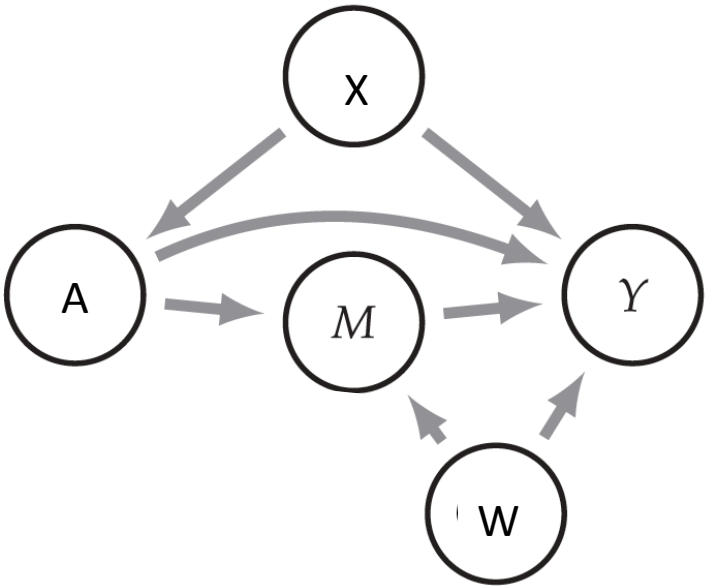
- The total effect comparing the outcome *Y* when we change the exposure from level *a** to *a* is given by $E[Y_a - Y_{a^*}]$

- For the total effect to have a causal interpretation, we must assume no unmeasured confounding of the treatment-outcome relation after accounting for potential confounding by covariates *X*

THE CONTROLLED DIRECT EFFECT

- The *controlled direct effect* (CDE), which is defined as the effect of the treatment on outcome had we intervened to set the mediator *M* to some fixed level *m*, $E[Y_{am} - Y_{a^*m}]$

- To identify the CDE, we must assume no unmeasured confounding of the total effect **(A1)** by *X*; we must also assume no unmeasured confounding by *W* of the mediator-outcome effect **(A2)**



NATURAL DIRECT AND INDIRECT EFFECTS

- In contrast to the CDE, the *natural direct effect* (NDE) sets the mediator for each individual to the level it would have been under the presence or absence of the treatment $E\left[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}\right]$

- Corresponding to the NDE is the *natural indirect effect* (NIE)

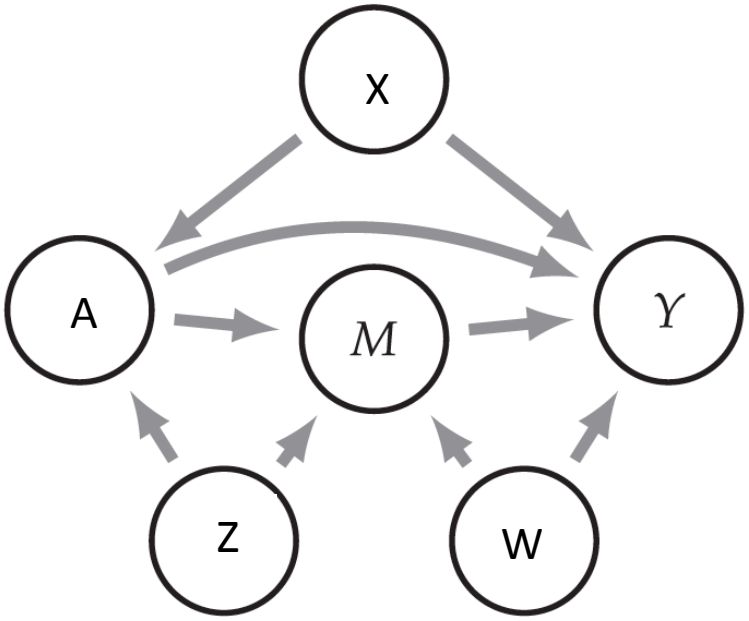
- Formally defined by $E\left[Y_{aM_a} - Y_{a_{M_{a^*}}}\right]$, the NIE answers the counterfactual question: if we were to hold the treatment constant, what would the effect be if we intervened to change the mediator from the value realized under the control condition to the value realized under the treatment condition?

NATURAL DIRECT AND INDIRECT EFFECTS

- In addition to the assumptions for identifying the CDE, additional conditions must be satisfied for a causal interpretation of the NDE & NIE

- First, no unmeasured confounding of the relation between the treatment and mediator **(A3)**, indicated by *Z*

- Second, no consequence of treatment that confounds the relation between the mediator and outcome **(A4)**



CONTROLLED OR NATURAL MEDIATED EFFECTS?

- With non-linearities, this distinction can become important
- Natural effects are attractive because they sum to the total effect
- Some have argued that natural effects are a "cross-world" counterfactual because they require us to estimate the outcome with the exposure set to a, but the mediator set to what it would have been under a* —these effects cannot be estimated in a RCT!
- From the policy standpoint, the CDE may be of greater interest
- **Note**: If we suspect *time-varying confounding* of the mediator-outcome effect then we cannot identify natural effects, but the CDE can be estimated using marginal structural models

Psychological Methods
2013, Vol. 18, No. 2, 137–150

1082-989X/13/$12.00 DOI: 10.1037/a0031034

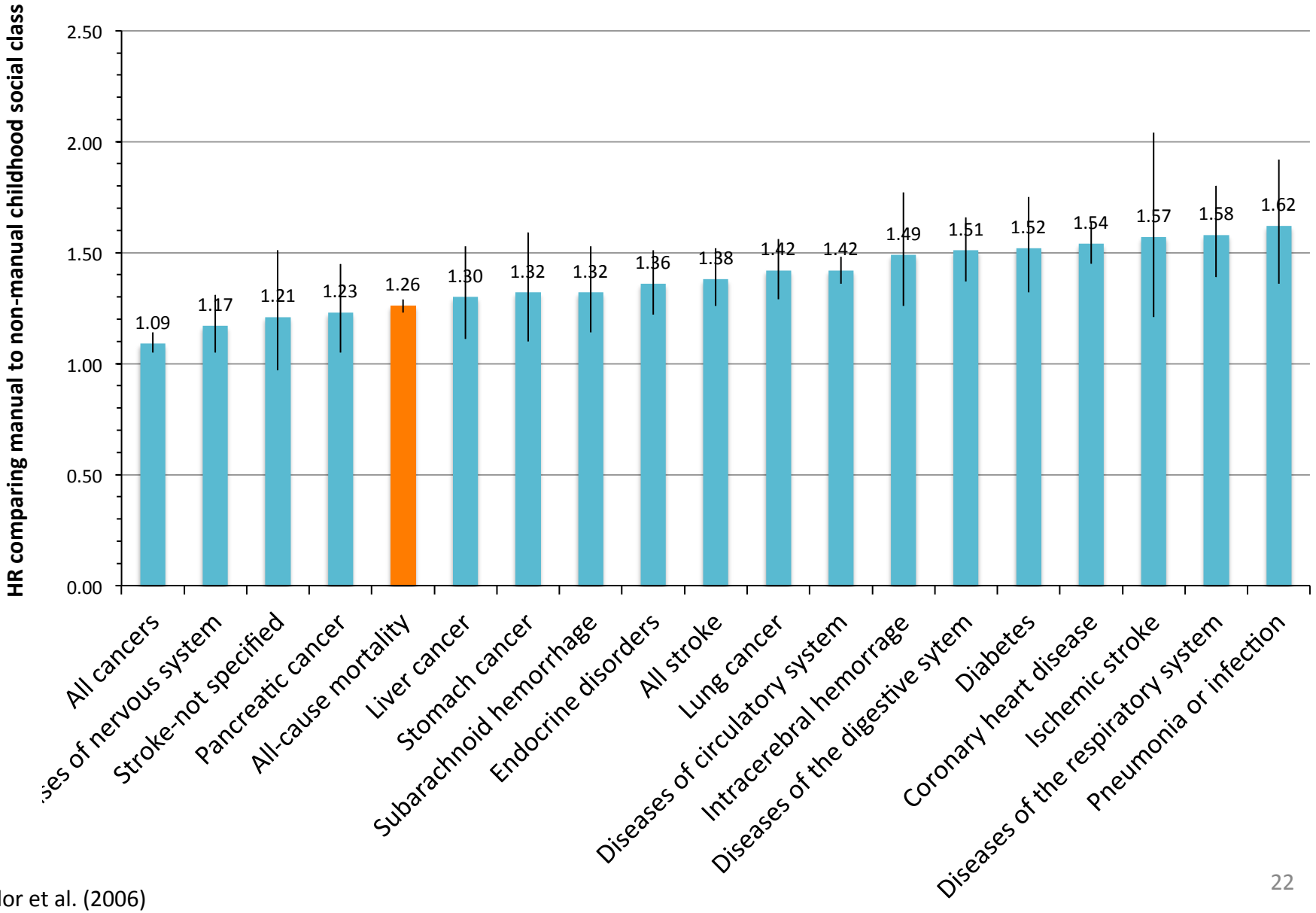
Mediation Analysis Allowing for Exposure–Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation With SAS and SPSS Macros

Linda Valeri and Tyler J. VanderWeele
Harvard University

Mediation analysis is a useful and widely employed approach to studies in the field of psychology and in the social and biomedical sciences. The contributions of this article are several-fold. First we seek to bring the developments in mediation analysis for nonlinear models within the counterfactual framework to the psychology audience in an accessible format and compare the sorts of inferences about mediation that are possible in the presence of exposure–mediator interaction when using a counterfactual versus the standard statistical approach. Second, the work by VanderWeele and Vansteelandt (2009, 2010) is extended here to allow for dichotomous mediators and count outcomes. Third, we provide SAS and SPSS macros to implement all of these mediation analysis techniques automatically, and we compare the types of inferences about mediation that are allowed by a variety of software macros.

(4) SES AND HEALTH OVER THE LIFE-COURSE

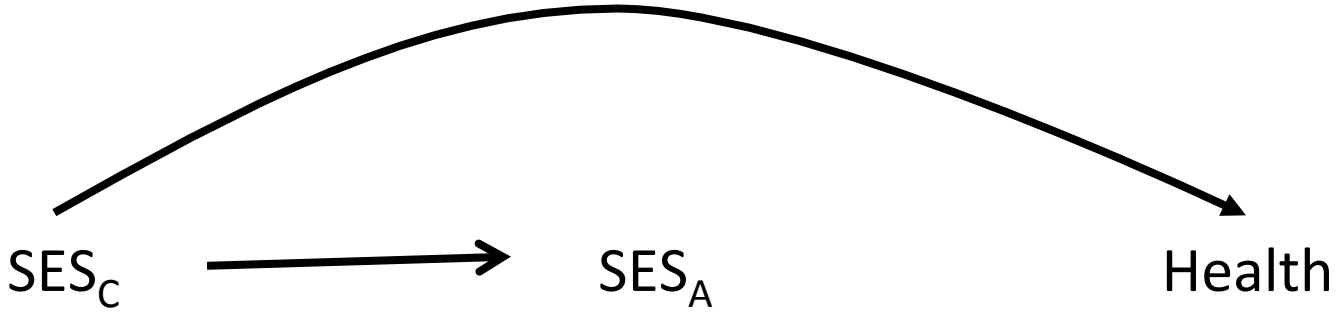
CHILDHOOD SES IS RELATED TO HEALTH IN ADULTHOOD



Lawlor et al. (2006)

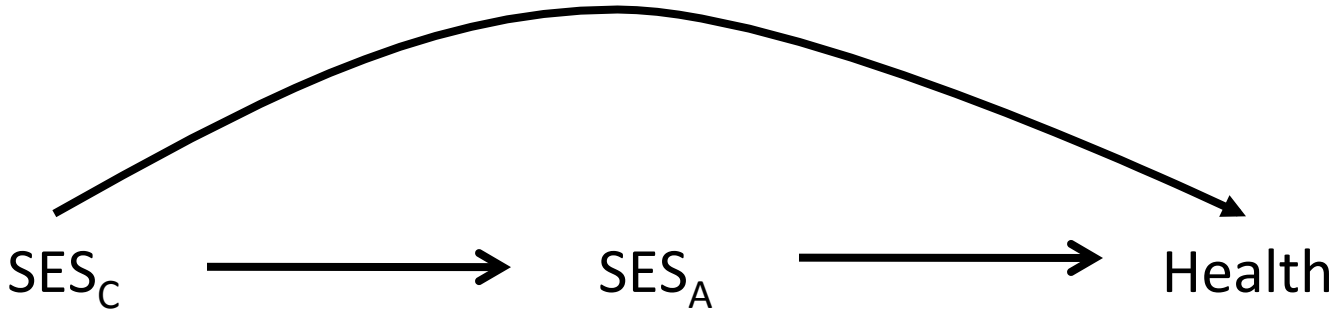
MEDIATION FOR TESTING LIFE-COURSE HYPOTHESES

①The latent effects/sensitive period model (e.g., “Barker” hypothesis)



②The pathway/social trajectory model (e.g., occupational health)

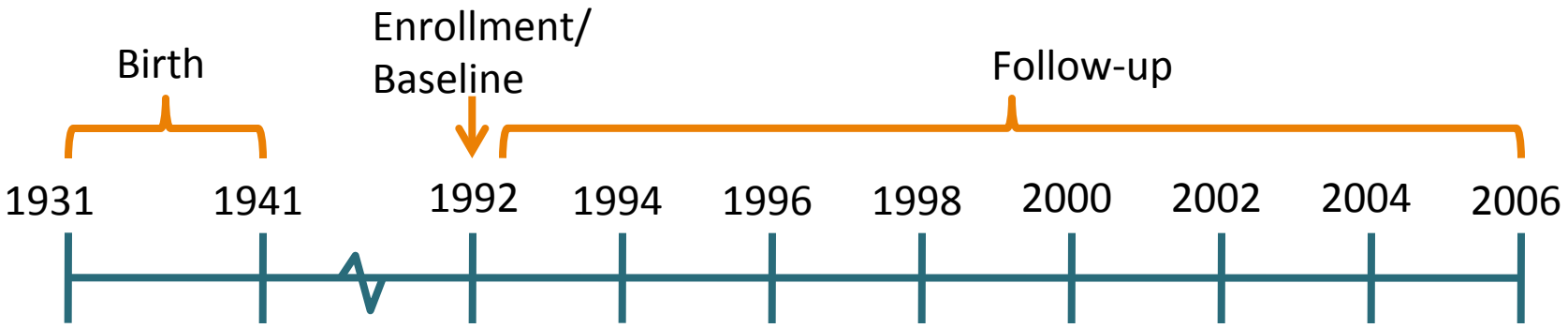
③ The cumulative exposure model



Kuh et al. (2003)

HEALTH AND RETIREMENT STUDY DESIGN

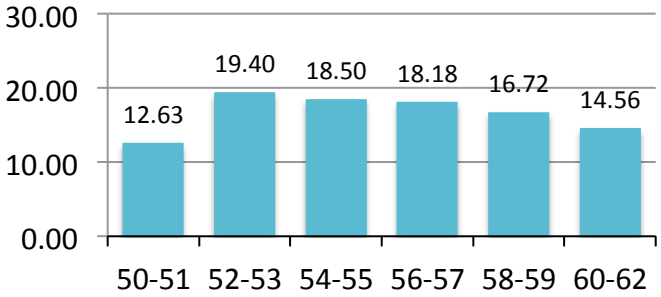
- National sample of US adults and their spouses
- Participants selected using multi-stage area probability sampling
- Biennial follow-up interviews (or proxy interviews for decedents) conducted through 2006
- 9760 enrollees in the 1992 cohort



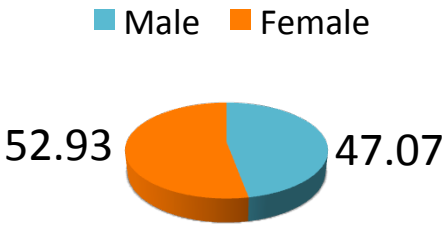
HEALTH AND RETIREMENT ANALYTICAL SAMPLE

N=9760 HRS enrollees

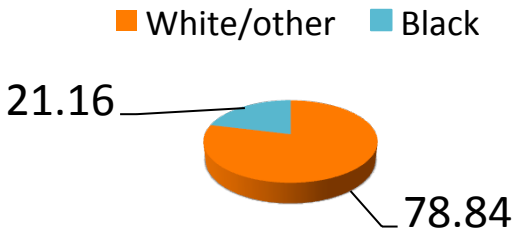
Baseline age (%)



Sex (%)



Race/ethnicity (%)



N=9265 participants with baseline and 1+ follow-up interview between 1992-2006

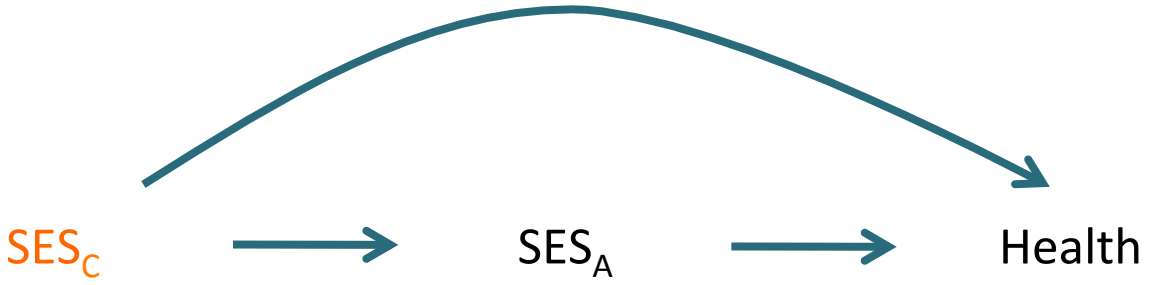
N=8318 CHD free participants (85%)	N=8278 diabetes free participants (85%)	N=9055 stroke free participants (93%)
708 cases (9%)	2001 cases (24%)	1461 cases (18%)

MEASURES

- Outcome: incident (CHD), diabetes, and stroke

MEASURES

- Outcome: incident (CHD), diabetes, and stroke
- SES_C=ƒ(father's education, mother's education, region of birth, rural/urban residence, father's occupation)

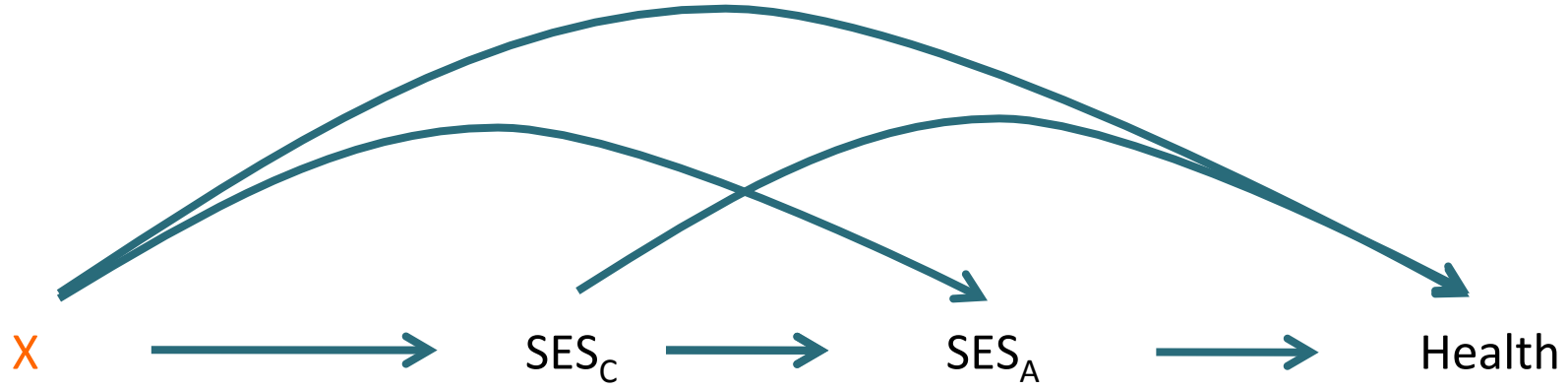


MEASURES

- Outcome: incident (CHD), diabetes, and stroke
- SES_C=ƒ(father's education, mother's education, region of birth, rural/urban residence, father's occupation)
- SES_A = ƒ(education, occupation, employment, income, wealth)

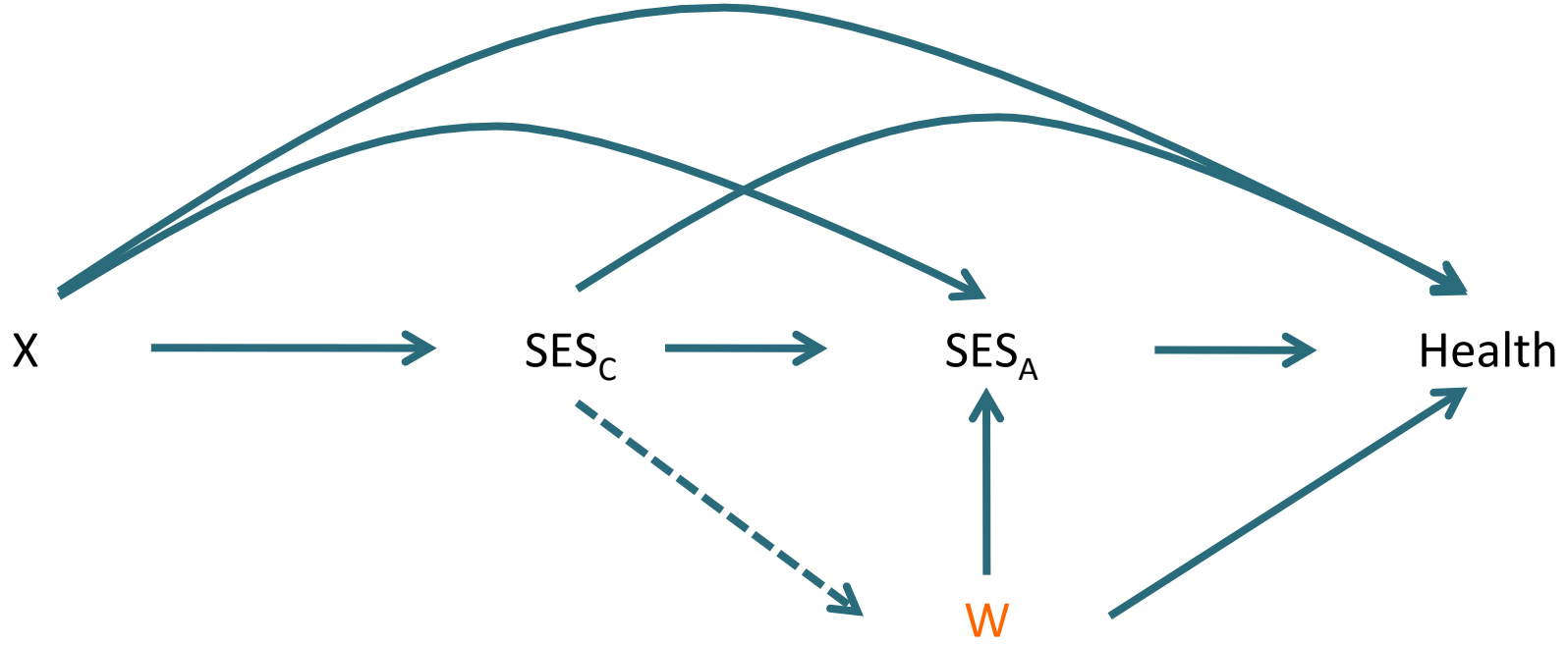
MEASURES

- Outcome: incident (CHD), diabetes, and stroke
- SES_C=ƒ(father's education, mother's education, region of birth, rural/urban residence, father's occupation)
- SES_A = ƒ(education, occupation, employment, income, wealth)
- Baseline characteristics (X) = age, sex, race, childhood health



MEASURES

- Outcome: incident (CHD), diabetes, and stroke
- SES_C=ƒ(father's education, mother's education, region of birth, rural/urban residence, father's occupation)
- SES_A = ƒ(education, occupation, employment, income, wealth)
- Baseline characteristics (X) = age, sex, race, childhood health
- Time-varying risk factors (W): smoking, high blood pressure, BMI, self-rated health, diabetes, functional limitations



STATISTICAL ANALYSES

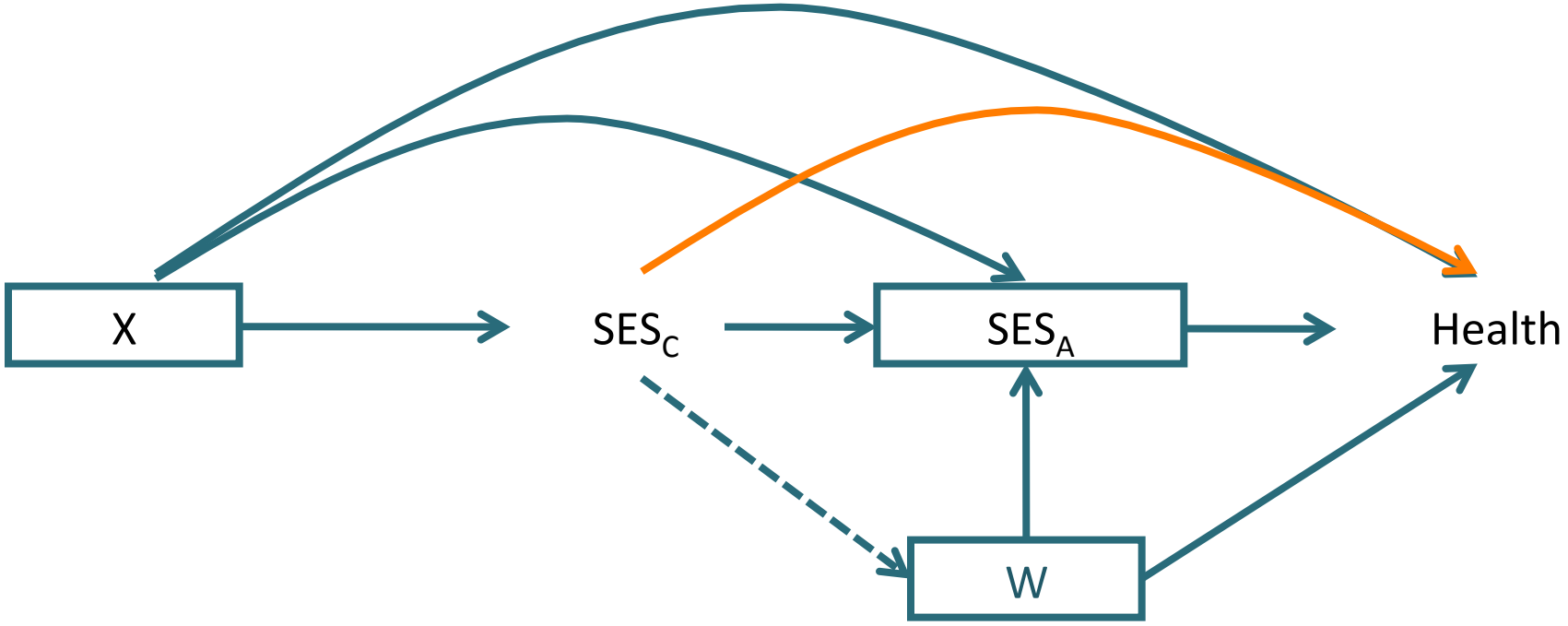
① Estimate the "total effect" of SES on mortality

R

STATISTICAL ANALYSES

① Estimate the "total effect" of SES_C on mortality

② Estimate the direct effect of SES_C on health by adjusting for SES_A (the B&K approach)

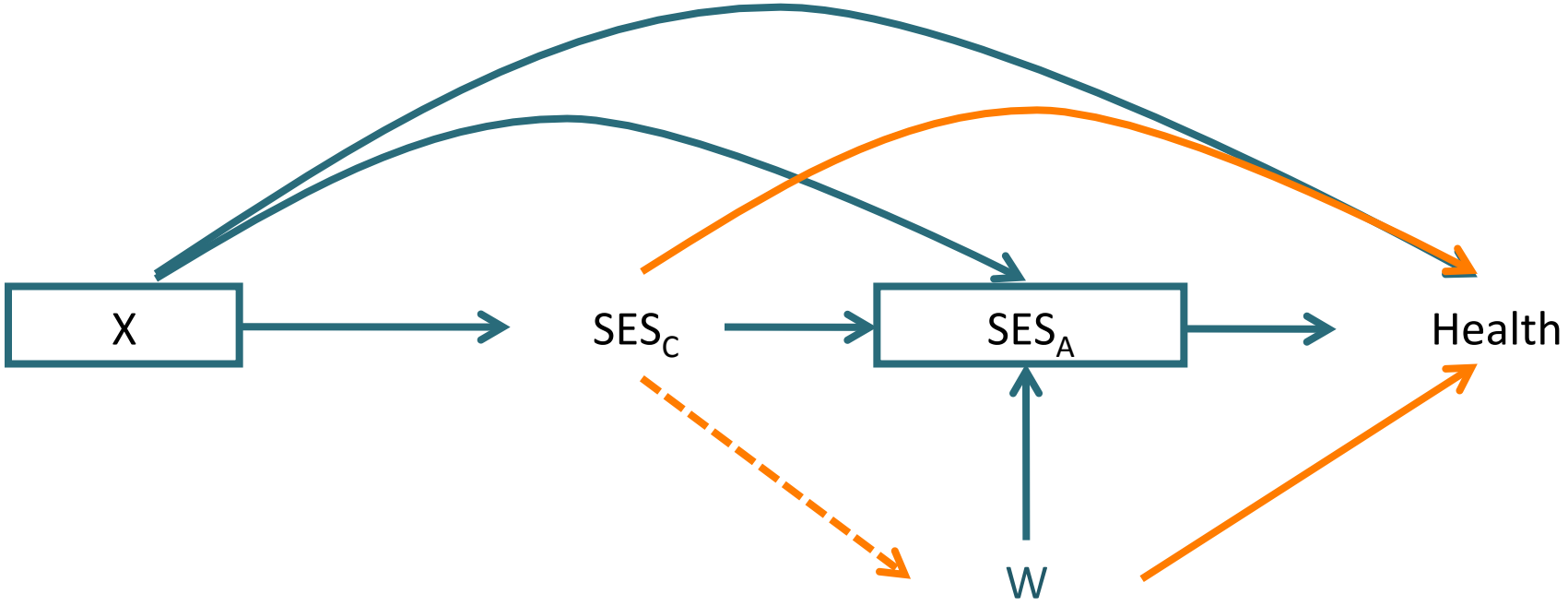


STATISTICAL ANALYSES

① Estimate the "total effect" of SES_C on health

② Estimate the direct effect of SES_C on health by adjusting for SES_A (the B&K approach)

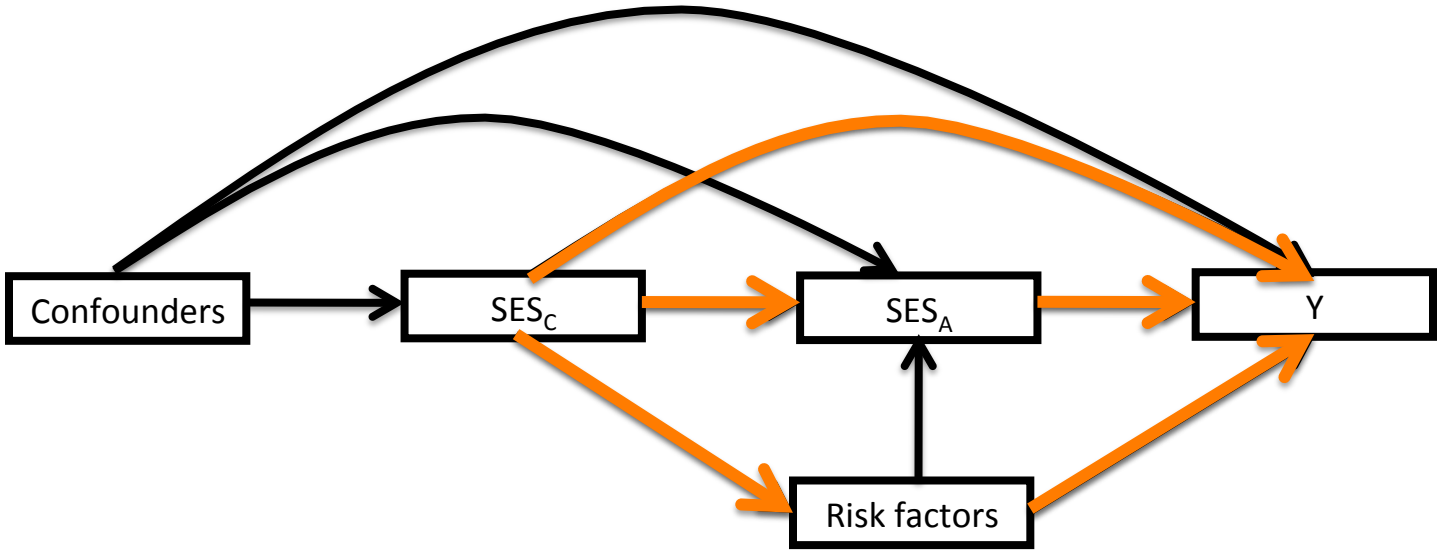
③ Estimate the direct effect of SES_C on health using a inverse probability weighted marginal structural model (MSM)



EFFECT DECOMPOSITION/REGRESSION APPROACH

① "Total effects" model:

$$\log \frac{\Pr(Y_i = 1)}{1 - \Pr(Y_i = 1)} = \beta_0 + \beta_1 SES_i^C + \beta_2 Confounders$$



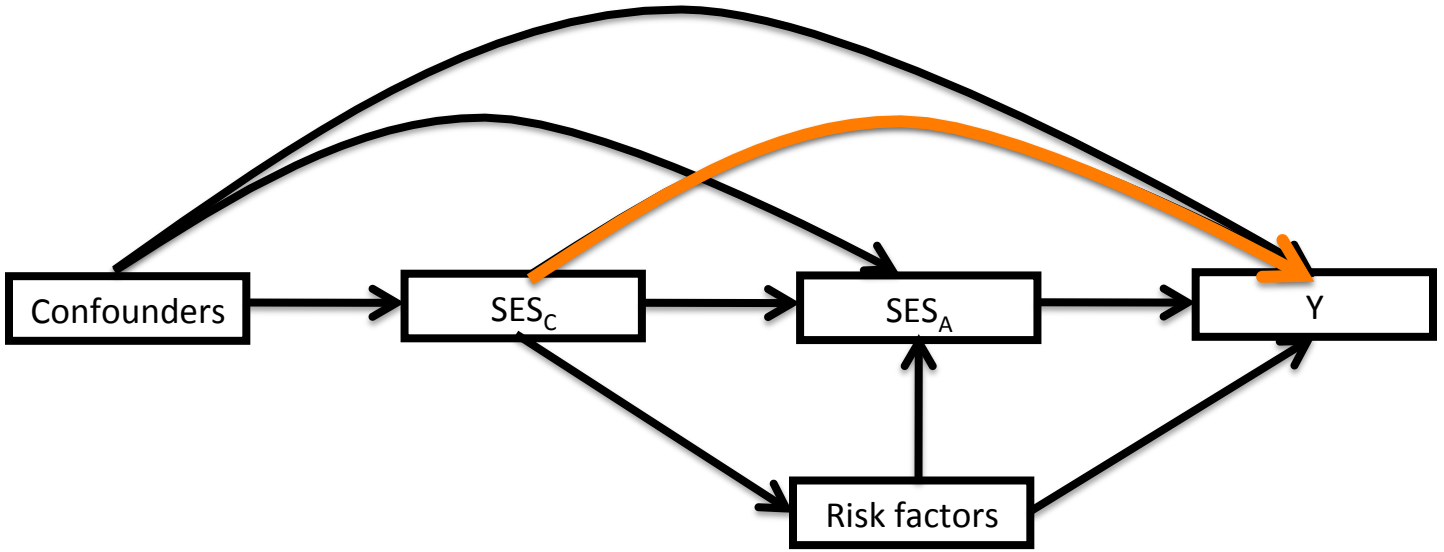
EFFECT DECOMPOSITION/REGRESSION APPROACH

① "Total effects" model:

$$\log\frac{\Pr(Y_i=1)}{1-\Pr(Y_i=1)} = \beta_0 + \beta_1 SES_i^C + \beta_2 Confounders$$

② "Mediation" model:

$$\log\frac{\Pr(Y_i=1)}{1-\Pr(Y_i=1)} = \beta_0 + \beta_1 SES_i^C + \beta_2 SES_i^A + \beta_3 Confounders + \beta_4 Risk\_factors$$

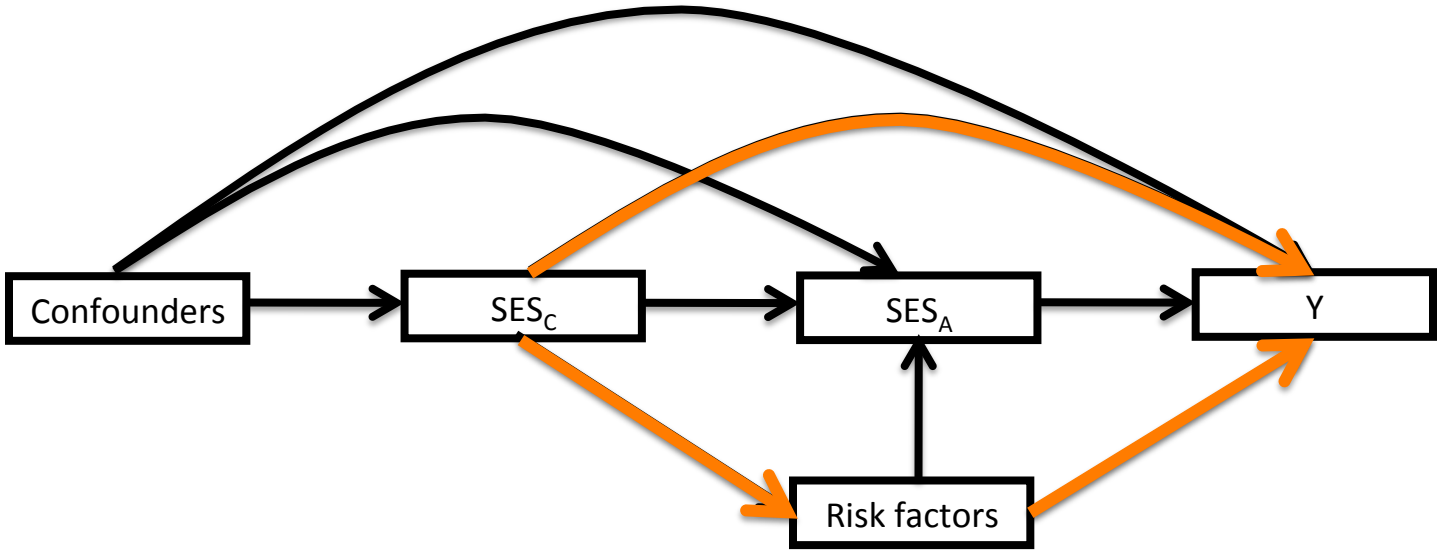


MARGINAL STRUCTURAL MODEL/IPW APPROACH

$$w_i^{SES^C} = \frac{P(SES^C = SES_i^C)}{P(SES^C = SES_i^C \mid Confounders_i)}$$

$$w_i^{SES^A} = \frac{P(SES^A = SES_i^A \mid SES_i^C)}{P(SES^A = SES_i^A \mid SES_i^C, Confounders_i, Risk\_factors_i)}$$

$$\log \frac{\Pr(Y_i = 1)}{1 - \Pr(Y_i = 1)} = \beta_0 + \beta_1 SES_i^C + \beta_2 SES_i^A \quad , W = w_i^{SES^C} * w_i^{SES^A}$$



VanderWeele (2009)

Marginal Structural Models for the Estimation of Direct and Indirect Effects

Tyler J. VanderWeele

Abstract: The estimation of controlled direct effects can be carried out by fitting a marginal structural model and using inverse probability of treatment weighting. To use marginal structural models to estimate natural direct and indirect effects, 2 marginal structural models can be used: 1 for the effects of the treatment and mediator on the outcome and 1 for the effect of the treatment on the mediator. Unlike marginal structural models typically used in epidemiologic research, the marginal structural models used to estimate natural direct and indirect effects are made conditional on the covariates.

(*Epidemiology* 2009;20: 18–26)

CHD: TOTAL EFFECT, B&K, AND MSM RESULTS

Life-course Socioeconomic Status	Total Effects (Model 1)[a] RR (95% CI)	Regression Adjustment for Adult SES, Confounders, and Risk Factors (Model 3)[c] RR (95% CI)	Marginal Structural Model (Model 4)[d] RR (95% CI)
Early-life SES[e]			
Quartile 1 (highest SES)[f]	1.00	1.00	1.00
Quartile 2	1.01 (0.90–1.13)	1.00 (0.87–1.14)	1.05 (0.93–1.18)
Quartile 3	1.12 (1.00–1.25)	1.05 (0.91–1.22)	1.23 (1.08–1.40)
Quartile 4 (lowest SES)	1.16 (1.04–1.30)	1.02 (0.86–1.21)	1.30 (1.12–1.51)
Adult SES[a]			
Quartile 1 (highest SES)[f]		1.00	1.00
Quartile 2		0.95 (0.83–1.09)	0.98 (0.87–1.09)
Quartile 3		0.83 (0.71–0.97)	0.85 (0.75–0.97)
Quartile 4 (lowest SES)		0.87 (0.73–1.05)	0.86 (0.74–1.00)

DIABETES: TOTAL EFFECT, B&K, AND MSM RESULTS

Life-course Socioeconomic Status	Total Effects (Model 1)[a] RR (95% CI)	Regression Adjustment for Adult SES, Confounders, and Risk Factors (Model 3)[c] RR (95% CI)	Marginal Structural Model (Model 4)[d] RR (95% CI)
Early-life SES[e]			
Quartile 1 (highest SES)[f]	1.00	1.00	1.00
Quartile 2	1.12 (0.97–1.29)	0.99 (0.83–1.17)	1.06 (0.91–1.23)
Quartile 3	1.31 (1.14–1.51)	0.99 (0.82–1.19)	1.15 (0.97–1.36)
Quartile 4 (lowest SES)	1.62 (1.41–1.86)	1.08 (0.88–1.32)	1.23 (1.02–1.48)
Adult SES[a]			
Quartile 1 (highest SES)[f]		1.00	1.00
Quartile 2		0.98 (0.83–1.17)	1.02 (0.87–1.18)
Quartile 3		1.06 (0.88–1.28)	1.02 (0.86–1.21)
Quartile 4 (lowest SES)		1.24 (1.00–1.54)	1.40 (1.17–1.68)

STROKE: TOTAL EFFECT, B&K, AND MSM RESULTS

Life-course Socioeconomic Status	Total Effects (Model 1)[a] RR (95% CI)	Regression Adjustment for Adult SES, Confounders, and Risk Factors (Model 3)[c] RR (95% CI)	Marginal Structural Model (Model 4)[d] RR (95% CI)
Early-life SES[e]			
Quartile 1 (highest SES)[f]	1.00	1.00	1.00
Quartile 2	0.86 (0.68–1.09)	0.74 (0.57–0.96)	0.75 (0.57–0.97)
Quartile 3	1.37 (1.11–1.69)	0.93 (0.71–1.21)	1.24 (0.95–1.61)
Quartile 4 (lowest SES)	1.59 (1.29–1.95)	0.90 (0.67–1.21)	1.29 (0.96–1.72)
Adult SES[a]			
Quartile 1 (highest SES)[f]		1.00	1.00
Quartile 2		0.98 (0.75–1.28)	0.98 (0.76–1.27)
Quartile 3		1.26 (0.95–1.67)	1.28 (0.98–1.68)
Quartile 4 (lowest SES)		1.35 (0.98–1.87)	1.49 (1.11–2.00)

Using Marginal Structural Models to Estimate the Direct Effect of Adverse Childhood Social Conditions on Onset of Heart Disease, Diabetes, and Stroke

Arijit Nandi,[a] *M. Maria Glymour,*[b] *Ichiro Kawachi,*[b] *and Tyler J. VanderWeele*[c,d]

Background: Early-life socioeconomic status (SES) is associated with adult chronic disease, but it is unclear whether this effect is mediated entirely via adult SES or whether there is a direct effect of adverse early-life SES on adult disease. Major challenges in evaluating these alternatives include imprecise measurement of early-life SES and bias in conventional regression methods to assess mediation. In particular, conventional regression approaches to direct effect estimation are biased when there is time-varying confounding of the association between adult SES and chronic disease by chronic disease risk factors.

Methods: First-reported heart disease, diabetes, and stroke diagnoses were assessed in a national sample of 9760 Health and Retirement Study participants followed biennially from 1992 through 2006. Early-life and adult SES measures were derived using exploratory and confirmatory factor analysis. Early-life SES was measured by parental education, father's occupation, region of birth, and childhood rural residence. Adult SES was measured by respondent's education, occupation, labor force status, household income, and household wealth. Using marginal structural models, we estimated the direct effect of early-life SES on chronic disease onset that was not mediated by adult SES. Marginal structural models were estimated with stabilized inverse probability-weighted log-linear models to adjust for risk factors that may have confounded associations between adult SES and chronic disease.

Results: During follow-up, 24%, 18%, and 9% of participants experienced first onset of heart disease, diabetes, and stroke, respectively. Comparing those in the most disadvantaged with the least disadvantaged quartile, early-life SES was associated with coronary heart disease (risk ratio = 1.30 [95% confidence interval = 1.12–1.51]) and diabetes (1.23 [1.02–1.48]) and marginally associated with stroke via pathways not mediated by adult SES.

Conclusions: Our results suggest that early-life socioeconomic experiences directly influence adult chronic disease outcomes.

(*Epidemiology* 2012;23: 223–232)

(5) CONCLUSIONS AND FUTURE DIRECTIONS

SUMMARY OF FINDINGS

- Regression estimates of the direct effect of early-life SES on health were attenuated relative to MSM estimates

- Marginal structural models showed a dose-response, direct effect of early-life SES on stroke and diabetes, but not CHD

- Most theories of life-course development of chronic disease suggest an underlying causal structure in which the marginal structural model would identify the causal effect of interest, but the conventional regression approach would not

THE VALUE OF MEDIATION ANALYSIS IN EPIDEMIOLOGY

- Mediating pathways are not only omnipresent in epidemiologic frameworks, but there may be value in mediation tests:

- One of the challenges of studying macro-level determinants of health over the life-course is that *"we are often at somewhat of a loss as to how to change the 'fundamental' cause of the outcome, but have more optimism that we could change a putative mediator, and the preferred policy response would obviously depend on the primary mediators"* (M Glymour)

Selected references

Ananth CV, VanderWeele TJ. Placental abruption and perinatal mortality with preterm delivery as a mediator: disentangling direct and indirect effects. Am. J. Epidemiol. Jul 1 2011;174(1):99---108.

Cole SR, Hernan MA. Fallibility in estimating direct effects. Int. J. Epidemiol. Feb 2002;31(1):163---165.

Hafeman DM. "Proportion explained": a causal interpretation for standard measures of indirect effect? Am. J. Epidemiol. Dec 1 2009;170(11):1443---1448.

Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychological methods. Dec 2010;15(4):309---334.

Kaufman JS. Marginalia: comparing adjusted effect measures. Epidemiology. Jul 2010;21(4):490---493.

Kaufman JS, Maclehose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. Epidemiol Perspect Innov. Oct 8 2004;1(1):4.

Kaufman JS. Invited commentary: Decomposing with a lot of supposing. Am. J. Epidemiol. Dec 15 2010;172(12):1349---1351; discussion 1355---1346.

Kaufman S, Kaufman JS, MacLehose RF, Greenland S, Poole C. Improved estimation of controlled direct effects in the presence of unmeasured confounding of intermediate variables. Stat. Med. Jun 15 2005;24(11):1683---1702.

Lange T, Hansen JV. Direct and indirect effects in a survival context. Epidemiology. Jul 2011;22(4):575---581.

Petersen ML, Sinisi SE, van der Laan MJ. Estimation of direct causal effects. Epidemiology. May 2006;17(3):276---284.

Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology. Mar 1992;3(2):143---155.

Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. Sep 2000;11(5):550---560.

Tchetgen Tchetgen EJ. On causal mediation analysis with a survival outcome. The international journal of biostatistics. 2011;7(1):Article 33.

Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. Psychological methods. 2013;18(2):137.

van der Laan MJ, Petersen ML. Direct effect models. The international journal of biostatistics. 2008;4(1):Article 23.

VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. Epidemiology. Jan 2009;20(1):18---26.

VanderWeele TJ. Mediation and mechanism. Eur. J. Epidemiol. 2009;24(5):217---224.

VanderWeele TJ. On the distinction between interaction and effect modification. Epidemiology. Nov 2009;20(6):863---871.

VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. Epidemiology. 2014;25(4):473-484.

Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. Am. J. Epidemiol. Dec 15 2010;172(12):1339---1348.

Vansteelandt S. Estimating direct effects in cohort and case---control studies. Epidemiology. Nov 2009;20(6):851---860.

Thanks!

INRICH organizers

Collaborators:
Maria Glymour
Ichiro Kawachi
Subu Subramanian
Tyler VanderWeele

McGill colleagues

Funding: Canada Research Chairs Program