Health Inequalities: Some Methodological Challenges

Sam Harper[1,2]

[1]Epidemiology, Biostatistics & Occupational Health, McGill University

[2]Institute for Health and Social Policy, McGill University

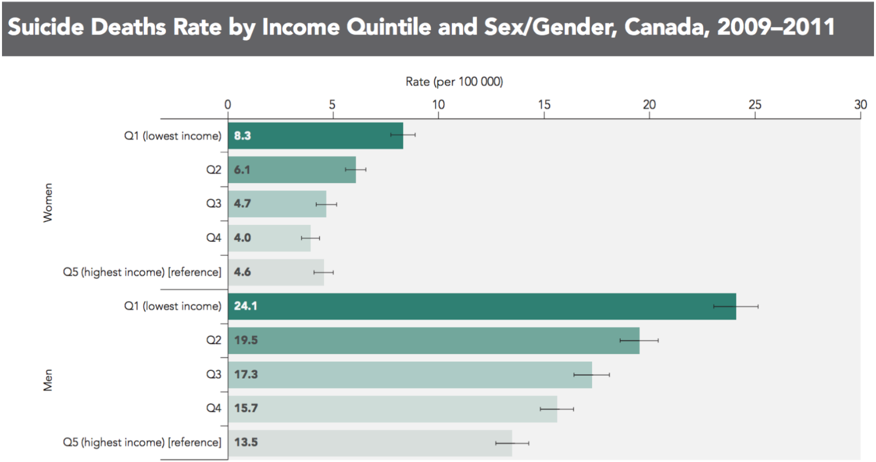
"Tackling Inequalities Master Class", Erasmus University, 1 May 2019

Where you left off...

Johan said:

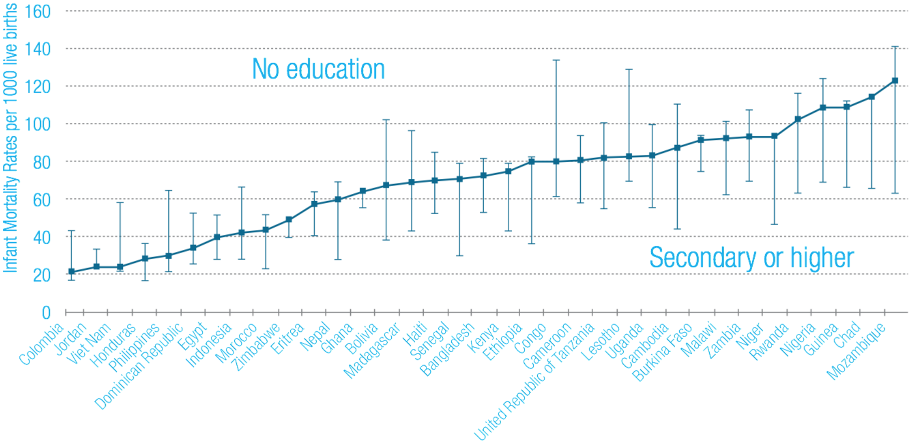
- Inequalities in material resources are still very large, even in the most generous welfare states.
- Changes in the structure of society have changed the composition and relative (dis)advantage of the lower and higher socioeconomic groups.
- There have been massive improvements, but these partly depend on behaviour change, which is easier for higher socioeconomic groups.

Much of the same story in Canada



Public Health Agency of Canada, 2018[1]

Motivating idea: what to do about health inequalities?



WHO CSDH, 2008[2]

Challenge #1

Thinking hard about causality.

"Normal science" in social epidemiology

1. Follow-up of individuals in different social groups for various health outcomes (incidence, mortality, risk factors)
2. Adjustment for various confounders/mediators (are inequalities "explained" by....X, Y, Z?).

- "Our results demonstrate that"...we should:
  - *raise* education levels
  - *increase* economic assistance to the poor
  - *remove* noxious exposures from home environments/neighborhoods
  - *reduce* psychosocial workplace hazards
  - *eliminate* hierarchies, and the like.
- These statements are based on **causal** thinking.

Advocates of the "Social" Approach: Marmot Review

- The "Marmot Review" of current and future prospects for reducing health inequalities in England emphasized several policy objectives:
  1. early childhood interventions;
  2. reducing inequalities in education;
  3. increasing employment;
  4. more progressive taxation;
  5. increasing social cohesion; and
  6. expanding preventive health care
- Argues that a strong evidence base exists for these recommendations ("But the evidence matters. Good intentions are not enough.")

https://www.local.gov.uk/marmot-review-report-fair-society-healthy-lives

Critiques of the Marmot Review: "Casual about causality"

- The policy recommendations of the Marmot report were critiqued, primarily on methodological grounds:
  - "...evidence that low-grade occupations cause poor health is far from clear"
  - "...changes in income fail to predict future changes in health"
  - "...increments to education seem to have heterogeneous effects"
  - "...the health-income link is to a large extent driven by the effect of health on income"
- These are critiques about whether social factors are **causes** of ill-health.

Stylized 'forms' of questions asked in social epidemiology

What question do most studies in social epidemiology answer?

- Do individuals who are disadvantaged with respect to social position have worse health than those who are advantaged?

Other kinds of questions that could be asked:

- Would individuals who are disadvantaged with respect to social position have better health if they were to become advantaged?
- Would individuals who are advantaged with respect to social position have worse health if they were to become disadvantaged?

These are causal questions.

Causal questions

- Will attending this class improve your proposal?
- How would you know?
  - Easy. We just need to know what your proposal looks like after attending this class, right?
  - Not exactly. We also need to know: what your proposal looks like after not attending this class.
- Problem: You are all already attending this class (oh...☹).
- We can't have data on both of these *potential* outcomes.
- This is the **fundamental problem of causal inference**.

One potential solution: random assignment

- Before class: we flip a coin.
- 50% of you get "assigned" to class; others can skip ☺
- The "treated" and "control" groups are:
  - Similar in all ways.
  - Except for which group they were assigned to!
- Later, we assess the quality of proposals.
- Any differences in proposals we can causally attribute to the class.*

*Note: the intervention may actually be useless.

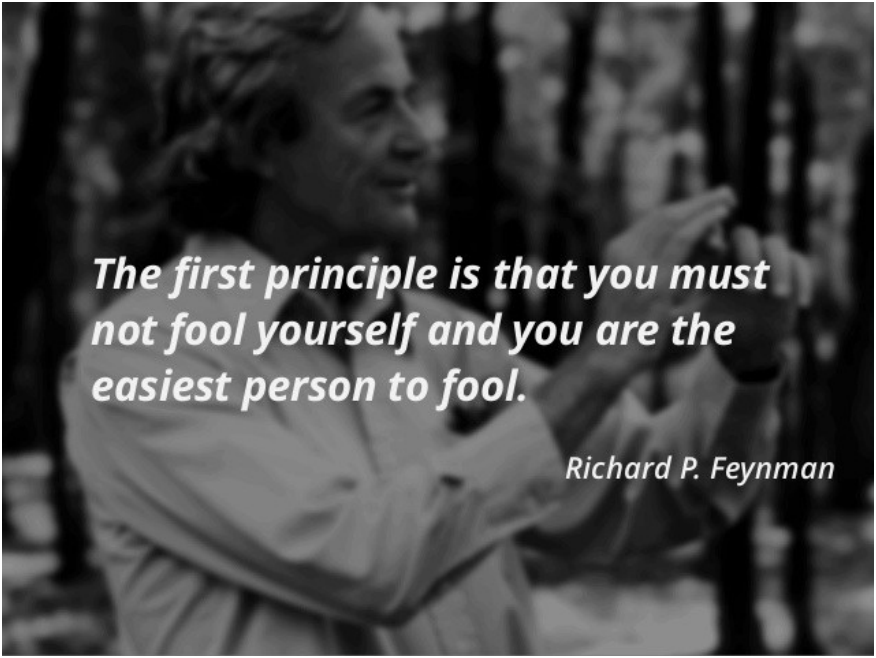
How to interpret statistical associations of health inequality?

- We have lots of statistical associations between social exposures and health.

$$X \rightarrow Y$$

- Some possible situations *consistent* with statistical associations:

1. True cause $X \rightarrow Y$
2. Chance $X \quad Y$
3. Reverse causation $Y \rightarrow X$
4. Confounding $X \leftarrow C \rightarrow Y$
5. Selection bias $X \rightarrow S \leftarrow Y$


The first principle is that you must not fool yourself and you are the easiest person to fool.
Richard P. Feynman

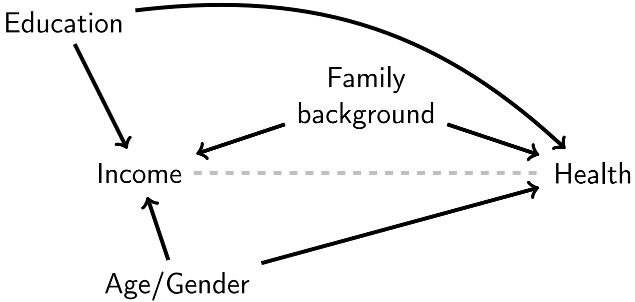
Causal relationships are challenging

- We want to know about the effect of income on health.
- What are you worried about?

$$\text{Income} \longrightarrow \text{Health}$$

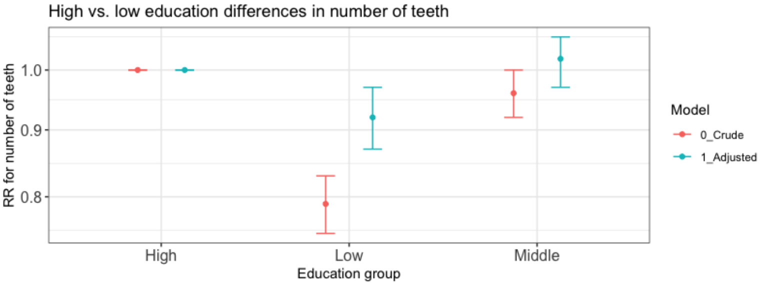
Confounding ($X \leftarrow C \rightarrow Y$)

- Might reflect other determinants (C) of income and health.
- Can you measure them all?



Controlling for confounding makes a difference

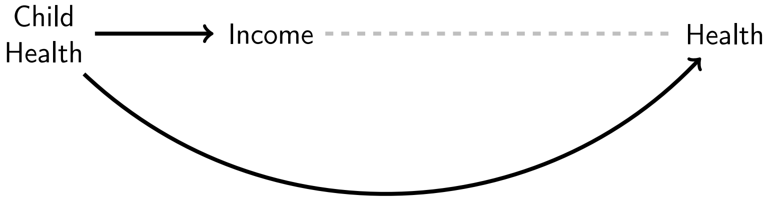
- Impact of adjustment for measured factors (housing tenure, dental attendance, smoking, cultural activities, number of owned books.) on inequalities in the number of teeth:



Duijster et al. (2017), Dutch Globe Study

Reverse causation ($Y \rightarrow X$)

- What about reverse causation?
- Could you account for it?



Whitehall II Study

- Whitehall II follows ~10,000 UK civil servants (men and women ages 35-55) recruited in 1985
- Multiple investigators have found in multiple studies that
  - Occupational health gradients exist.
  - They are not 'explained' by conventional risk factors.
  - Psychosocial factors also matter.
- Reverse causation usually not considered.

Changing the Question: Whitehall II Study

- Case and Paxson (2010) ask a different question: Does change in occupational grade ($\triangle G$) affect change in health ($\triangle H$), or does change in health affect change in occupational grade?
  - $\triangle G \rightarrow \triangle H$?
  - $\triangle H \rightarrow \triangle G$?
- They compare the **same** individuals who experienced changes in occupation or health to those who didn't change, which controls for any fixed individual characteristics, such as:
  - parental SEP
  - childhood health
  - education
- Still not random assignment to occupational grade, but likely closer than simply measuring outcomes among those achieving different grades.

Changing the Question: Whitehall II Study

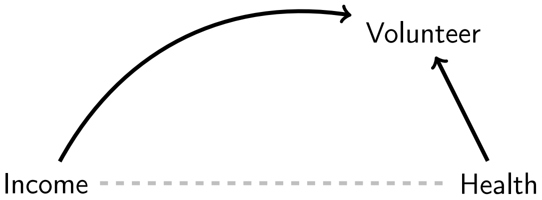
Changing the question *does* matter!

- No evidence that lagged civil service grade affects health.
- Significant effects of prior health on future employment (healthier individuals more likely to be promoted).
- Future civil service grade also predicted current health, suggesting health-related selection or unmeasured confounding.

Case and Paxson (2010)

Selection bias ($X \rightarrow S \leftarrow Y$)

- Conditioning on a **common effect** of exposure and outcome:



Does this look familiar?

- "I have no explanation for it, but I do firmly believe that modest drinking improves longevity," Kawas stated.
- Why do the most newsworthy headlines always seem the least trustworthy?

Drinking alcohol key to living past 90

01:08
Drinking alcohol key to living past 90

BY
JOE DZIEMIANOWICZ
FOLLOW

NEW YORK DAILY NEWS
Monday, February 19, 2018, 12:07 PM

Cheers to life — seriously.

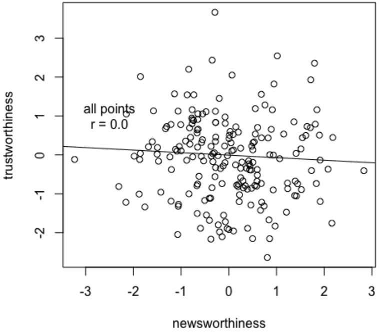
When it comes to making it into your 90s, booze actually beats exercise, according to a long-term study.

The research, led by University of California neurologist Claudia Kawas, tracked 1,700 nonagenarians enrolled in the 90+ Study that began in 2003 to explore impacts of daily habits on longevity.

How to create an association (without really trying)

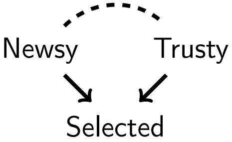
Newsy → Trusty

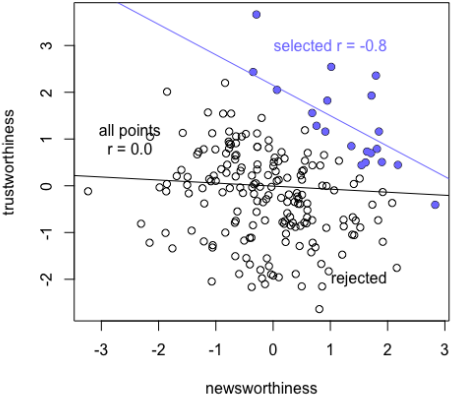
- Newsworthiness and trustworthiness shouldn't be correlated among research ideas.



McElreath, 2019

How to create an association (without really trying)



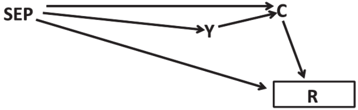


- But if I look only at the "top" proposals, I create a negative association!
- We call this "selection bias".

McElreath, 2019

Example from social epidemiology

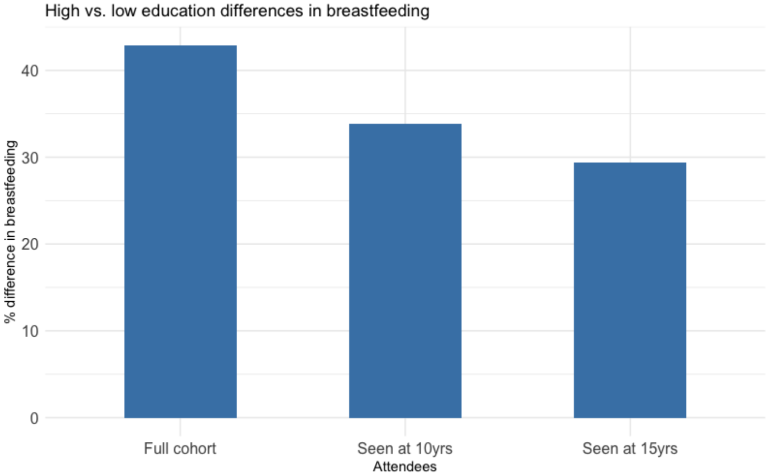
- Real example of underestimated inequalities in breastfeeding because of selection bias.
- We can only study (or "select") people that continue to participate in studies.
- Dropout (R) may be correlated with lower education (SEP), but also with difficulty in breastfeeding (C), a consequence of actual breastfeeding (Y):



Howe et al., 2013[3]

Resulting underestimate of inequalities in breastfeeding

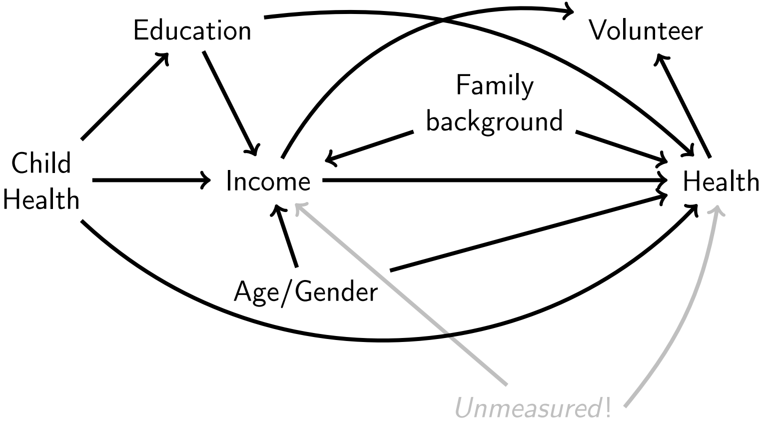
- Lower inequalities among those still participating at 15 years.



Howe et al., 2013[3]

Causal relationships are challenging

- Forget something? What about things you *didn't* measure?
- Are you convinced no other sources of bias?



Be careful, and skeptical

- Correlations between social factors and health are easy to find.
- They do not necessarily reflect **causal** relationships.
- Most important thing is to search hard for alternative explanations.
- Important to consider these factors in designing a study about health inequalities.

Challenge #2

What to do about inequalities?

How we got here

- Longstanding concerns about persistent health inequalities.
- Challenges with causal inference of social exposures.
- Much of social epidemiology focused on trying to “explain” away inequalities.
- More recent calls to think about interventions.

Policymakers' Context for Health Inequalities

- Interviews with UK health policymakers in the early 2000s were disappointing for those wanting their research to have "impact".
- The "inverse evidence law" (Petticrew 2004[4]): "...relatively little [evidence] about some of the wider social economic and environmental determinants of health, so that with respect to health inequalities we too often have the right answers to the wrong questions."
- Problem of "policy-free evidence": an abundance of research that does not answer clear, or policy relevant questions.

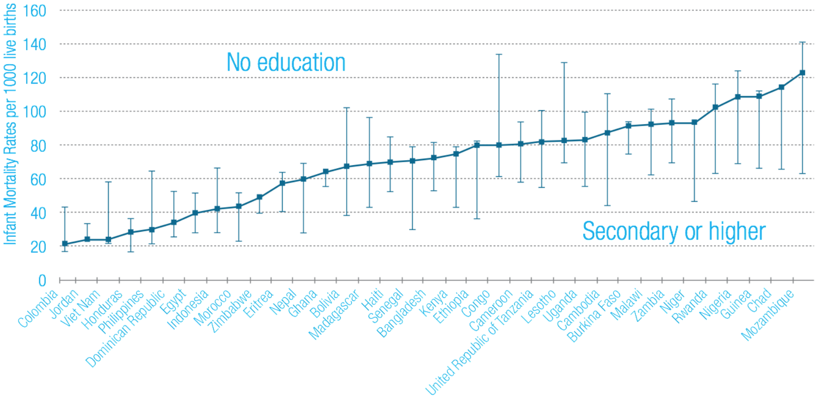
How to make social epidemiology relevant to policy?

..."researchers may improve the likelihood of their research having a wider policy impact by focusing less on describing the problem and more on ways to solve it, working closely with those who are charged with the task of tackling health inequalities, and others who can contribute to the creation of a climate in which reducing health inequalities is perceived to be not only politically possible but publicly desirable."

Bambra et al., 2011 "A labour of Sisyphus?"[5]

A thought experiment

Let's assume that the education-based gradients in infant mortality, perhaps counter to fact, reflect a causal effect, and you were charged with eliminating these inequalities. . . what would you do?



WHO CSDH, 2008[2]

How to reduce inequalities in infant mortality?

Should we:

- Increase secondary or higher education by making it free?
- Increase secondary education by making it compulsory?
- Increase secondary education by increasing school quality?
- Build more secondary schools?

- Increase access to maternal care among less-educated women?
- Increase immunization among kids of less-educated mothers?
- Increase access to family planning?
- Increase access to household resources among less-educated mothers?

- All of the above?
- Some of the above
- None of the above?

How will you know whether or not it worked?

What is impact evaluation?

- An impact evaluation "assesses the changes in the well-being of individuals that can be attributed to a particular project, program, or policy"(Gertler, 2016 [6]).
- "Impact evaluation asks about the difference between what happened with the program and what would have happened without it (referred to as the counterfactual)."(Savedoff, 2006[7]).
- The "impact" can be defined as the change in the outcome that can be causally attributed to the program (Ravallion, 2008 [8])
- Impact evaluation studies are among a range of complementary techniques for supporting evidence-based policymaking.

What's the problem?

- We are mainly (though not exclusively) interested in causal effects.
- We want to know:
  - Did the program work? If so, for whom? If not, why not?
  - If we implement the program elsewhere, should we expect the same result?
- These questions involve counterfactuals about what would happen **if** we intervened to do something.
- These are causal questions.

- **Causal effect**: Do individuals randomly assigned (i.e., SET) to treatment have better outcomes?

$$E\left(Y|SET\left[Treated\right]\right) - E\left(Y|SET\left[Untreated\right]\right)$$

- **Association**: Do individuals who happen to be treated have better outcomes?

$$E\left(Y|Treated\right) - E(Y|Untreated)$$

Randomized Trials and Quasi-Experimental Studies

RCTs, Defined

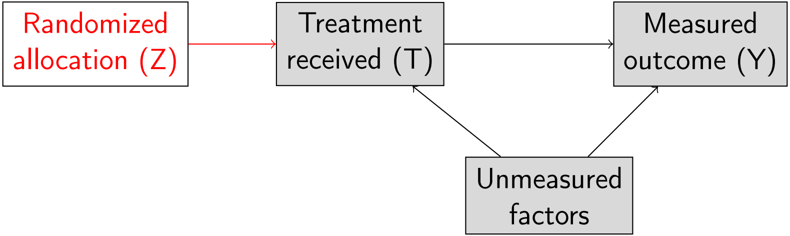
An RCT is characterized by: (1) comparing treated and control groups; (2) assigning treatment randomly; and (3) investigator does the randomizing.

- In an RCT, treatment/exposure is assigned by the investigator
- In observational studies, exposed/unexposed groups exist in the source population and are selected by the investigator.

- Good quasi-experiments do (1) and (2), but not (3).
- Because there is no control over assignment, the credibility of quasi-experiments hinges on how good "as-if random" approximates (2).

Benefits of randomization

- Randomization means that we can generally estimate the causal effect without bias.
- Randomization guarantees exchangeability on measured and unmeasured factors.



- Howden-Chapman et al. (2007)
- Cluster-RCT to retrofit houses with insulation.
- Randomization leads to:
    - balance on measured factors.
    - balance on unmeasured factors.
- Unmeasured factors cannot bias the estimate of the exposure effect.

Table 1
Baseline information for each group

	Control group	Intervention group
Household factors at baseline		
Number recruited	672	680
Number returned questionnaire	652	658
Dwelling reported in "poor" or "very poor" condition (%)	116/644 (18)	118/653 (18)
Condensation reported (%)	566/633 (89)	577/640 (90)
Non-condensation dampness reported (%)	413/613 (66)	437/641 (68)
Mould reported (%)	481/643 (75)	490/651 (75)
Dwelling cold "always" or "most of the time" (%)	452/647 (70)	473/651 (73)

Problem of Social Exposures

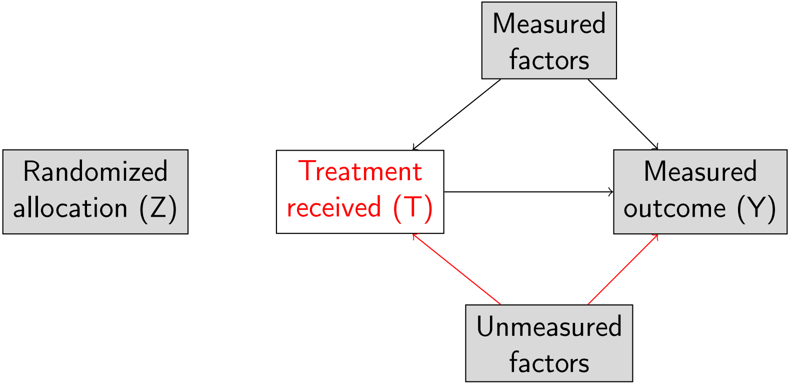
- Many social exposures/programs cannot be randomized by investigators:
  - Unethical (poverty, parental social class, job loss)
  - Impossible (ethnic background, place of birth)
  - Expensive (neighborhood environments)
- Some exposures are hypothesized to have long latency periods (many years before outcomes are observable).
- Effects may be produced by complex, intermediate pathways.
- We need alternatives to RCTs.

Consequences of non-randomized treatment assignment

- If we are not controlling treatment assignment, then who is?
- Policy programs do not typically select people to treat at random.
  - Programs target those that they think are most likely to benefit.
  - Programs implemented decisively non-randomly (e.g., provinces passing drunk driving laws in response to high-profile accidents).
  - Governments deciding to tax (or negatively tax) certain goods.
- People do not choose to participate in programs at random.
  - Screening programs and the worried well.
  - People who believe they are likely to benefit from the program.

Illustration of the problem

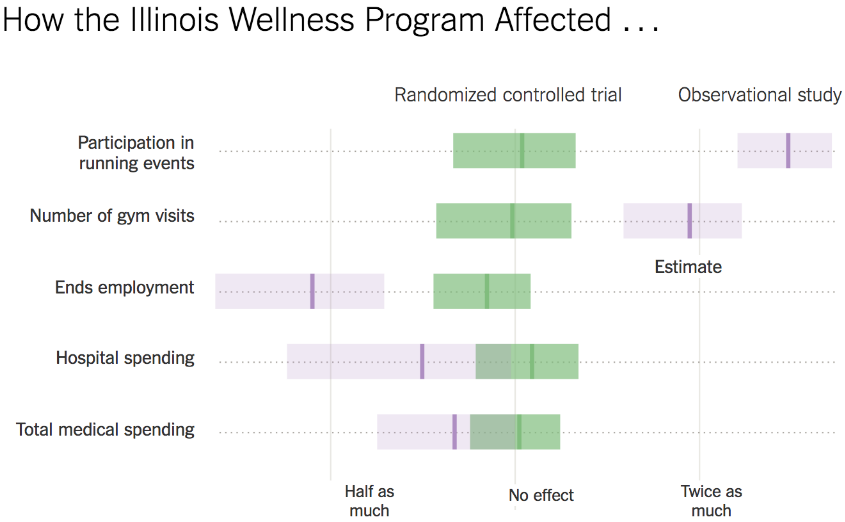
- Non-randomized designs typically start with observing treated and untreated groups, so more assumptions are necessary.
- In particular we should be worried about unmeasured (or mismeasured) factors that may lead to bias:



Why we worry about observational studies

- Recent evaluation of "Workplace Wellness" program in US state of Illinois
- Treatment: biometric health screening; online health risk assessment, access to a wide variety of wellness activities (e.g., smoking cessation, stress management, and recreational classes).
- Randomized evaluation:
  - 3,300 individuals assigned treated group.
  - 1,534 assigned to control (could not access the program).
- Also analyzed as an observational study:
  - comparing "participants" vs. non-participants in treated group.

Jones et al. 2018 [9]



Carroll, *New York Times*, Aug 6, 2018.

Observational studies may not convince skeptics

- Many observational studies show higher IQs for breastfed children.
- All generally rely on regression adjustment.
- Hard to avoid the issue of residual confounding.
  - "I would argue that in the case of breastfeeding, this issue is impossible to ignore and therefore any study that simply compares breast-fed to formula-fed infants is deeply flawed. That doesn't mean the results from such studies are necessarily wrong, just that we can't learn much from them."
- Can quasi-experiments convince a skeptic like this?

Oster (2015). http://fivethirtyeight.com/features/everybody-calm-down-about-breastfeeding/

Ex: Education and CVD incidence in Australia

Many low p-values. Is "no other unmeasured differences" credible?

Table 1 Characteristics of 38 355 subjects in the Melbourne Collaborative Cohort Study at baseline (1990–1994)

		Highest level of education attained				
		Completed tertiary* n=8588	**Completed secondary†** n=7882	**Some secondary‡** n=14543	**Primary only§** n=7342	**p Value for trend¶**
Male	n (%)	4025 (47%)	3776 (48%)	4680 (32%)	2780 (38%)	<0.001
Female	n (%)	4563 (53%)	4106 (52%)	9863 (68%)	4562 (62%)	<0.001
Age (years)	(Mean, SD)	51.6 (8.4)	54.5 (8.8)	55.7 (8.5)	57.8 (7.1)	<0.001
Country of birth, n (%)	Australia, New Zealand or northern Europe (n=28 835)	8263 (96%)	6814 (86%)	12696 (87%)	1062 (14%)	<0.001
	Southern Europe (n=9520)	325 (4%)	1068 (14%)	1847 (13%)	6280 (86%)	<0.001
Behavioural risk factors						
Current smoker	n (%)	574 (7%)	960 (12%)	1828 (13%)	947 (13%)	<0.001
Vegetable intake (times/day)	Mean (SD)	5.7 (3)	5.3 (3)	5.2 (3)	5.8 (4)	1.000
Fruit intake (times/day)	Mean (SD)	4.4 (3)	4.0 (3)	3.9 (3)	4.7 (4)	0.007
Saturated fat intake (g/day)	Mean (SD)	35.0 (15)	34.3 (16)	33.7 (16)	30.3 (18)	<0.001
Current drinker	n (%)	7061 (82%)	5883 (75%)	9397 (65%)	3666 (50%)	<0.001
Alcohol intake, current drinkers (g/day)	Median (IQR)	14 (5,26)	13 (4,26)	10 (3, 23)	15 (4,30)	<0.001
Physical activity (% inactive)	n (%)	1224 (14%)	1520 (19%)	3238 (22%)	2546 (35%)	<0.001
Social connection						
Living alone	n (%)	1514 (18%)	1274 (16%)	2250 (15%)	498 (7%)	<0.001

Beauchamp et al., 2010 [10]

Challenge #3

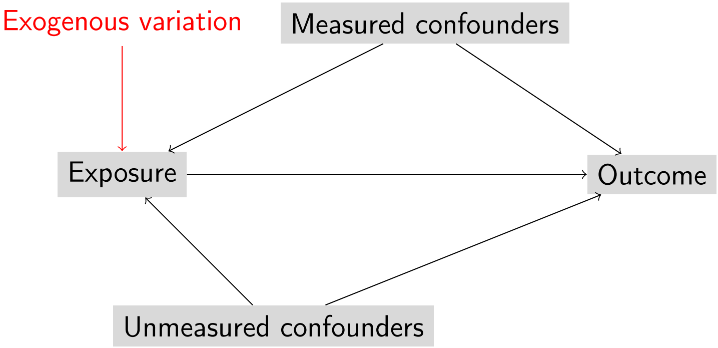
Effective study designs

How do quasi-experiments help?

- Quasi-experiments aim to mimic RCTs.
- Typically "accidents of chance" that create:
  1. Comparable treated and control units
  2. Random or "as-if" random assignment to treatment.
- Well-designed quasi-experiments control for (some) sources of bias that cannot be adequately controlled using regression adjustment.
- More credible designs also help us to understand the relevance of other factors that may be implicated in generating inequalities.

Selection on "observables" and "unobservables"

- Observables: Things you measured or can measure
- Unobservables: Things you can't measure (e.g., innate abilities)
- Exogenous variation: predicts exposure but (**we assume**) not associated with anything else [mimicking random assignment].



Strategies based on observables and unobservables

- Most observational study designs select on observables:
  - Stratification
  - Regression adjustment
  - Matching (propensity scores, etc.)
- Quasi-experimental strategies that select on unobservables:
  - Interrupted time series (ITS)
  - Difference-in-differences (DD)
  - Synthetic controls (SC)
  - Instrumental variables (IV)
  - Regression discontinuity (RD)

Some *potential* sources of natural experiments

- Differential distance to care (people rarely choose neighborhood based on health care services).
- Law changes (unlikely to be influenced or chosen by participants).
- Eligibility for social programs (thresholds are arbitrary).
- Genes (segregation of alleles during meiosis is random).
- Weather events (hard to predict).
- Clinical guidelines (arbitrary thresholds).
- Historical geographic features of environment (can't be chosen by current residents).

Pre-post and ITS approaches

Essence of ITS studies

Interrupted time series studies use routine data collected at equally spaced intervals of time before and after an intervention, and do not necessarily require a control group.

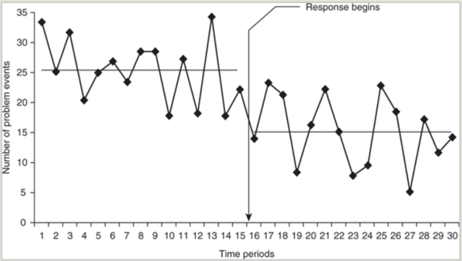


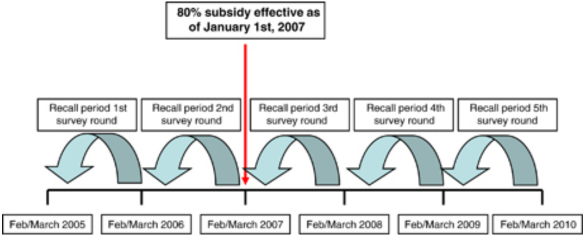
Figure 4.1. Hypothetical Example of a Simple Interrupted Time Series Design.

Thyer 2012 [11]

ITS/pre-post example

De Allegri et al. The **impact** of targeted subsidies for facility-based delivery on access to care and equity – Evidence from a population-based study in rural Burkina Faso. *J Public Health Policy* 2012;33:439–453

> *...the first population-based impact assessment of a financing policy introduced in Burkina Faso in 2007 on women's access to delivery services. The policy offers an 80 per cent subsidy for facility-based delivery. We collected information on delivery... from 2006 to 2010 on a representative sample of 1050 households in rural Nouna Health District. Over the 5 years, the proportion of facility-based deliveries increased from 49 to 84 per cent (P<0.001).*

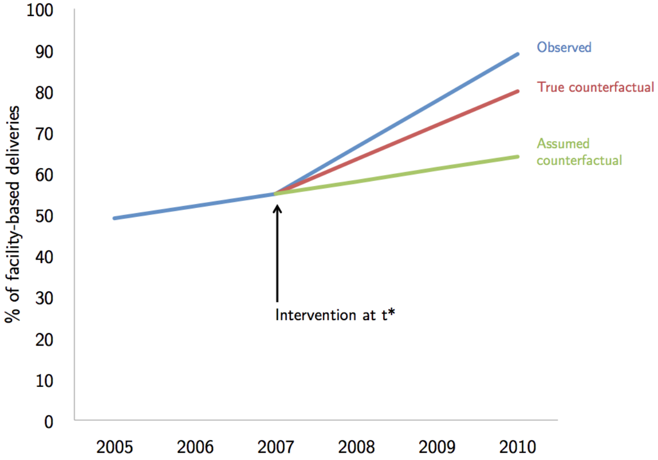


Potential possibilities with a single treated unit

- Can consider this as an interrupted time-series.
- Authors are making (implicit) assumptions about the trajectory of counterfactual outcomes.
- Certainly possible that the true impact of the intervention could have raised rates from 49% to 84%.
- But is it **plausible**? What assumptions are needed?

How credible are your assumptions?

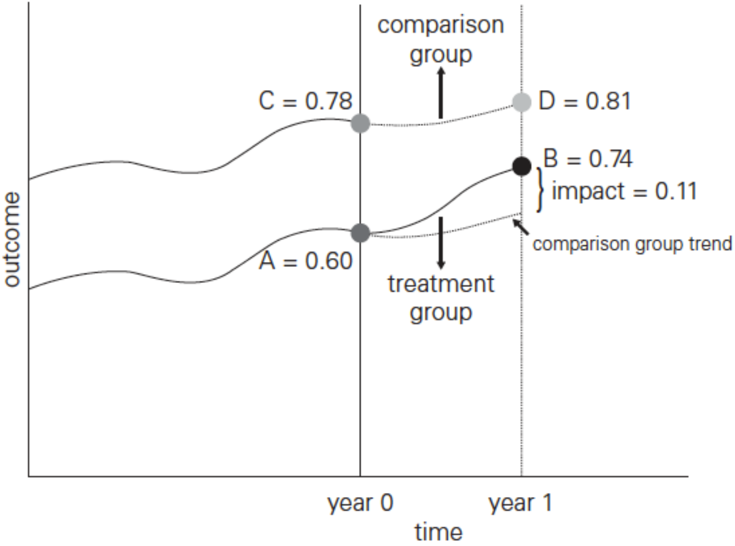
- Assumption by extrapolating pre-intervention trend may substantially overestimate the effect of any intervention



How does adding a control group help?

- Pre/post in a control group helps by differencing out any **time-invariant** characteristics of both groups.
  - Many observed factors don't change over the course of an intervention (e.g., geography, parents' social class, birth cohort).
  - Any time-invariant *unobserved* factors also won't change over intervention period.
  - We can therefore effectively control for them.
- Measuring same units before and after a program cancels out any effect of all of the characteristics that are unique to units of observation and that do not change over time.
- This leads to **difference-in-differences**.

Visual Intuition of (good) Difference-in-Differences



A US example

Effect of Massachusetts healthcare reform on racial and ethnic disparities in admissions to hospital for ambulatory care sensitive conditions: retrospective analysis of hospital episode statistics

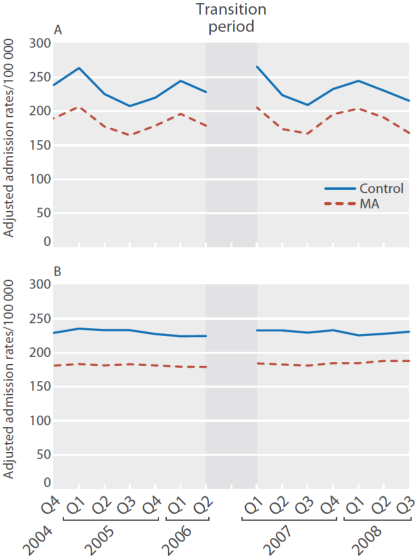
Danny McCormick,[1] Amresh D Hanchate,[2,3] Karen E Lasser,[3] Meredith G Manze,[3] Mengyun Lin,[3]
Chieh Chu,[3] Nancy R Kressin[2,3]

- Evaluated impact of MA reform on hospital admissions.
- Compared MA to nearby states: NY, NJ, PA.
- Intervention "worked": % uninsured halved (12% to 6%) from 2004-06 to 2008-09.
- No change in disparities in admission rates between blacks and whites (-1.9%, -8.5% to 5.1%)

McCormick et al. 2015 [12]

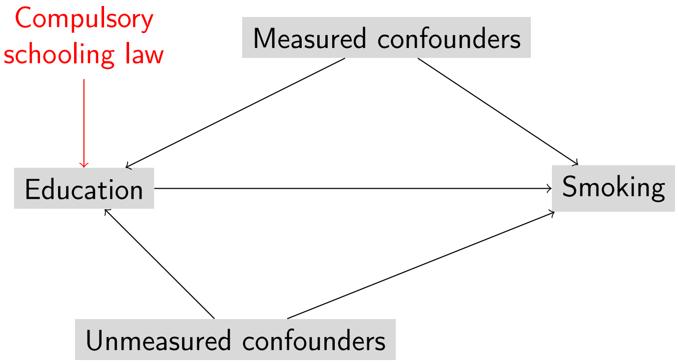
Visual evidence: comparable pre-intervention trends

- Adds credibility to assumption that post-intervention trends would have been similar in the absence of the intervention.
- "Null" results help focus on alternative mechanisms linking disadvantage to hospital admissions.



Example of instrumental variable: Policies

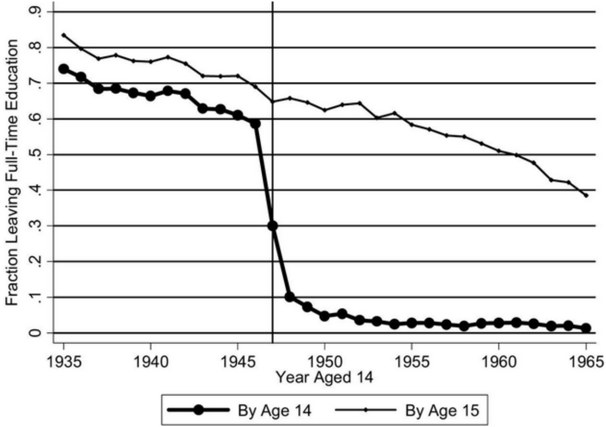
- Does education (T) affect smoking (Y)?
- **Instrument**: changes in compulsory schooling laws [mimicking random assignment].



Glymour et al. 2008[13], Hamad et al. 2018[14], Galama et al. 2018[15]

What does a quasi-experiment look like?

Fraction left full-time education by year aged 14 and 15 (Great Britain)



The lower line shows the proportion of British-born adults aged 32 to 64 from the 1983 to 1998 General Household Surveys who report leaving full-time education at or before age 14 from 1935 to 1965. The upper line shows the same, but for age 15. The minimum school-leaving age in Great Britain changed in 1947 from 14 to 15 [Oreopoulos 2006].

Contents lists available at ScienceDirect

Social Science & Medicine

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

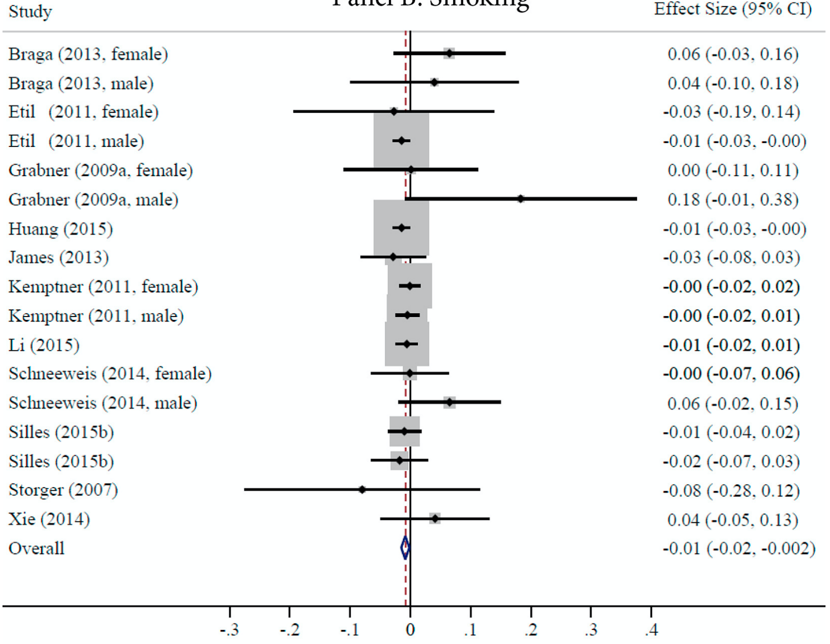
Review article

How and why studies disagree about the effects of education on health: A systematic review and meta-analysis of studies of compulsory schooling laws

Rita Hamad[a,*], Holly Elser[b,c], Duy C. Tran[c], David H. Rehkopf[c], Steven N. Goodman[c]

[a] University of California San Francisco, Philip R. Lee Institute for Health Policy Studies, Department of Family & Community Medicine, 995 Potrero Avenue, Building 80, Ward 83, San Francisco, CA, 94110, USA
[b] University of California Berkeley, School of Public Health, Division of Epidemiology, Berkeley, CA, USA
[c] Stanford University, School of Medicine, Stanford, CA, USA

Panel B. Smoking



Potential drawbacks of quasi-experimental approaches

- How good is "as-if" random? (need "shoe-leather")
- Credibility of additional (modeling) assumptions.
- Relevance of the intervention.
- Relevance of population.

Freedman 1991 [16], Rosenbaum 2017 [17]

What are quasi-experiments good for?

1. To understand the effect of treatments *induced by policies* on outcomes, e.g., Policy $\rightarrow$ Treatment $\rightarrow$ Outcome:
   - Environmental exposures.
   - Education/income/financial resources.
   - Access to health care.
   - Health behaviors.

2. To understand the effect of policies on outcomes, e.g., Policy $\rightarrow$ Outcome:
   - Taxes, wages.
   - Environmental legislation.
   - Food policy.
   - Employment policy.
   - Civil rights legislation.

Glymour 2014 [18]

Also, consider experimenting!

- RCT $\neq$ Gold standard, but can be very powerful and convincing.
- We can control aspects of programs/policies to experimentally increase the probability of exposure in one group vs. another:
  - **Access**: we can randomly select which people are offered access to a program (most common).
  - **Timing**: we can randomly select when people are offered access to a program.
  - **Encouragement**: we can randomly select which people are given encouragement or incentive to participate.
- Each of these aspects can be varied for individuals or groups.

Concluding thoughts

- Causal inference with social exposures is hard.
- Serious consideration of alternative explanations are needed.
- Quasi-experiments are useful in theory, but difficult to find in practice.
- They are still observational: key issue is credibility of assumptions.
- Actual experiments may also be relevant for policy.

Acknowledgements

- Canadian Institutes for Health Research
- Salary award from Fonds de recherche du Québec – Santé
- Smarter Choices for Better Health Initiative, Erasmus University

Thank you!

sam.harper@mcgill.ca @sbh4th

References I

[1] Public Health Agency of Canada. Key Health Inequalities in Canada: A National Portrait. Public Health Agency of Canada; 2018. Available from: https://www.canada.ca/content/dam/phac-aspc/documents/services/publications/science-research/hir-full-report-eng_Original_version.pdf.

[2] Commission on Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health: final report of the commission on social determinants of health. Geneva: World Health Organization; 2008.

[3] Howe LD, Tilling K, Galobardes B, Lawlor DA. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. Epidemiology. 2013 Jan;24(1):1–9.

[4] Petticrew M, Whitehead M, Macintyre SJ, Graham H, Egan M. Evidence for public health policy on inequalities: 1: the reality according to policymakers. J Epidemiol Community Health. 2004 Oct;58(10):811–6.

[5] Bambra C, Smith KE, Garthwaite K, Joyce KE, Hunter DJ. A labour of Sisyphus? Public policy and health inequalities research from the Black and Acheson Reports to the Marmot Review. J Epidemiol Community Health. 2011 May;65(5):399–406.

[6] Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CM. Impact evaluation in practice. World Bank Publications; 2016.

[7] Savedoff WD, Levine R, Birdsall N. When will we ever learn?: Improving lives through impact evaluation. Center for Global Development; 2006.

[8] Ravallion M. Evaluating anti-poverty programs. Handbook of development economics. 2007;4:3787–3846.

[9] Jones D, Molitor D, Reif J. What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study. National Bureau of Economic Research; 2018. Available from: http://www.nber.org/papers/w24229.

References II

[10] Beauchamp A, Peeters A, Wolfe R, Turrell G, Harriss LR, Giles GG, et al. Inequalities in cardiovascular disease mortality: the role of behavioural, physiological and social risk factors. J Epidemiol Community Health. 2010 Jun;64(6):542–8.

[11] Thyer BA. Quasi-experimental research designs. Pocket guides to social work research methods. New York: Oxford University Press; 2012.

[12] McCormick D, Hanchate AD, Lasser KE, Manze MG, Lin M, Chu C, et al. Effect of Massachusetts healthcare reform on racial and ethnic disparities in admissions to hospital for ambulatory care sensitive conditions: retrospective analysis of hospital episode statistics. BMJ. 2015;350:h1480.

[13] Glymour MM, Kawachi I, Jencks CS, Berkman LF. Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. J Epidemiol Community Health. 2008 Jun;62(6):532–7.

[14] Hamad R, Collin DF, Rehkopf DH. Estimating the Short-term Effects of the Earned Income Tax Credit on Child Health. Am J Epidemiol. 2018 Sep;.

[15] Galama TJ, Lleras-Muney A, van Kippersluis H. The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. National Bureau of Economic Research; 2018. Available from: www.nber.org/papers/w24225.

[16] Freedman DA. Statistical models and shoe leather. Sociological methodology. 1991;21(2):291–313.

[17] Rosenbaum PR. Observation and experiment: an introduction to causal inference. Harvard University Press; 2017.

[18] Glymour MM. Policies as tools for research and translation in social epidemiology. In: Berkman LF, Kawachi I, Glymour MM, editors. Social epidemiology. Oxford University Press; 2014. p. 452–77.