To IV or not to IV? That is the question.

Neil M Davies

neil.davies@bristol.ac.uk @nm_davies

MRC IEU at the University of Bristol

Instrumental Variables and Causal Inference Meeting

Copenhagen 16-17th November 2016

MRC Integrative Epidemiology Unit

NHS National Institute for Health Research

Outline

- The problem
  - When will an instrumental variable estimator give a less biased estimate than a multivariable adjusted estimate?

- Potential methods of assessing bias
  - Tables of covariate balance
  - Bias component plots
  - Negative control outcomes
  - Negative control populations

Motivation

- If multivariable adjusted and instrumental variable regression estimates from the same sample are contradictory, which should we recommend?

- Instrumental variables increasing used in the epidemiological literature
- Instrumental variable estimates can be less biased than conventional estimates.
- BUT IV estimates can easily be far more biased than other approaches.

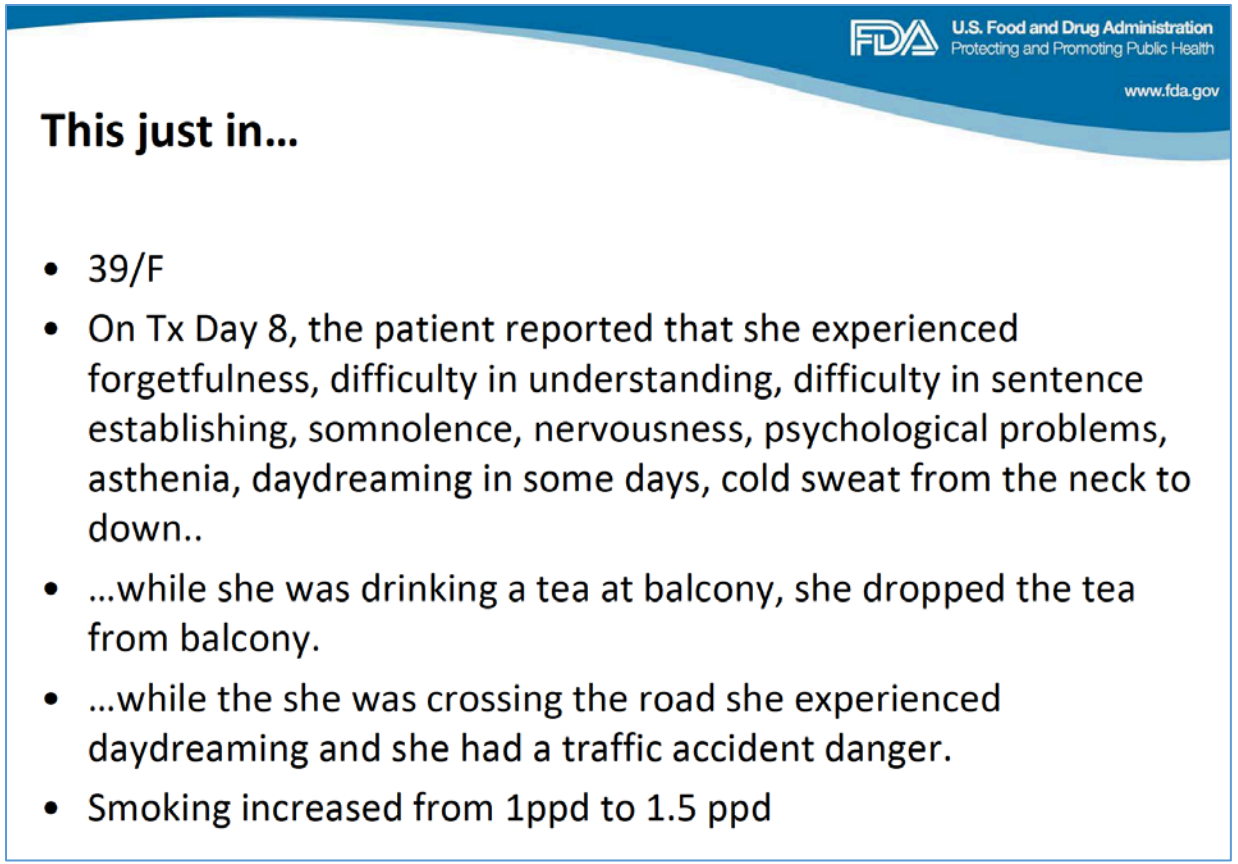
- Need to assess the plausibility of each approaches assumptions.

- Note: for the purpose of this talk I will ignore PS matching because it depends on very similar assumptions as multivariable adjusted regression (i.e. no unmeasured confounding).

Motivating example: does varenicline affect suicide and self-harm?

- Varenicline and nicotine replacement therapy (NRT) are smoking cessation medications
- Anecdotal reports linked varenicline to suicide and self-harm

- What should the FDA and EMA advise patients and clinicians?
- Is a multivariable adjusted or instrumental variable estimate the best estimate of the causal effect of smoking cessation treatments?
- Can we assess the plausibility of assumptions.



Presented on 16th October 2014 to a FDA review of varenicline.
"Celia Jaffe Winchell, Medical Team Leader , Addiction Products Division of Anesthesia, Analgesia, and Addiction Products Center for Drug Evaluation and Research U.S. Food and Drug Administration."

WARNING: SERIOUS NEUROPSYCHIATRIC EVENTS

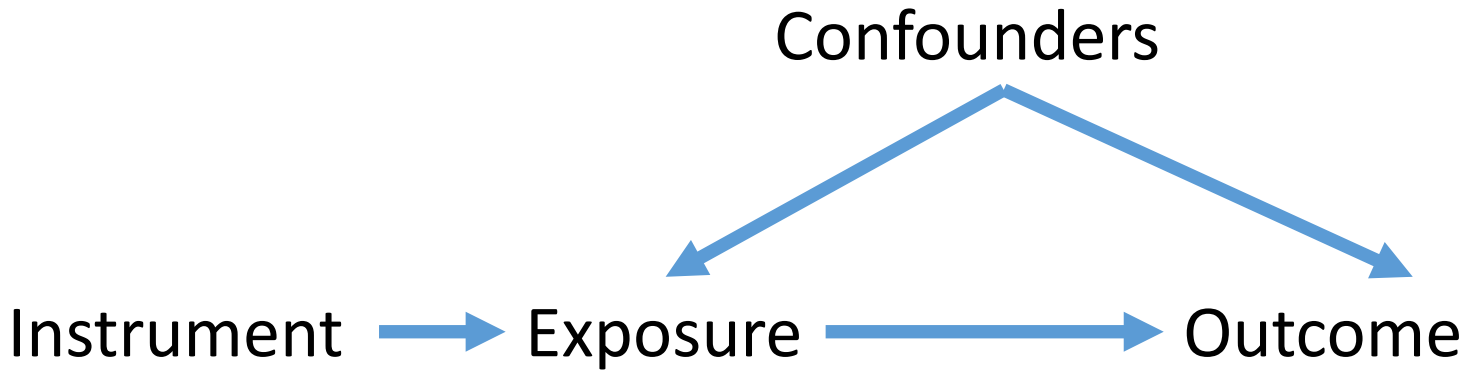
See full prescribing information for complete boxed warning.

- **Serious neuropsychiatric events have been reported in patients taking CHANTIX. (5.1 and 6.2)**
- **Advise patients and caregivers that the patient should stop taking CHANTIX and contact a healthcare provider immediately if agitation, hostility, depressed mood, or changes in behavior or thinking that are not typical for the patient are observed, or if the patient develops suicidal ideation or suicidal behavior while taking CHANTIX or shortly after discontinuing CHANTIX. (5.1 and 6.2)**
- **Weigh the risks of CHANTIX against benefits of its use. CHANTIX has been demonstrated to increase the likelihood of abstinence from smoking for as long as one year compared to treatment with placebo. The health benefits of quitting smoking are immediate and substantial. (5.1 and 6.2)**

Multivariable adjusted regression

- Assumes the exposure is independently assigned conditional on observed covariates (conditional exchangeability)
- Pleiotropy will almost certainly violate this assumption
- Genetic variants which associate with both the exposure and the outcome will be potentially unmeasured confounders.
- Can potentially account for observed differences
- Conditional exchangeability is unverifiable

An instrumental variables DAG



The core IV assumptions

1) Relevance assumption (verifiable)
2) Exclusion restriction (unverifiable, but falsifiable)
3) Independence assumption (unverifiable, but falsifiable)

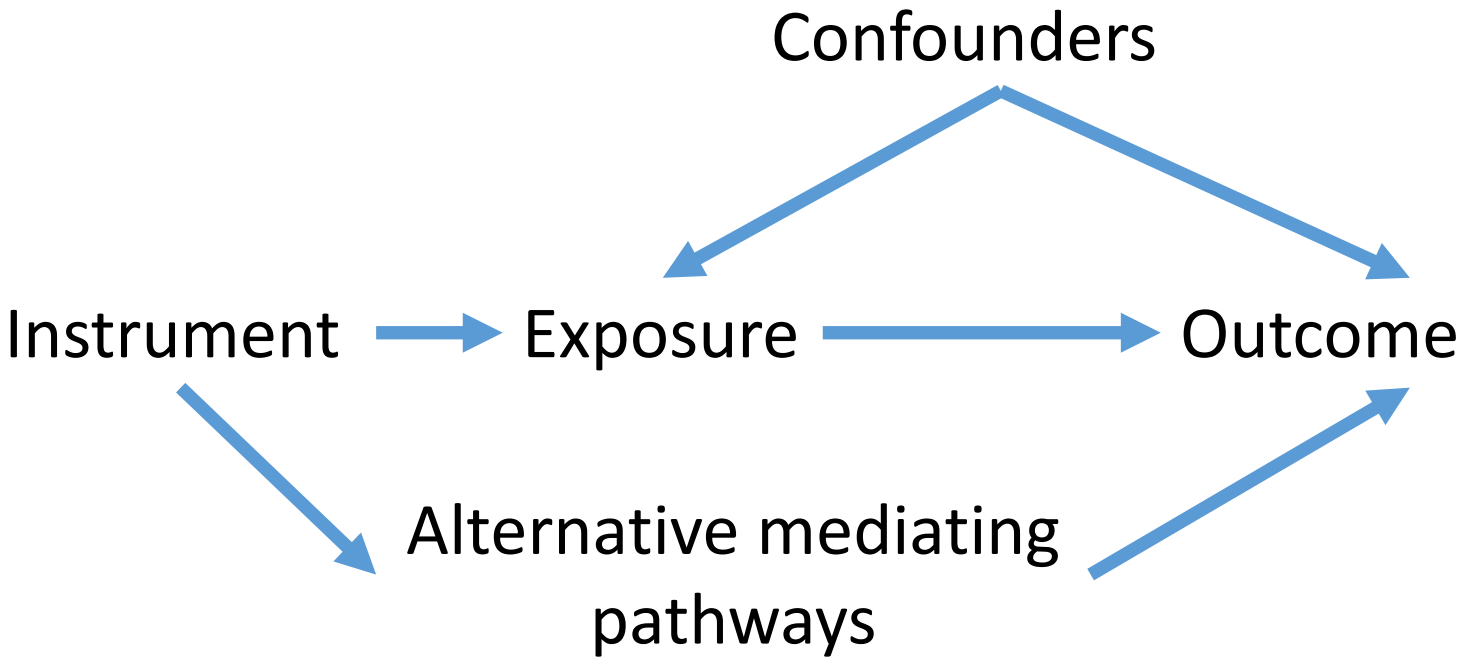
Hernán MA, Robins JM (2016). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming.

Verifying the relevance assumption

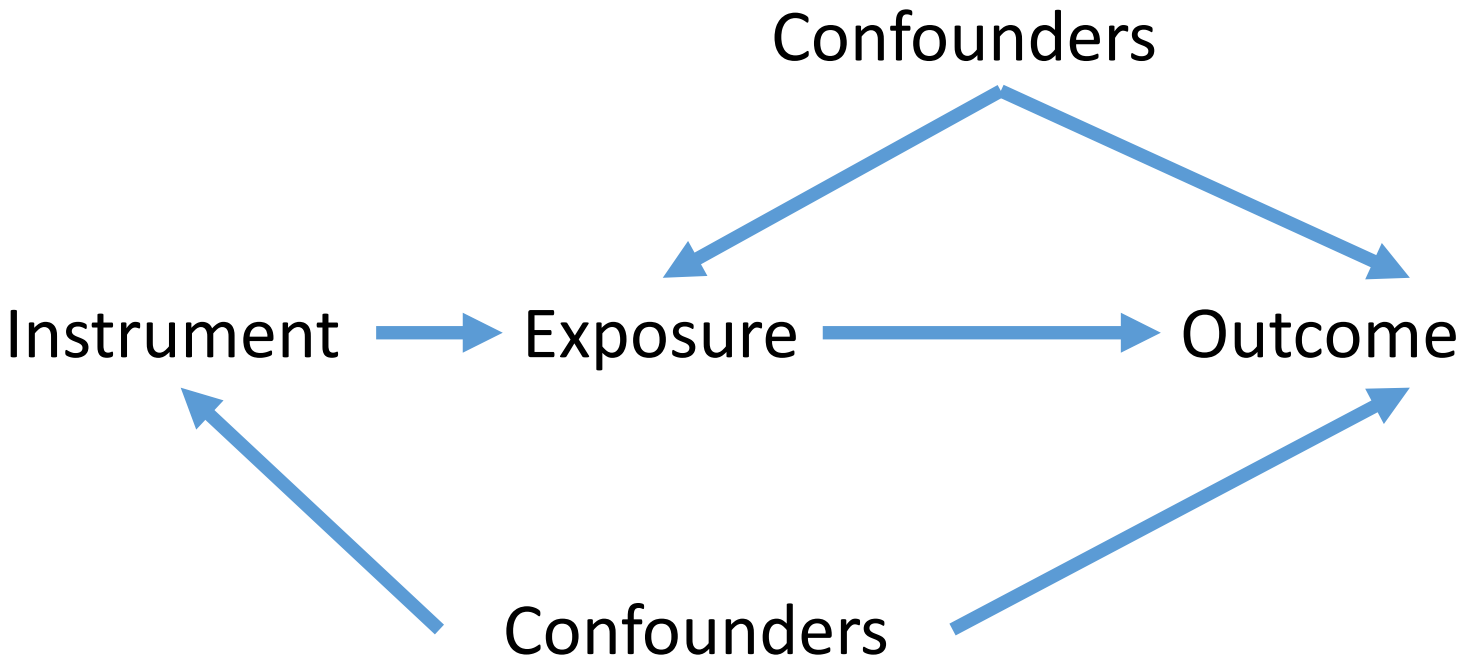
- Estimate the instrument-exposure association
- Common statistics for reporting instrument strength:
  - Risk difference
  - Partial r^2
  - Partial f-statistic
- A weak instrument is trivial to detect
- Many weak instruments slightly more tricky (see Davies et al. 2015)
- A strong instrument does not guarantee sufficient power to detect effects of interest

Davies et al. “The Many Weak Instruments Problem and Mendelian Randomization.” *Statistics in Medicine* 34, no. 3 (2015): 454–68. doi:10.1002/sim.6358.

Falsifying the exclusion restriction



Falsifying the independence assumption



Instrumental variables estimator

- With an outcome Y, and a binary exposure X and instrument Z

$$\alpha_{IV} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[X|Z=1] - E[X|Z=0]}$$

- The relevance assumption relates to the denominator
- Violations of the exclusion restriction and independence assumptions can bias the numerator.

A model of bias

- Consider the standard linear model, where the outcome Y is a function of the binary exposure x, and a single binary confounder C. The causal effect of the exposure is assumed to be equal to the constant α_1.

$$Y(x) = \alpha_0 + \alpha_1 x + \alpha_2 C + \epsilon_x$$

$$bias_{ols} = (E[Y|C = 1, X = x] - E[Y|C = 0, X = x]) \times (E[C|X = 1] - E[C|X = 0])$$

$$bias_{iv} = (E[Y|C = 1, X = x] - E[Y|C = 0, X = x]) \times \frac{E[C|Z = 1] - E[C|Z = 0]}{E[X|Z = 1] - E[X|Z = 0]}$$

Methods for assessing bias: Tests of covariate balance

- Simple test of the exclusion restriction and independence assumption
- How strongly are the covariates associated with the proposed instrument compared to the exposure?

Table 4.—Patient Characteristics by Differential Distance to a Catheterization or Revascularization Hospital*

Characteristic	Differential Distance ≤2.5 Miles (n=102 516)	Differential Distance >2.5 Miles (n=102 505)
Demographic Characteristics		
Female	51.3	49.5
Black	7.1	4.3
Mean age, y (SD)	76.1 (7.3)	76.1 (7.2)
Rural	6.5	52.4
Comorbid Disease Characteristics		
Cancer	1.9	1.9
Pulmonary disease, uncomplicated	10.4	10.9
Dementia	0.99	0.94
Diabetes	18.1	18.0
Renal disease, uncomplicated	2.0	1.9
Cerebrovascular disease	4.8	4.8
Treatments		
Initial admit to catheterization hospital†	34.4	5.0
Initial admit to revascularization hospital†	41.7	10.7
Initial admit to high-volume hospital†	67.1	36.5
Catheterization within 7 d	20.7	11.0
Catheterization within 90 d	26.2	19.5
CABG‡ within 90 d	8.6	6.9
PTCA§ within 90 d	6.4	4.3

McClellan, M, BJ McNeil, and JP Newhouse. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables." *JAMA* 272, no. 11 (1994): 859–66.

Methods for assessing bias: Tests of covariate balance

- Limitations – covariate-instrument associations have a much larger impact on the bias than covariate-exposure associations
- Falsely reassuring instrument will often be less associated than actual exposure
- Potentially underpowered

Methods for assessing bias: Prevalence Difference Ratios

- Brookhart and Schneeweiss (2007)

$$E[X|Z=1] - E[X|Z=0] > \frac{E[C|Z=1] - E[C|Z=0]}{E[C|X=1] - E[C|X=0]}$$

- If the strength of the instrument is greater than the ratio of the difference in the prevalence, then IV less biased.
- This method is not widely used.
- Not graphical, not intuitive

Methods for assessing bias 2: Bias components terms

- Jackson and Swanson (2015) bias components

$$bias_{ols} = (E[Y|C=1, X=x] - E[Y|C=0, X=x]) \times (E[C|X=1] - E[C|X=0])$$

$$bias_{iv} = (E[Y|C=1, X=x] - E[Y|C=0, X=x]) \times \frac{E[C|Z=1] - E[C|Z=0]}{E[X|Z=1] - E[X|Z=0]}$$

- Assess whether:

$$E[C|X=1] - E[C|X=0] > \frac{E[C|Z=1] - E[C|Z=0]}{E[X|Z=1] - E[X|Z=0]}$$

- Recommend plotting bias terms graphically

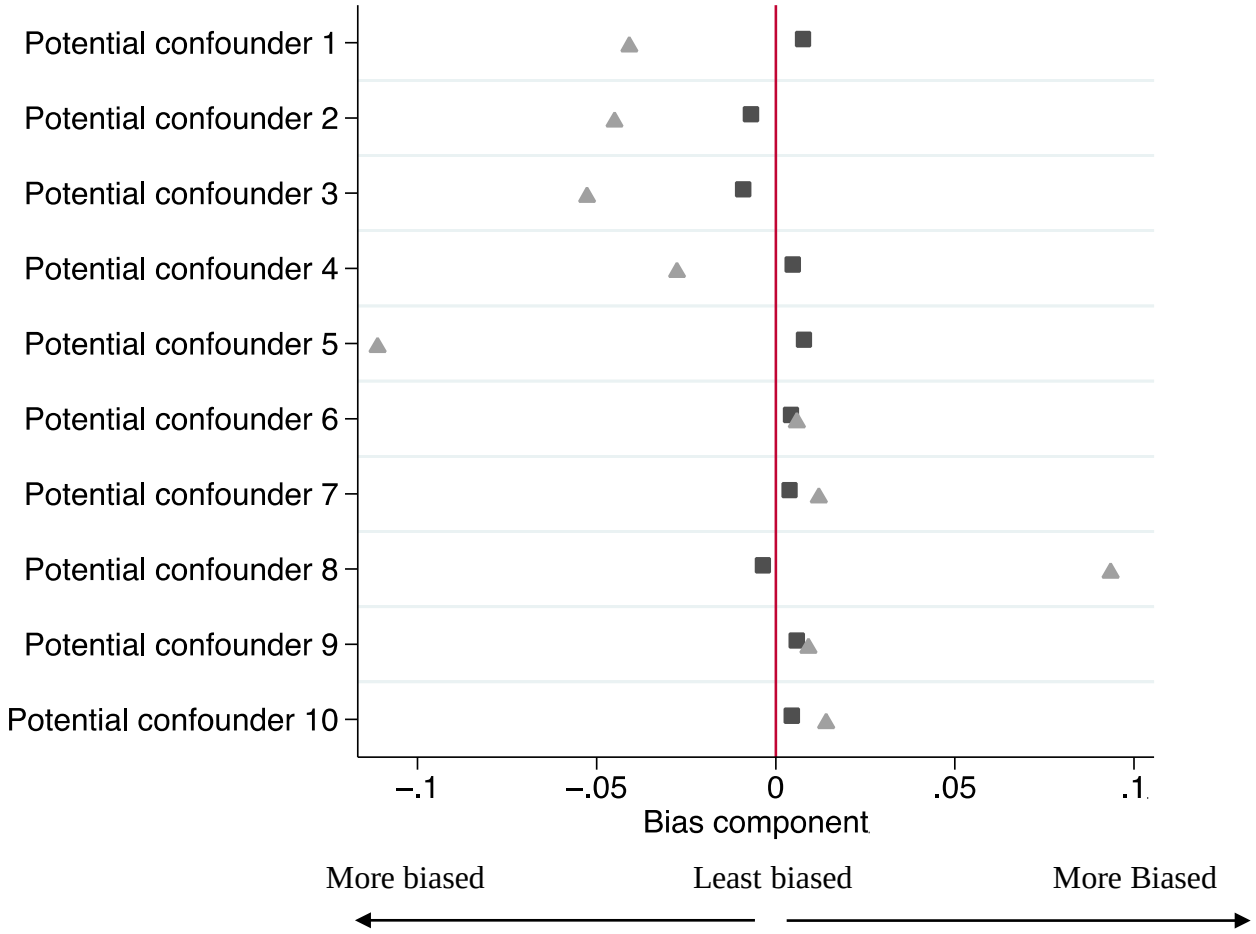
Simulation of bias terms

- Simulated linear model as before
- 10 binary covariates
- A single binary exposure and instrument
- A continuous outcome

- Set the effect of the exposure, $\alpha_1 = 0.5$, and $N = 10{,}000$.

Plots of bias component terms

Bias component plots: difference in patients' diagnoses in the previous year by actual exposure (■) and proposed instrument (▲).

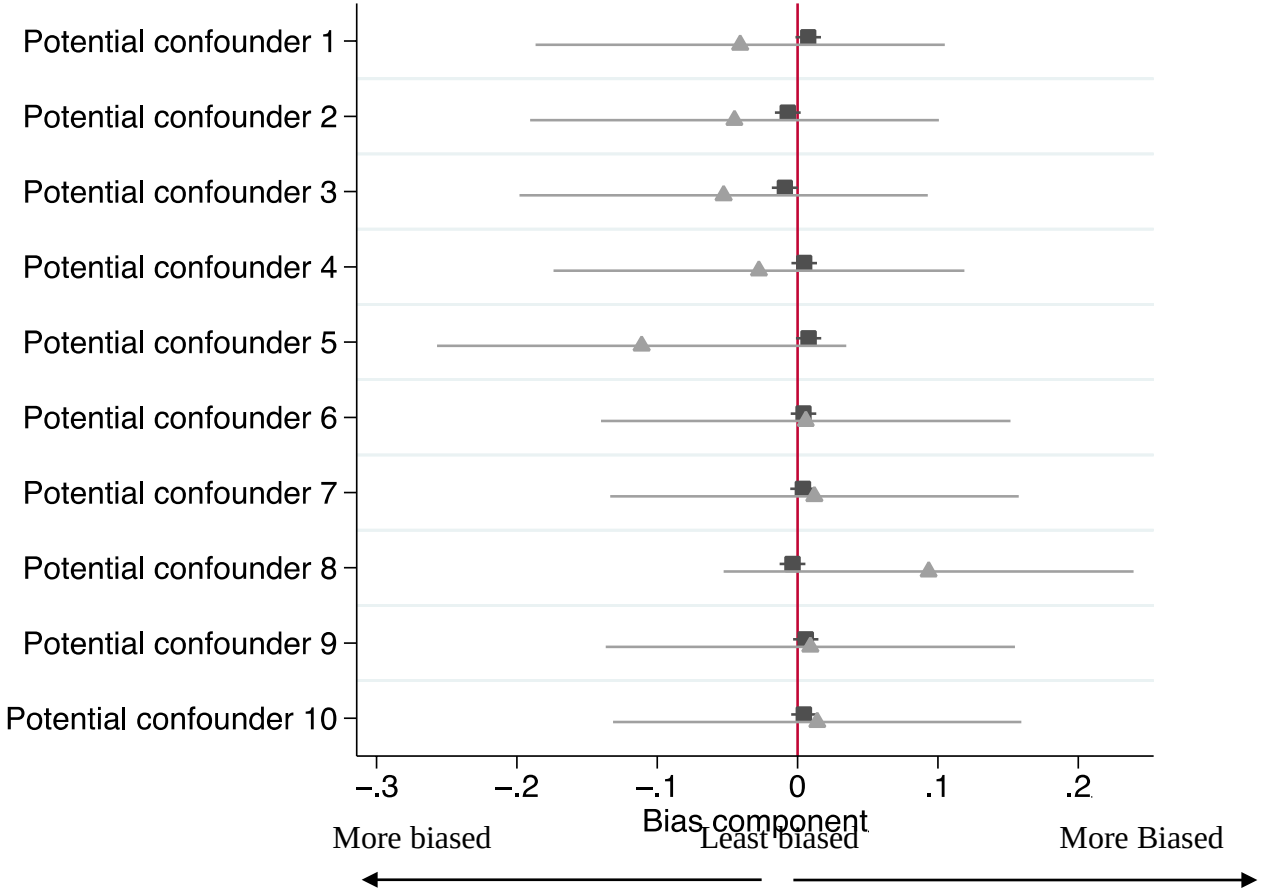


Plots of bias component terms

- Limitations: these plots do not account for estimation error
- All estimates of the bias terms will be estimated with error
- The instrumental variable bias terms are likely to be more variable than the OLS bias terms
- Because the IV bias term will only use a portion of variation in the exposure
- Therefore we can only interpret these plots if they include CIs

Bias component plots: with CIs

Bias component plots: difference in patients' diagnoses in the previous year by actual exposure (■) and proposed instrument (▲).



Negative control outcomes and populations

- **Negative control outcome**: An outcome which the researcher believes should not be affected by the exposure or the proposed instrumental variable.
- **Negative control population**: A population in which the researcher believes the exposure or instrumental variable will not affect or be related to the outcome.

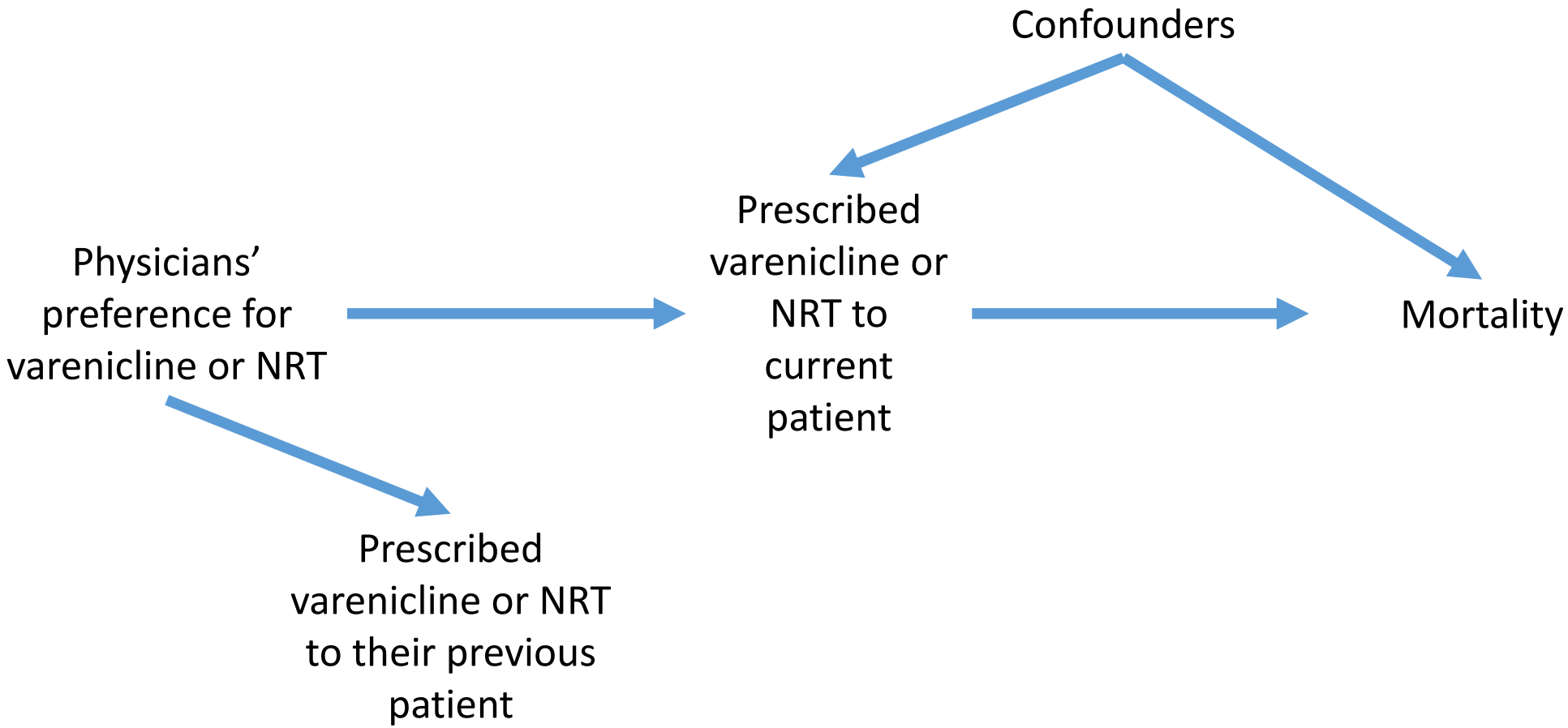
A brief note on selection bias

- These methods can potentially detect selection bias
- Swanson et al. (2015) argue that selection bias can be induced if patients given some treatments are omitted.
- We used the simulation described in the paper to assess whether these bias assessment methods can detect selection bias
- We modified their simulation to have a proxy (measured) confounder which had only a weak correlation with the true confounder (r2=0.01)
- When we restricted the analysis to treated patients, the instrumental variable bias component was detectable and an order of magnitude larger than the linear regression bias component.

An empirical example

- Hypothesis: do smoking cessation treatments affect mortality?
- Data: 280,000 patients from the Clinical Practice Research Datalink
- Prescribed either
  - Nicotine replacement therapy (NRT) (control)
  - Varenicline (drug) (treatment)
- Followed from first prescription (as per Hernan et al. 2008)
- Information on a wide range of baseline diagnoses and treatments

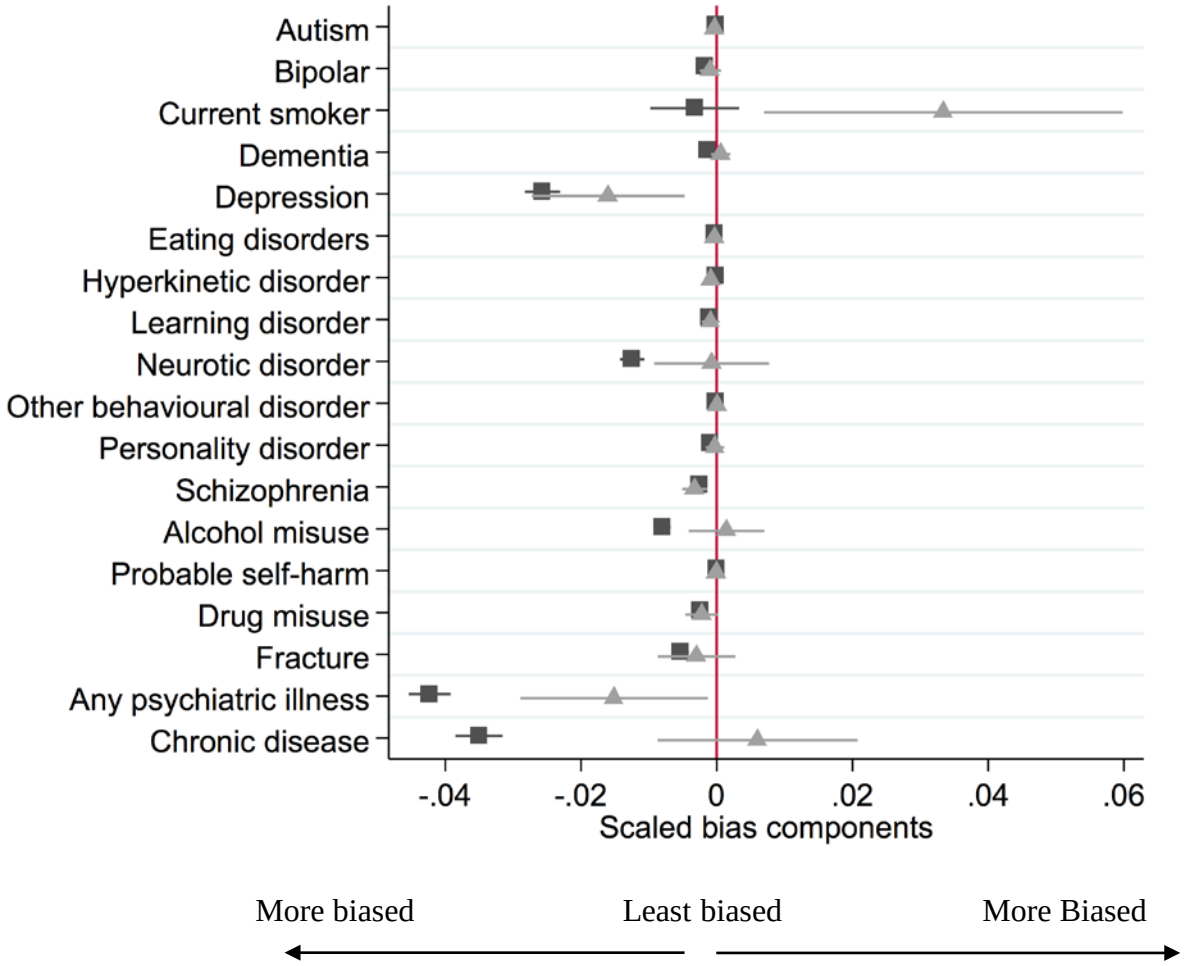
Physicians' prescribing preferences: a potential instrumental variable



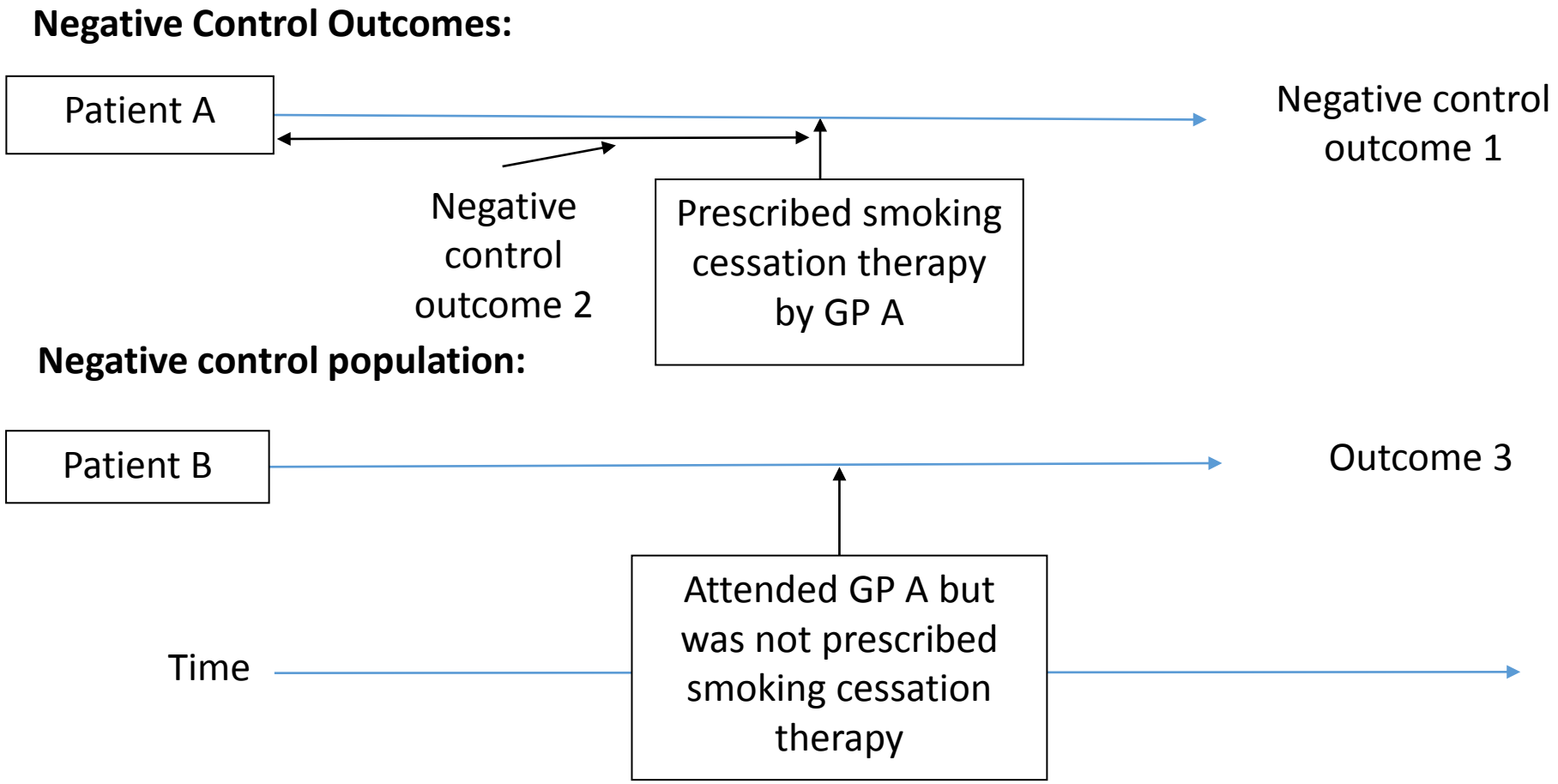
Relevance assumption

- Physicians who prescribed varenicline to their previous patient were 24 percentage points (95%CI: 23, 25) more likely to prescribe varenicline to their subsequent patients than physicians who previously prescribed NRT
- Partial F-statistic=1011.5
- Instruments relevant
- Are the proposed instruments excludable and independent of covariates?

Figure: Bias component plots: difference in patients' diagnoses in the previous year by actual exposure (■) and proposed instrument (▲).

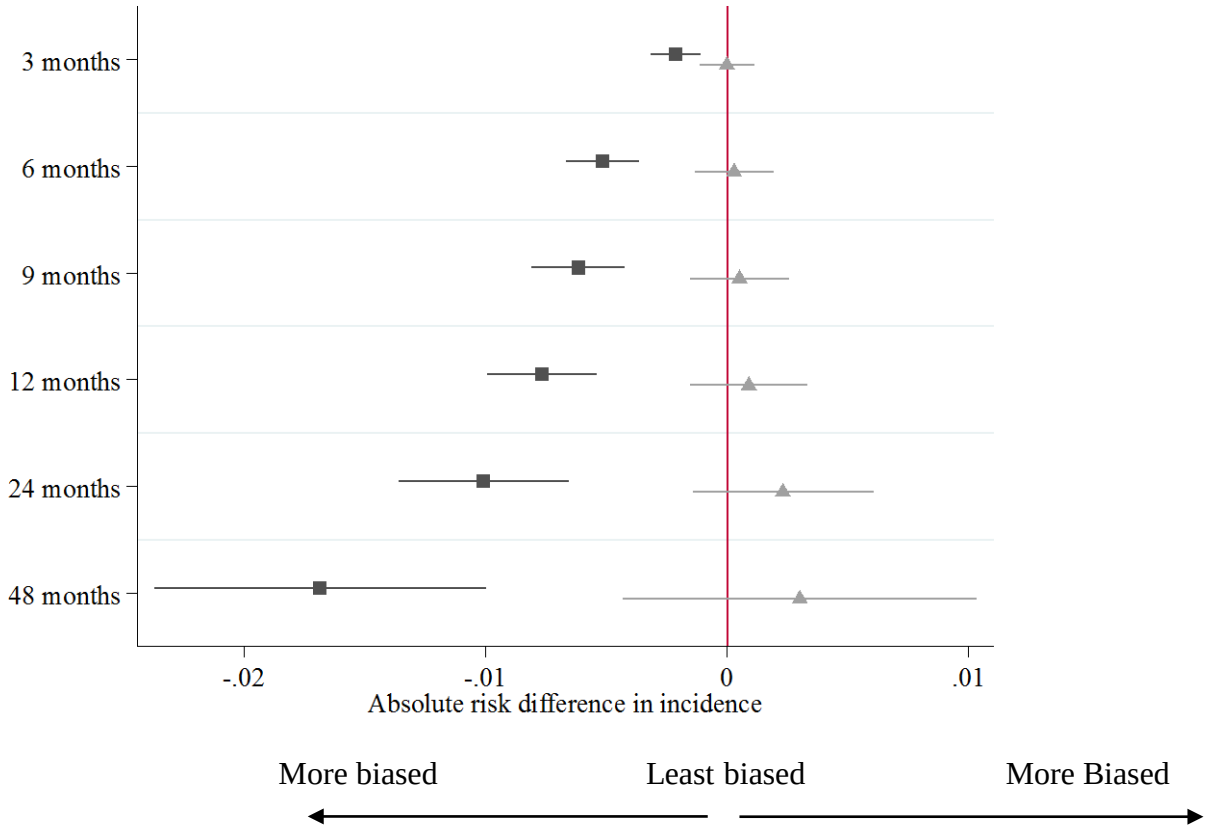


Potential negative control outcomes and negative control populations



Negative control outcomes

Difference in the incidence of urinary tract infections in the four years after smoking cessation treatment for the index patients by actual prescription (■) and the proposed instrument (▲).



Limitations

All these bias assessment methods

- Have limited power
- Assume a homogenous treatment effect
- Assume the observed confounders are indicative of the unobserved confounders

Conclusions

To IV or not to IV?

- The relative plausibility of the IV and multivariable adjusted regression assumptions can be assessed using:
  1. Tests of relevance assumption
  2. Bias component plots
  3. Negative control outcomes
  4. Negative control populations

- The observed data can provide an indication of the relative plausibility of different approaches
- The exclusion restriction and independence assumptions are not verifiable, but they are falsifiable.

Thank you! Questions, comments?

References

- Davies et al. “The Many Weak Instruments Problem and Mendelian Randomization.” *Statistics in Medicine* 34, no. 3 (2015): 454–68. doi:10.1002/sim.6358.
- Jackson, John W., and Sonja A. Swanson. “Toward a Clearer Portrayal of Confounding Bias in Instrumental Variable Applications.” *Epidemiology* 26, no. 4 (July 2015): 498–504. doi:10.1097/EDE.0000000000000287.
- Davies, Neil M. “Commentary: An Even Clearer Portrait of Bias in Observational Studies?” *Epidemiology (Cambridge, Mass.)* 26, no. 4 (July 2015): 505–8. doi:10.1097/EDE.0000000000000302.
- Brookhart, M Alan, Philip S Wang, Daniel H Solomon, and Sebastian Schneeweiss. “Evaluating Short-Term Drug Effects Using a Physician-Specific Prescribing Preference as an Instrumental Variable.” *Epidemiology* 17, no. 3 (May 2006): 268–75. doi:10.1097/01.ede.0000193606.58671.c5.
- Brookhart, M Alan, and Sebastian Schneeweiss. “Preference-Based Instrumental Variable Methods for the Estimation of Treatment Effects: Assessing Validity and Interpreting Results.” *The International Journal of Biostatistics* 3, no. 1 (2007): 14.
- Hernán, Miguel A, and JM Robins. “Instruments for Causal Inference: An Epidemiologist’s Dream?” *Epidemiology* 17, no. 4 (2006): 360–72. doi:10.1097/01.ede.0000222409.00878.37.
- Hernán MA, Robins JM (2016). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming.
- Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;**29**:722–9.
- Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc* 1996;**91**:444–55.
- Jackson, John W. “Diagnostics for Confounding of Time-Varying and Other Joint Exposures:” *Epidemiology* 27, no. 6 (November 2016): 859–69. doi:10.1097/EDE.0000000000000547.
- McClellan, M, BJ McNeil, and JP Newhouse. “Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables.” *JAMA* 272, no. 11 (1994): 859–66.
- Hernán, Miguel A, Alvaro Alonso, Roger Logan, Francine Grodstein, Karin B Michels, Walter C Willett, Joann E Manson, and James M Robins. “Observational Studies Analyzed like Randomized Experiments: An Application to Postmenopausal Hormone Therapy and Coronary Heart Disease.” *Epidemiology (Cambridge, Mass.)* 19, no. 6 (November 2008): 766–79. doi:10.1097/EDE.0b013e3181875e61.
- Swanson, S. A., J. M. Robins, M. Miller, and M. A. Hernan. “Selecting on Treatment: A Pervasive Form of Bias in Instrumental Variable Analyses.” *American Journal of Epidemiology* 181, no. 3 (February 1, 2015): 191–97. doi:10.1093/aje/kwu284.

Acknowledgements







Our funders:
This project was funded by the National Institute for Health Research HTA programme (project number 14/49/94), and the MRC IEU unit (MC_UU_12013/6,MC_UU_12013/9, MR/N01006X/1).

Department of Health Disclaimer:
The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HTA programme, NIHR, NHS or the Department of Health.

MRC Integrative Epidemiology Unit

NHS National Institute for Health Research