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
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.”

This just in...

• 39/F
• On Tx Day 8, the patient reported that she experienced forgetfulness, difficulty in understanding, difficulty in sentence establishing, somnolence, nervousness, psychological problems, asthenia, daydreaming in some days, cold sweat from the neck to down...
• ...while she was drinking a tea at balcony, she dropped the tea from balcony.
• ...while the she was crossing the road she experienced daydreaming and she had a traffic accident danger.
• Smoking increased from 1ppd to 1.5 ppd
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)

Verifying the relevance assumption

• Estimate the instrument-exposure association
• Common statistics for reporting instrument strength:
  • Risk difference
  • Partial r²
  • 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

Falsifying the exclusion restriction

- Instrument
- Exposure
- Outcome
- Confounders
- Alternative mediating pathways
Falsifying the independence assumption

Instrument → Exposure → Outcome

Confounders

Confounders
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 $\alpha_1$.

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

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

\[ bias_{iv} = (E[Y|C = 1, X = x] - E[Y|C = 0, X = x]) \times \frac{E[C|Z = 1] - E[C|Z = 0]}{E[X|Z = 1] - E[X|Z = 0]} \]
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>
<thead>
<tr>
<th>Characteristic</th>
<th>Differential Distance ≤2.5 Miles</th>
<th>Differential Distance &gt;2.5 Miles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=102,516)</td>
<td>(n=102,505)</td>
</tr>
<tr>
<td>Female</td>
<td>51.3</td>
<td>49.5</td>
</tr>
<tr>
<td>Black</td>
<td>7.1</td>
<td>4.3</td>
</tr>
<tr>
<td>Mean age, y (SD)</td>
<td>76.1 (7.3)</td>
<td>76.1 (7.2)</td>
</tr>
<tr>
<td>Rural</td>
<td>6.5</td>
<td>52.4</td>
</tr>
<tr>
<td>Cancer</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Pulmonary disease, uncomplicated</td>
<td>10.4</td>
<td>10.9</td>
</tr>
<tr>
<td>Dementia</td>
<td>0.99</td>
<td>0.94</td>
</tr>
<tr>
<td>Diabetes</td>
<td>18.1</td>
<td>18.0</td>
</tr>
<tr>
<td>Renal disease, uncomplicated</td>
<td>2.0</td>
<td>1.9</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Initial admit to catheterization hospital†</td>
<td>34.4</td>
<td>5.0</td>
</tr>
<tr>
<td>Initial admit to revascularization hospital†</td>
<td>41.7</td>
<td>10.7</td>
</tr>
<tr>
<td>Initial admit to high-volume hospital†</td>
<td>67.1</td>
<td>36.5</td>
</tr>
<tr>
<td>Catheterization within 7 d</td>
<td>20.7</td>
<td>11.0</td>
</tr>
<tr>
<td>Catheterization within 90 d</td>
<td>26.2</td>
<td>19.5</td>
</tr>
<tr>
<td>CABG‡ within 90 d</td>
<td>8.6</td>
<td>6.9</td>
</tr>
<tr>
<td>PTCA§ within 90 d</td>
<td>6.4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

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)

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

- 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: 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 (r^2=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

Physicians’ preference for varenicline or NRT

Prescribed varenicline or NRT to current patient

Confounders

Prescribed varenicline or NRT to their previous patient

Mortality
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:

- Patient A
  - Negative control outcome 2
  - Prescribed smoking cessation therapy by GP A

Negative control population:

- Patient B
  - Attended GP A but was not prescribed smoking cessation therapy
  - Outcome 3

Time
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?


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.