ISSUE FIVE October 2014

Evidence Notes

Bridge Medical
Evidence that matters

Methods to Minimise Bias from Observational Study Data – Part 2 *Handling Unknown or Unmeasured Confounders*

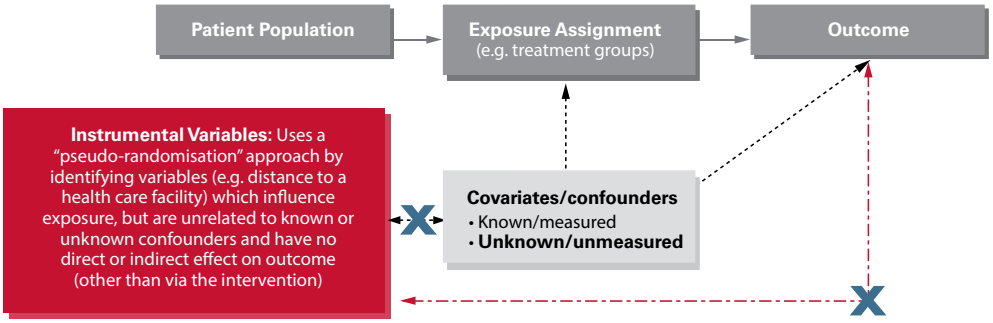
Welcome to this new edition of Evidence Notes. Our aim with these newsletters is to write short, informative articles on a range of topics in the evidence space. Handling bias and confounding is a well known challenge. In our last edition we discussed the application of propensity score matching to real world data sets where confounders are known. In this second article we will describe the use of **instrumental variables** as a method to address bias associated with unknown or unmeasured confounders in observational research. We describe this approach without complex statistics – the aim of this piece is simply to give our readers some idea about the circumstances in which the technique might be applied, and some of the key drawbacks. Our next two editions will be policy focused rather than data focused. The first will provide an update on "where are we" with regards to adaptive study designs and regulatory approval, and the next edition will provide an update on "where are we" with adaptive licensing.

In situations where RCTs are not feasible, ethical or practical and where unknown or unmeasured confounders are likely, understanding and minimising biases relating to treatment assignment is of key importance. Thus, it is important to have a means by which unmeasured confounders are addressed and instrumental variable (IV) methodologies offer one way to do this.

The history of its use dates back to the 1920's, although the inventor is unknown. Since then, IVs have become commonly used in economics and are being increasingly applied in epidemiology. The first study using IV-methods in prescription drug research was an oncology study (Earle et al 2001). However, despite their appeal, valid IVs are not common. They are rarely used in medical research and, where applied, they are most often used in retrospective administrative database studies where the presence of unmeasured confounders is likely to be high.

The aim of using an IV is to eliminate the effects of confounders *without needing to measure them*. Whereas an RCT does this by randomly assigning patients to treatment, the IV approach applies randomness by

Figure 1 – Graphical Illustration of a Statistical Approach Used in Observational Research to Handle Unknown Confounding



using a characteristic of the world or environment which has no effect on outcome but is associated with a subject being more likely to receive a particular exposure/treatment.

Thus, in essence, it is a "pseudo-randomisation" approach which balances measured and unmeasured covariates between intervention groups. Randomisation is an example of an "ideal" IV in which randomisation determines the treatment assignment and only influences outcome via treatment.

So, what constitutes a good IV? Most importantly it must meet 3 conditions. The IV must:

- Influence exposure (e.g. treatment)
- Be unrelated to confounders whether observed (e.g. patient characteristics) or unobserved/unknown; i.e. it is a factor that is "effectively" randomly assigned
- Have no direct or indirect effect on outcome (other than via the intervention)

One example of an IV is "distance to a healthcare facility". If, for example, one wishes to study whether an invasive approach (catheterisation) to the management of acute myocardial infarction (AMI) has mortality benefits vs a more conservative approach it may be ethically difficult to perform an RCT because the invasive approach

has been adopted in many regions. Instead, patients can be grouped using an IV of "distance to a hospital with catheterisation facilities" (and hence probability of receiving a particular treatment). It is irrespective of health status, is unrelated to any known or unknown confounders and is not likely to affect outcome other than via the treatment received. The most commonly used IVs include the following and case studies for each are provided in Table 1:

- Distance to a specific healthcare facility
- Number of local hospitals
- Geographic region
- Physician preference

In selecting an IV it is important to carefully assess its strength and weaknesses for the specific study being undertaken. For example, physician preference has been criticised by some because in some circumstances preference may be related to patient characteristics e.g. physicians with a higher prescription rate for a drug may tend to see more severely ill patients and this may confound the choice of IV. Also, geographic region or distance to a healthcare facility offering a specific treatment may be compromised by other concomitant treatments associated with the IV. Thus, the outcome of the study may be related to more than just the particular treatment under study.

A graphic illustration of IV is provided in Figure 1.

Analysis approaches vary but are based on the concept that the probability of receiving intervention is a function of the IV and other covariates. The basic approach can be broadly summarised as follows:

- Build a model using known confounders and fit it to data representing actual treatment received
- Look at the difference between model prediction and actual treatment. If the propensity for one treatment is higher than expected, an alternative confounder is likely
- Find a variable (instrumental variable) that is correlated with the error term i.e. the difference between model prediction and actual.

Because IV is most suited to situations where unknown confounding is likely, it is no surprise that most of the literature is based on retrospective studies of administrative databases (see Table 1). Use of standard analytical methods may be more appropriate where the distribution of unmeasured confounders are likely to be similar such as when prospectively studying interventions with similar clinical indications and risk (e.g. typical vs atypical antipsychotics in schizophrenia). In contrast, IV methods may be better suited to observational studies of patients selected for invasive or surgical procedures as they are more likely to differ in unknown or unmeasurable ways from patients who are not. For a guide on the use of IV methods see Swanson et al 2013.

Though not systematically studied, IV methods are generally comparable with RCT findings (see Table 1). Some believe an IV approach may produce less biased estimates than multivariable regression and propensity methods in retrospective database research. For example, a retrospective database study of invasive vs conservative treatment for AMI showed that IV methods (regional catheterization rate) produced survival benefits that were similar in magnitude to those observed in RCTs and much lower than observed

Table 1: Administrative Database Case Studies

Study Objective	Reason for use of IV	IV	Rationale for Instrument	Outcome
Examine effectiveness of admission to long-term acute care hospitals (LTAC) vs acute care ICU for elderly patients after severe chronic illness (Kahn et al 2013)	To account for selection bias (LTACs select patients on the basis of degree of sickness for LT care) and unmeasured confounding	Distance to nearest LTAC and the number of LTAC beds in local area	1. Admission to LTACs is more likely the nearer or more numerous they are 2. The proximity or number of LTACs are not likely to affect patients outcome other than via use of LTAC	IV analysis showed no difference in survival after admission to LTAC. In contrast, unadjusted & adjusted data (multivariate regression) showed shorter survival on LTAC admission
Examine effectiveness of chemotherapy in elderly patients with stage IV non-small-cell lung cancer (NSCLC) (Earle et al 2001)	To account for unknown prognostic features (confounders)	Geographic variation in chemotherapy use	1. Data shows significant regional variation in use of palliative chemotherapy for advanced lung cancer 2. Effectiveness of chemotherapy not region dependent 3. Prognostic features of patients in high and low use areas are similar	The IV estimate indicated and improvement in survival with chemotherapy consistent with RCT data in younger highly selected subjects
Examine risk of gastrointestinal (GI) complications and myocardial infarction (MI) in first-time prescriptions of COX-2 inhibitors (COX-2s) or NSAIDs in the elderly (Davies et al 2013*; Schneeweiss et al 2006)	To account for important risk factors for cardiac events & GI toxicity e.g. use of aspirin, BMI, smoking, alcohol use and residual confounding via unobserved factors	Physician's previous prescriptions as a surrogate for physician preference	1. Data shows that prescribing of specific NSAID in elderly is driven more strongly by physician preference than by recorded patient characteristics	IV analysis suggested that COX-2s reduce risk of GI events consistent with RCTs. In contrast, regression analysis showed little influence of COX-2s on GI protective effects*

using multivariable regression and propensity methods (Stukel et al 2007).

IVs have limitations, however, and a summary of pros and cons is shown in Table 2

In summary, it is not always possible to know or measure all potential confounders and in these circumstances, instrumental variables are a useful alternative to standard methodologies or propensity scoring. This is particularly true in retrospective database studies where such confounding is likely to be high.

Further information on Instrumental Variables can be found in the references provided.

Martin Jones, Aiden Flynn and Paul Gandhi

Corresponding author:
martinjones@bridgemedical.org

Table 2: Summary of pros & cons

Pros	Cons
Addresses unmeasured confounding	Suitable instruments not always available
Useful if IV is valid, sample size is large, and unmeasured confounding expected to be great	Assumptions are either difficult to verify or unverifiable e.g. if IV only weakly associated with intervention or if IV influences outcome by unknown pathways
Assumptions are different from those underlying conventional approaches	Choice of IV is critical as even a small effect on outcome would undermine the analysis
	Interpretation may be limited to a subset of patients whose treatment choices vary with the instrument
	Very large sample sizes required
	Less precise (i.e. have more variability) than conventional analyses
	Not appropriate in studies with rare outcomes

References

1. Alemayehu D, Alvir JM, Jones B et al. Statistical issues with the analysis of nonrandomized studies in comparative effectiveness research. J Manag Care Pharm. 2011 Nov-Dec;17(9 Suppl A):S22-6. http://www.ncbi.nlm.nih.gov/pubmed/22074671
2. Armstrong K. Methods in comparative effectiveness research. J Clin Oncol. 2012 Dec 1;30(34):4208-14. http://www.ncbi.nlm.nih.gov/pubmed/23071240
3. Beck CA, Penrod J, Gyorkos TW et al. Does aggressive care following acute myocardial infarction reduce mortality? Analysis with instrumental variables to compare effectiveness in Canadian and United States patient populations. Health Serv Res. 2003 Dec;38(6 Pt 1):1423-40. http://www.ncbi.nlm.nih.gov/pubmed/14727781
4. Brookhart MA, Wang PS, Solomon DH et al. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiology. 2006 May;17(3):268-75. http://www.ncbi.nlm.nih.gov/pubmed/16617275
5. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010 Jun;19(6):537-54. http://www.ncbi.nlm.nih.gov/pubmed/20354968
6. Chen Y, Briesacher BA. Use of instrumental variable in prescription drug research with observational data: a systematic review. J Clin Epidemiol. 2011 Jun;64(6):687-700. http://www.ncbi.nlm.nih.gov/pubmed/21163621
7. Crown WH. There's a reason they call them dummy variables: a note on the use of structural equation techniques in comparative effectiveness research. Pharmacoeconomics. 2010;28(10):947-55. http://www.ncbi.nlm.nih.gov/pubmed?term=2010%5Bpdat%5D+AND+Crown+WH%5Bfirst+author%5D&cmd=detailssearch
8. Davies NM, Smith GD, Windmeijer F et al. Issues in the reporting and conduct of instrumental variable studies: a systematic review. Epidemiology. 2013 May;24(3):363-9. http://www.ncbi.nlm.nih.gov/pubmed/23532055
9. Earle CC, Tsai JS, Gelber RD et al. Effectiveness of chemotherapy for advanced lung cancer in the elderly: instrumental variable and propensity analysis. J Clin Oncol. 2001 Feb 15;19(4):1064-70. http://www.ncbi.nlm.nih.gov/pubmed/11181670
10. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology. 2006 Jul;17(4):360-72. http://www.ncbi.nlm.nih.gov/pubmed/16755261Erratum in Epidemiology. 2014 Jan;25(1):164.
11. Iwashyna TJ, Kennedy EH. Instrumental variable analyses. Exploiting natural randomness to understand causal mechanisms. Ann Am Thorac Soc. 2013 Jun;10(3):255-60. http://www.ncbi.nlm.nih.gov/pubmed/23802827
12. Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report--Part III. Value Health. 2009 Nov-Dec;12(8):1062-73. http://www.ncbi.nlm.nih.gov/pubmed/19793071
13. Kahn JM, Werner RM, David G et al. Effectiveness of long-term acute care hospitalization in elderly patients with chronic critical illness. Med Care. 2013 Jan;51(1):4-10. http://www.ncbi.nlm.nih.gov/pubmed/22874500
14. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA. 1994 Sep 21;272(11):859-66. http://www.ncbi.nlm.nih.gov/pubmed/8078163
15. Normand SL. Some old and some new statistical tools for outcomes research. Circulation. 2008 Aug 19;118(8):872-84. http://www.ncbi.nlm.nih.gov/pubmed/18711024
16. Pilote L, Beck CA, Eisenberg MJ et al. Comparing invasive and noninvasive management strategies for acute myocardial infarction using administrative databases. Am Heart J. 2008 Jan;155(1):42-8. http://www.ncbi.nlm.nih.gov/pubmed/18082487
17. Schneeweiss S, Solomon DH, Wang PS et al. Simultaneous assessment of short-term gastrointestinal benefits and cardiovascular risks of selective cyclooxygenase 2 inhibitors and nonselective nonsteroidal antiinflammatory drugs: an instrumental variable analysis. Arthritis Rheum. 2006 Nov;54(11):3390-8. http://www.ncbi.nlm.nih.gov/pubmed/17075817
18. Schneeweiss S, Gagne JJ, Glynn RJ et al. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. Clin Pharmacol Ther. 2011 Dec;90(6):777-90. http://www.ncbi.nlm.nih.gov/pubmed/22048230
19. Stukel TA, Fisher ES, Wennberg DE et al. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. JAMA. 2007 Jan 17;297(3):278-85. http://www.ncbi.nlm.nih.gov/pubmed/17227979
20. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). Epidemiology. 2013 May;24(3):370-4. http://www.ncbi.nlm.nih.gov/pubmed/23549180