An Introduction to Instrumental Variables Analysis: Part 1

Derrick A. Bennett

Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

Key Words

Controlled trials • Epidemiological methods • Statistics

Abstract

There are several examples in the medical literature where the associations of treatment effects predicted by observational studies have been refuted by evidence from subsequent large-scale randomised trials. This is because of the fact that non-experimental studies are subject to confounding – and confounding cannot be entirely eliminated even if all known confounders have been measured in the study as there may be unknown confounders. The aim of this 2-part methodological primer is to introduce an emerging methodology for estimating treatment effects using observational data in the absence of good randomised evidence known as the method of instrumental variables.

Copyright © 2010 S. Karger AG, Basel

The randomised controlled trial (RCT) is considered to be the most robust study design for assessing treatment effects and for inferring causality. In many situations however, RCTs are not possible – for financial, ethical or other reasons [1]. In such situations, standard analytical epidemiological studies such as cohort and case-control studies can be used as an alternative source of evidence.

However, due to inherent problems their levels of evidence are considered weaker than those produced by an RCT [2]. One of the main problems of non-experimental studies is that they are subject to confounding – and confounding cannot be entirely eliminated even if all known confounders have been measured in the study as there may be unknown confounders [3]. An advantage of RCTs is that confounders, whether measured or not, are spread evenly between the intervention and control groups due to randomisation. If randomisation is conducted using a constrained randomisation technique (such as minimisation), then the possibility of chance imbalances is reduced, which is particularly important for small trials. As a consequence of this, if an effect of an intervention is observed, then the magnitude of treatment effect can be reliably estimated [4]. But what if it is not possible to conduct an RCT to ascertain an estimate of the treatment effect? Is it possible to use information from observational studies to estimate the treatment effect? This is a problem in all areas of epidemiology but is a particular problem in neuroepidemiology as discussed in a recent issue of this journal [5]. The aim of this 2-part methodological primer is to introduce an emerging methodology for estimating treatment effects using observational data in the absence of good randomised evidence known as the method of instrumental variables (IVs).

Fig. 1. An illustration of an IV analysis. The assumptions of an IV are: (1) the IV must predict treatment but that prediction does not have to be perfect; (2) a valid IV will not directly be related to outcome, except by the direct effect of treatment; (3) a valid IV will not be related to outcome by either measured or unmeasured paths.

What Are the Problems when Using Observational Studies to Estimate Treatment Effects?

There are several examples in the medical literature where the associations of treatment effects predicted by observational studies have been refuted by evidence from subsequent large-scale randomised trials [6]. One example of misleading estimates of treatment effect from observational data was that vitamin C was beneficial for coronary heart disease (CHD) [7]. However, the subsequent randomised evidence indicated that the vitamin C supplements had no effect on CHD risk [8]. The reason proposed for the discrepancy in the observational evidence and the RCT evidence for vitamin C and CHD is that there is confounding between vitamin C levels and other exposures that could affect the risk of CHD [9]. Lawlor et al. [9] noted that there was considerable confounding from factors across the life course that predict an increased risk of CHD in the British Women's Heart and Health Study, in which women with high levels of plasma vitamin C were less likely to be in a manual social class, be a smoker, be obese and drink alcohol on a daily basis [10]. Observational studies of the effects of exposure to treatment are also particularly prone to confounding by indication (or by contra-indication). This is due to the fact that the allocation to treatment in observational studies is not randomised and the indication (or contra-indication) for treatment may be related to the risk of future health outcomes [11]. Confounding by indication can lead to bias in the size and direction of treatment ef-

fects depending on the relationship between the confounder and outcome [12]. Although relatively new to the field epidemiology, IV methods have been used for several decades by econometricians, where an RCT is not possible, and when assumptions of no unmeasured confounding are not justified [13].

What Is an IV?

An IV is a variable that can be thought to mimic the treatment allocation process in a randomised study. If an appropriate and valid instrument can be found, then the effects of measured and unmeasured confounding can be eliminated [14]. A well-known study by Newhouse and McClellan [15] assessed the impact of cardiac catheterization (bypass surgery or angioplasty) on post-myocardial-infarction survival rates by using variation in distance from a patient's home and the nearest hospital with the relevant facilities as an instrument for whether the patient received cardiac catheterization after myocardial infarction. Typically, after a myocardial infarction, a patient is taken as quickly as possible by ambulance to the nearest hospital regardless of whether the hospital has the relevant catheterization facilities. Hence the catheterization capabilities of the hospital nearest the patient's home are a strong instrument as it determines whether the patient gets the procedure and should be uncorrelated with the outcome. A more recent IV analysis by Wang et al. [16] investigated the treatment effect of antipsychotic drugs prescribed for dementia, delirium and affective disorders among other things using an IV analysis. The authors used the prescribing physician's preference for conventional or atypical medications (based on the most recent new prescription of an antipsychotic agent before the first prescription was written) as the instrument. This was also considered to be a strong instrument. A strong instrument is one that is a good predictor of actual treatment, with its predictive effect independent of other measured variables. An estimate of treatment effect that uses an IV that is weakly associated with treatment (i.e. one that does not predict treatment assignment well) may give misleading results [17].

Figure 1 illustrates the assumptions of a valid IV in terms of its relationship to treatment outcome and unmeasured confounders. In an RCT situation, these assumptions are met by the design of the study and the conduct of randomization. In an IV analysis, these assumptions must be checked empirically based on the subject matter knowledge and scientific context [17]. Investing

an effort into finding plausibly valid instruments is worthwhile because the biases in treatment effects from conventional analyses may be large enough to make the ensuing estimates worthless [15, 18].

How to Estimate the Effect of Treatment Using an IV

IV analysis is typically done using a 2-stage least-squares estimation as illustrated below. In order to assess the strength of an instrument, an F test can be used in first-stage regression which predicts treatment as a function of the instrument and covariates. The F test is used to test the hypothesis that α_1 is significantly different from 0 (in equation 1 described below) [19]. An F statistic >10 is often used as a 'rule of thumb' to indicate that an instrument is not weak [20], but this may not be the case if multiple instruments are available. A partial r^2 (the square of the correlation between the instrument and the treatment adjusted for other covariates) can also be used to assess the proportion of the variance explained by the addition of the IV to the regression model. A large partial r^2 is an indication that the instrument makes a large contribution to the prediction of treatment [19].

Description of the Two-Stage Least-Squares Regression

Stage 1: regression: $T_i = \alpha_0 + \alpha_1 Z_i + v_i$ (1)

where T_i = treatment; Z_i = IV; v_i = error term for treatment $\alpha_1 \neq 0$; Z_i and T_i can be either continuous or binary (can also adjust for measured confounders).

Stage 2: regression: $Y_i = \beta_0 + \beta_1 \hat{T}_i + e_i$ (2)

where $\hat{T}_i = \hat{\alpha}_0 + \hat{\alpha}_1 Z_i$; Y_i = outcome; $\hat{T}_i$ = estimated treatment effect; e_i = error term for outcome (can also adjust for measured confounders).

Substituting equation 1 into equation 2:

$$Y_i = \gamma_0 + \gamma_1 Z_i + \varepsilon_i$$

where

$$\gamma_0 = \beta_0 + \eta_1 \hat{\alpha}_0; \gamma_1 = \beta_1 \hat{\alpha}_1; \varepsilon_i = \beta_1 v_i + e_i.$$

In order to estimate the *direct treatment effect* (β_1) of treatment (T_i) on outcome (Y_i), we need to use the information from equations 1 and 2:

$$\beta_1 = \hat{\gamma}_1 / \hat{\alpha}_1.$$

In the first stage, a regression estimate ($\hat{\alpha}_1$) is obtained by regressing treatment (T) on the IV (Z) in equation 1 (can also adjust for relevant measured confound-

ers). In the second stage, the predicted value of the treatment ($\hat{T}$) is used in a regression of the outcome (Y) on treatment ($\hat{T}$) (can once again adjust for relevant measured confounders) to obtain an estimate $\hat{\gamma}_1 = \beta_1 \hat{\alpha}_1$ in equation 2. This 2-stage approach eliminates the bias that would have occurred in a conventional regression of outcome on actual treatment received using our observational data [18]. The estimated direct treatment effect (β_1) is calculated as the ratio of $\hat{\gamma}_1 / \hat{\alpha}_1$. So for example, if our outcome of interest was continuous and our treatment of interest was dichotomous, the estimated direct treatment effect would be the ratio of the difference in mean outcomes between treatment groups and the difference in treatment assignment predicted by the instrument between the two groups [21]. One of the first published examples of a 2-stage least-squares regression approach was that of Permutt and Hebel [22] who investigated the effect of an intervention that encouraged women to stop smoking whilst pregnant on the birth weight of the newborn infant. The IV was encouraging expectant mothers to stop smoking. The first-stage regression of the difference in the mean number of cigarettes smoked was estimated for those who received the intervention and those who did not (our estimate of 'treatment'). The estimate was found to be -6.4 (i.e. our estimate of $\hat{\alpha}_1$ above). The difference in mean birth weight of newborns between intervention-group mothers and non-intervention-group mothers was 92 g (i.e. our estimate of β_1 above). The direct effect of smoking is estimated to be $92 \text{ g} / -6.4$ (i.e. our estimate $\hat{\gamma}_1 / \hat{\alpha}_1$ above), that is -15 g. This estimate can be interpreted as for every additional cigarette smoked the effect is to reduce the birth weight of the newborn infant by 15 g. Due to the 2-stage nature of the model and the imperfect estimation of the direct treatment effect by the IV, IV analyses will generally produce wider confidence intervals than conventional epidemiological regression analyses [23].

Conclusion

IV should be used carefully as the estimates derived can be more biased than estimates produced by conventional methods if the instruments are not sufficiently strong predictors of treatment [18, 23, 24]. Several assumptions exist that need to be satisfied when an IV analysis is used, and departures from any of these can lead to biased treatment effects [23, 25, 26]. Although situation specific, the general rules for a valid instrument are that

there must be some correlation between the IV and exposure, the relationship between the IV and the exposure is not confounded by other variables, and there is no correlation between the IV and other factors explaining the outcome. In addition, large sample sizes are generally required in order to obtain reasonable precision of the direct treatment effect estimate [25]. As will be discussed in

the second part of this methodological primer, IV analysis has recently gained popularity with genetic epidemiologists where specific genetic markers are used as instruments in order to assess potential treatment effects in the absence of randomised evidence to assess the causal role of emerging risk factors or biomarkers.

References

- Black N: Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215–1218.
- Grimes DA, Schulz KF: Bias and causal associations in observational research. *Lancet* 2002;359:248–252.
- Grimes DA, Schulz KF: An overview of clinical research: the lay of the land. *Lancet* 2002;359:57–61.
- Collins R, MacMahon S: Reliable assessment of the effects of treatment on mortality and major morbidity. I. Clinical trials. *Lancet* 2001;357:373–380.
- Feigin V, Kurtzke JF, Korczyn A, Beghi E, Brown A: Bridging the gap between experimental and nonexperimental neuroepidemiology, and ultimately between neuroepidemiological research and practice: round table discussion at the First International Congress on Clinical Neurology and Epidemiology. *Neuroepidemiology* 2009;33:296–304.
- Tatsioni A, Bonitsis NG, Ioannidis JP: Persistence of contradicted claims in the literature. *JAMA* 2007;298:2517–2526.
- Osganian SK, Stampfer MJ, Rimm E, et al: Vitamin C and risk of coronary heart disease in women. *J Am Coll Cardiol* 2003;42:246–252.
- Heart Protection Study Collaborative Group: MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002;360:23–33.
- Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S: Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet* 2004;363:1724–1727.
- Lawlor DA, Ebrahim S, Kundu D, Bruckdorfer KR, Whincup PH, Smith GD: Vitamin C is not associated with coronary heart disease risk once life course socioeconomic position is taken into account: prospective findings from the British Women's Heart and Health Study. *Heart* 2005;91:1086–1087.
- Signorello LB, McLaughlin JK, Lipworth L, Friis S, Sorensen HT, Blot WJ: Confounding by indication in epidemiologic studies of commonly used analgesics. *Am J Ther* 2002;9:199–205.
- MacMahon S, Collins R: Reliable assessment of the effects of treatment on mortality and major morbidity. II. Observational studies. *Lancet* 2001;357:455–462.
- Angrist J, Krueger A: Instrumental variables and the search for identification: from supply and demand to natural experiments. *J Econ Perspect* 2001;15:69–85.
- Greenland S: An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;29:722–729.
- Newhouse JP, McClellan M: Econometrics in outcomes research: the use of instrumental variables. *Annu Rev Public Health* 1998;19:17–34.
- Wang PS, Schneeweiss S, Avorn J, et al: Risk of death in elderly users of conventional versus atypical antipsychotic medications. *N Engl J Med* 2005;353:2335–2341.
- Klunger OH, Martens EP, Psaty BM, et al: Methods to assess intended effects of drug treatment in observational studies are reviewed. *J Clin Epidemiol* 2004;57:1223–1231.
- Martens EP, Pestman WR, de Boer A, Belitser SV, Klunger OH: Instrumental variables: application and limitations. *Epidemiology* 2006;17:260–267.
- Bound J, Jaeger DA, Baker R: Problems with instrumental variable estimation when the correlation between the instruments and endogenous explanatory variables is weak. *J Am Stat Assoc* 1995;90:443–450.
- Cragg JG, Donald SG: Testing identifiability and specification in instrumental variable models. *Econ Theory* 1993;9:222–240.
- Normand SL: Some old and some new statistical tools for outcomes research. *Circulation* 2008;118:872–884.
- Permutt T, Hebel JR: Simultaneous-equation estimation in a clinical trial of the effect of smoking on birth weight. *Biometrics* 1989;45:619–622.
- Rassen JA, Brookhart MA, Glynn RJ, Mittelman MA, Schneeweiss S: Instrumental variables. I. Instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *J Clin Epidemiol* 2009;62:1226–1232.
- Penrod JD, Goldstein NE, Deb P: When and how to use instrumental variables in palliative care research. *J Palliat Med* 2009;12:471–474.
- Grootendorst P: A review of instrumental variables estimation of treatment effects in the applied health sciences. *Health Serv Outcomes Res Methodol* 2007;7:159–179.
- Rassen JA, Brookhart MA, Glynn RJ, Mittelman MA, Schneeweiss S: Instrumental variables. II. Instrumental variable application – in 25 variations, the physician prescribing preference generally was strong and reduced covariate imbalance. *J Clin Epidemiol* 2009;62:1233–1241.