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 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

---

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

In the figure:

- **Patient Population**
- **Exposure Assignment** (e.g. treatment groups)
- **Outcome**
- **Instrumental Variables:** Uses a “pseudo-randomisation” approach by identifying variables (e.g. distance to a health care facility) which influence exposure, but are unrelated to known or unknown confounders and have no direct or indirect effect on outcome (other than via the intervention)
- **Covariates/confounders**
  - Known/measured
  - Unknown/unmeasured

---

bridgemedical.org
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>
<thead>
<tr>
<th>Study Objective</th>
<th>Reason for use of IV</th>
<th>IV</th>
<th>Rationale for Instrument</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>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)</td>
<td>To account for selection bias (LTACs select patients on the basis of degree of sickness for LT care) and unmeasured confounding</td>
<td>Distance to nearest LTAC and the number of LTAC beds in local area</td>
<td>1. Admission to LTACs is more likely to differ in unknown or unmeasured ways from patients who are not. For a guide on the use of IV methods see Swanson et al. 2013.</td>
<td>IV analysis showed no difference in survival after admission to LTAC. In contrast, unadjusted &amp; adjusted data (multivariate regression) showed shorter survival on LTAC admission</td>
</tr>
<tr>
<td>Examine effectiveness of chemotherapy in elderly patients with stage IV non-small-cell lung cancer (NSCLC) (Earle et al 2001)</td>
<td>To account for unknown prognostic features (confounders)</td>
<td>Geographic variation in chemotherapy use</td>
<td>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</td>
<td>The IV estimate indicated and improvement in survival with chemotherapy consistent with RCT data in younger highly selected subjects</td>
</tr>
<tr>
<td>Examine risk of gastrointestinal (GI) complications and myocardial infarction (MI) in first-time prescriptions of COX-2 inhibitors (CDX-2s) or NSAIDs in the elderly (Davies et al 2013*, Schneeweiss et al 2006)</td>
<td>To account for important risk factors for cardiac events &amp; GI toxicity e.g. use of aspirin, BMI, smoking, alcohol use and residual confounding via unobserved factors</td>
<td>Physician's previous prescriptions as a surrogate for physician preference</td>
<td>1. Data shows that prescribing of specific NSAID in elderly is driven more strongly by physician preference than by recorded patient characteristics</td>
<td>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*</td>
</tr>
</tbody>
</table>

Table 1: Administrative Database Case Studies
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@bridgeme.org |

Table 2: Summary of pros & cons

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addresses unmeasured confounding</td>
<td>Suitable instruments not always available</td>
</tr>
<tr>
<td>Useful if IV is valid, sample size is large, and unmeasured confounding expected to be great</td>
<td>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</td>
</tr>
<tr>
<td>Assumptions are different from those underlying conventional approaches</td>
<td>Choice of IV is critical as even a small effect on outcome would undermine the analysis</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References