PM R 10 (2018) 303-308

www.pmrjournal.org

Statistically Speaking

Instrumental Variables: Uses and Limitations

Kristin L. Sainani, PhD

Introduction

Observational studies are limited by the issue of confounding. For example, light-to-moderate drinkers tend to have lower risk of cardiovascular disease than nondrinkers. This association could be due to a causal effect of alcohol on the body, or it could be due to other attributes of light-to-moderate drinkers, such as high socioeconomic status or the ability to practice moderation across a range of behaviors. Statistical adjustment only imperfectly addresses confounding [1]: Confounders such as socioeconomic status may be crudely measured, resulting in residual confounding; plus, some confounders—such as intangible attributes like the ability to practice moderation—may remain unmeasured.

To address this statistical dilemma, instrumental variable analysis addresses both residual and unmeasured confounding by exploiting "natural experiments." Instrumental variables are widely used in economics and social science research but have only more recently been applied to medical studies [2]. This article reviews instrumental variables, including what they are, how they work, how they are evaluated, and their advantages and limitations. Besides their use in observational studies, instrumental variable analysis also can be applied to the case of randomized trials with noncompliance.

What Is an Instrumental Variable?

An instrumental variable is a naturally occurring phenomenon that imperfectly randomizes people to an exposure or treatment. Also called an instrument, it has to meet 3 conditions: (1) it must be related to the exposure or treatment; (2) it must be unrelated to confounders (at least after adjusting for measured confounders); and (3) it must have no direct effect on the outcome except through its effect on exposure/treatment.

For example, *ALDH2* genotype has been used as an instrument for studying the association between alcohol use and cardiovascular disease [3]. The gene *ALDH2* is involved in alcohol metabolism; people who carry inactive copies of the gene experience facial flushing and other adverse symptoms after drinking. *ALDH2* genotype appears to be an appropriate instrument since it is: (1) strongly related to alcohol consumption; (2) unlikely related to confounders, such as socioeconomic status or one's ability to practice moderation; and (3) unlikely to directly affect cardiovascular disease risk other than through its effect on alcohol intake. This is depicted graphically in Figure 1 [3]. Carrier status can be viewed as a random event that partially determines one's alcohol exposure but is otherwise unrelated to cardiovascular disease risk. Thus, by relating *only* the alcohol exposure that is determined by this gene to cardiovascular disease risk, one is able to estimate an unconfounded effect of alcohol on cardiovascular disease risk.

Another commonly used instrument is distance to specialty care. Proximity may dictate where one receives care during an acute episode; but, presumably, people do not choose their residences based on this factor. In one study, researchers used differential distance to a designated stroke center—the distance from a patient's residence to the nearest stroke center minus the distance from a patient's residence to the nearest hospital of any kind—as an instrument for studying the impact of admissions to a stroke center on mortality in acute stroke patients [4]. Differential distance strongly predicted admissions to a stroke center and appeared unrelated to age or comorbidities. Differential distance was related to race and residence in an urban versus rural setting, which are potential confounders. Differential distance is still a valid instrumental variable if: (1) it is randomly determined within groups defined by these measured confounders (eg, within race groups) and (2) researchers adjust for these measured confounders in their analyses. In this case, the association between differential distance and mortality will still be unconfounded.

Other commonly used instruments include policy changes, physician or institution preference for one

1934-1482/$ - see front matter
https://doi.org/10.1016/j.pmrj.2018.02.002

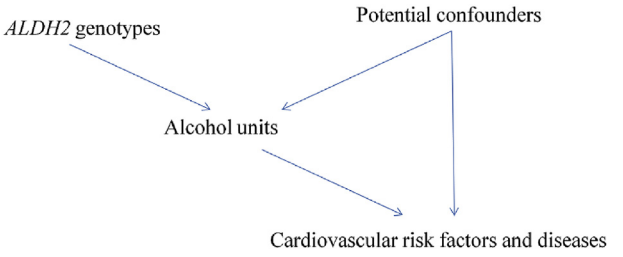

Figure 1. The *ALDH2* genotype appears to be an appropriate instrument as it is affects alcohol intake but is unlikely to affect cardiovascular risk other than through alcohol intake. Reprinted from Au Yeung et al [3].

treatment over another, and prescribing trends over time (see Table 1 [3-8] for examples). Note that multiple instruments may be used in the same analysis to improve precision. Treatment assignment in a randomized trial with noncompliance also can be treated as an instrumental variable. For example, in a randomized trial of integrated care versus usual care for improving the psychosocial health of children with special needs, only 48% of those assigned to the integrated care group complied with treatment (all control patients received usual care) [8]. Treatment assignment in the randomized trial was therefore an imperfect randomizer and was treated as an instrumental variable.

How Are Effects Calculated?

Instrumental variable analysis works by isolating the variation in exposure/treatment that is explained by the instrument and then relating *solely* this variation in exposure/treatment to the outcome. For simple cases, this can be accomplished by calculating the effect of the instrument on the outcome and then rescaling this effect to reflect units of the exposure/treatment rather than those of the instrument (see the In-Depth box for more mathematical details). For example, in the *ALDH2* genotype study, carriers of 2 copies of the inactive *ALDH2* gene ($ALDH2^{-}/^{-}$) had a 1-mm Hg lower diastolic blood pressure on average than noncarriers. (Note: I have simplified the example by ignoring heterozygotes, $ALDH2^{+}/^{-}$.) Individuals with $ALDH2^{-}/^{-}$ also imbibed an average of just 0.09 standard drinks/day versus 0.90 standard drinks/day for noncarriers. Thus, we can estimate the effect of alcohol use on blood pressure as follows:

$$\text{Effect of genotype on blood pressure} = -1 \text{ mm Hg}$$

$$\text{Effect of genotype on alcohol consumption} = -0.81 \text{ standard drinks/day}$$

$$\text{Effect of alcohol consumption on blood pressure} = \frac{-1 \text{ mm Hg}}{-0.81 \text{ standard drinks/day}} = 1.2 \text{ mm Hg per 1 standard drink/day}$$

Complier: someone whose treatment/exposure level depends on the instrument. For example, in the case of a randomized trial with noncompliance, a complier is a person who would have taken the treatment if assigned to the treatment group and would have taken the control if assigned to the control group.

Noncomplier: someone whose treatment/exposure level does not depend on the instrument. For example, in the case of a randomized trial with noncompliance, a noncomplier is someone who would make the same treatment decision irrespective of how they were randomized.

Importantly, we cannot always tell who is a complier and who is not (for example, someone who would never take the treatment and is assigned to the control group is indistinguishable from a complier assigned to the control group). But we can still estimate the complier effect.

Importantly, because the estimate is based only on a portion of the variation in alcohol consumption (that dictated by genotype), the effective sample size is reduced and the results may not be generalizable to everyone. To see this, imagine that there are only 2 types of people: **Compliers** are those whose drinking behavior is completely determined by genotype; and **noncompliers** are people whose drinking behavior is completely unrelated to genotype—for example, people who would never drink regardless of genotype due to religious reasons. Instrumental variable analysis effectively omits noncompliers from the analysis (see In-Depth box for more details). This means that: (1) the results may not be generalizable to noncompliers, and (2) our effective sample size is smaller than the total N, resulting in a loss of precision. For example, the 95% confidence interval for the effect of alcohol use on diastolic blood pressure was 0.23 to 2.07 from the instrumental variable analysis, whereas it was 0.35 to 0.64 from a direct regression of blood pressure on alcohol intake.

For randomized trials with noncompliance, the effect of the instrument on the outcome is simply the intention-to-treat estimate; and the effect of the instrument on treatment is just the difference in the proportion of people receiving the treatment in the randomization groups. For example, in the integrated