Original article

Power calculator for instrumental variable analysis in pharmacoepidemiology

Venexia M Walker,[1,2]* Neil M Davies,[1,2] Frank Windmeijer,[3,2]
Stephen Burgess[2,4,5] and Richard M Martin[1,2]

[1]School of Social and Community Medicine, [2]MRC University of Bristol Integrative Epidemiology Unit,
[3]Department of Economics, University of Bristol, Bristol, UK, [4]Department of Public Health and Primary
Care, University of Cambridge, Cambridge, UK and [5]MRC Biostatistics Unit, Cambridge, UK

*Corresponding author. School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove,
Bristol BS8 2BN, UK. E-mail: venexia.walker@bristol.ac.uk

Abstract

Background: Instrumental variable analysis, for example with physicians' prescribing preferences as an instrument for medications issued in primary care, is an increasingly popular method in the field of pharmacoepidemiology. Existing power calculators for studies using instrumental variable analysis, such as Mendelian randomization power calculators, do not allow for the structure of research questions in this field. This is because the analysis in pharmacoepidemiology will typically have stronger instruments and detect larger causal effects than in other fields. Consequently, there is a need for dedicated power calculators for pharmacoepidemiological research.
Methods and Results: The formula for calculating the power of a study using instrumental variable analysis in the context of pharmacoepidemiology is derived before being validated by a simulation study. The formula is applicable for studies using a single binary instrument to analyse the causal effect of a binary exposure on a continuous outcome. An online calculator, as well as packages in both R and Stata, are provided for the implementation of the formula by others.
Conclusions: The statistical power of instrumental variable analysis in pharmacoepidemiological studies to detect a clinically meaningful treatment effect is an important consideration. Research questions in this field have distinct structures that must be accounted for when calculating power. The formula presented differs from existing instrumental variable power formulae due to its parametrization, which is designed specifically for ease of use by pharmacoepidemiologists.

Key words: Pharmacoepidemiology, instrumental variable, power, binary exposure, continuous outcome, prescribing preference

---

Key Messages

- Research questions using instrumental variable analysis in pharmacoepidemiology have distinct structures that have previously not been catered for by instrumental variable analysis power calculators.
- Power can be calculated for studies using a single binary instrument to analyse the causal effect of a binary exposure on a continuous outcome in the context of pharmacoepidemiology using the presented formula, an online power calculator or packages available for use in both R and Stata.
- The use of this power calculator will allow investigators to determine whether a pharmacoepidemiology study is likely to detect clinically meaningful treatment effects before the study's commencement.

Introduction

Pharmacoepidemiological studies risk irrelevance if they are insufficiently powered to detect clinically meaningful treatment effects. Before starting a study, the statistical power to calculate a given treatment effect can be calculated. This type of calculation is becoming increasingly important for grant and data request applications, which look to value the contribution of such studies.

The number of pharmacoepidemiology studies using instrumental variable analysis, for example with physicians' prescribing preferences as an instrument for exposure, continues to grow.[1–6] This is partly because instrumental variable analyses have the potential to overcome some of the issues associated with conventional statistical approaches, such as residual confounding and reverse causation. As the demand to provide power calculations to support applications increases, there is a more pressing need to be able to provide power calculations for this method.

There are power calculators for instrumental variable analysis in other settings, such as Mendelian randomization, which uses germline genetic variants as proxies for exposures in disease-related research.[7,8] However, pharmacoepidemiological research questions have distinct structures that are not sufficiently catered for by these existing calculators. Unlike Mendelian randomization studies, which often use a case-control study design, pharmacoepidemiology studies typically use a cohort study design. Further to this, pharmacoepidemiology studies usually report a risk difference for a binary exposure using a binary instrument, whereas Mendelian randomization studies report on a continuous exposure using a discrete or continuous genetic instrument (count of alleles or allele score respectively). As a result of these differences, as well as the stronger instruments and larger causal effects seen in pharmacoepidemiology, there is a need for a dedicated power calculator for instrumental variable analysis in the context of this field.

This paper will address how to conduct power calculations for pharmacoepidemiological studies using a single binary instrument to analyse the causal effect of a binary exposure on a continuous outcome. The formula to calculate power will be derived and then validated by a simulation study. The formula is distinct from existing instrumental variable power formulae due to its parametrization, which is designed specifically for ease of use by pharmacoepidemiologists. An online calculator, as well as packages in both R and Stata, are provided for the implementation of the formula by others.

Methods and Results

Let us consider physicians' prescribing preferences for two different treatments–for example a treatment of interest and a control treatment–as an instrument for exposure to these treatments. Physicians' preferences are generally not directly observable, so each physician's prescriptions to previous patients are used as a proxy for their preferences. This results in a binary instrument that takes a value of one if the physician issued a prescription for the treatment of interest to their previous patient and a value of zero if they prescribed the control treatment. We will derive the formula for the power of studies that use this instrument to measure the causal effect of a drug exposure on a continuous outcome, for example systolic blood pressure or low-density lipoprotein cholesterol.

Formula derivation

The instrumental variable analysis we consider requires the following three variables; namely a binary instrument Z, a binary exposure X and a continuous outcome Y. The outcome for patient i, for $i = 1, \ldots, n$, is modelled as follows:

$$Y_i = \alpha + \beta X_i + U_i$$

where U_i is a zero-mean error term containing unobserved confounders, determining both the outcome Y_i and the treatment X_i. The instrument Z_i affects treatment X_i, but is not associated with the unobserved confounders and has no direct effect on the outcome.

Let $\tilde{Y}_i = Y_i - \overline{Y}$, $\tilde{X}_i = X_i - \overline{X}$ and $\tilde{Z}_i = Z_i - \overline{Z}$, where $\overline{Y}$, $\overline{X}$ and $\overline{Z}$ are sample averages. Denote by $\tilde{y}$, $\tilde{x}$ and $\tilde{z}$ the n-vectors of observations on $\tilde{Y}_i$, $\tilde{X}_i$ and $\tilde{Z}_i$, respectively. The two-stage least squares (2SLS) estimator of β is then given by:

$$\hat{\beta} = (\tilde{z}'\tilde{x})^{-1}\tilde{z}'\tilde{y}.$$

The variance of the 2SLS estimator is:

$$\mathrm{Var}(\hat{\beta}) = \sigma^2(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}$$

where $P_{\tilde{z}} = \tilde{z}(\tilde{z}'\tilde{z})^{-1}\tilde{z}'$ and $\sigma^2 = E(U_i^2)$ is the residual variance. Note that conditional homoscedasticity holds, so the variance is constant for all values of the instrument i.e. $E(U_i^2) = E(U_i^2|Z_i) = \sigma^2$ for $i = 1, \ldots, n$.

Consider the term $\tilde{x}'P_{\tilde{z}}\tilde{x}$:

$$\tilde{x}'P_{\tilde{z}}\tilde{x} = \tilde{x}'\tilde{z}(\tilde{z}'\tilde{z})^{-1}\tilde{z}'\tilde{x} = n\left(\frac{\tilde{x}'\tilde{z}}{n}\right)\left(\frac{\tilde{z}'\tilde{z}}{n}\right)^{-1}\left(\frac{\tilde{z}'\tilde{x}}{n}\right)$$

Let $p_Z = P(Z = 1)$, $p_X = P(X = 1)$ and $p_{XZ} = P(X = 1|Z = 1)$. In large samples:

$$\left(\frac{\tilde{z}'\tilde{z}}{n}\right) \approx \mathrm{Var}(\tilde{z}) = p_Z(1 - p_Z)$$

$$\left(\frac{\tilde{x}'\tilde{z}}{n}\right) = \left(\frac{x'z}{n} - \overline{XZ}\right) \approx p_Z(p_{XZ} - p_X)$$

Hence $\tilde{x}'P_{\tilde{z}}\tilde{x}$ can be presented in the following way:

$$\tilde{x}'P_{\tilde{z}}\tilde{x} \approx \frac{n\left(p_Z(p_{XZ} - p_X)\right)^2}{p_Z(1 - p_Z)}$$

Now consider the instrumental variable estimator of β. Using the asymptotic distribution $\hat{\beta} \sim N\left(\beta, \sigma^2(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}\right)$, the distribution of the t-test statistic under the null hypothesis $H_0 : \beta = \beta_0$ is:

$$t = \frac{\hat{\beta} - \beta_0}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}} \sim N(0, 1)$$

The distribution of the test statistic under the alternative hypothesis $H_1 : \beta = \beta_0 + \delta$ is:

$$t = \frac{\hat{\beta} - \beta_0}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}} = \frac{\hat{\beta} - \beta_0 - \delta}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}} + \frac{\delta}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}}$$

$$\sim N\left(\frac{\delta}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}}, 1\right)$$

The null hypothesis is rejected if $|t| > c_\alpha$ where c_α is the critical value at significance level α.

The power is the probability that the test statistic will exceed the critical value, which is:

$$P(t > c_\alpha) + P(t < -c_\alpha) = \Phi\left(-c_\alpha + \frac{\delta}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}}\right) + \Phi\left(-c_\alpha - \frac{\delta}{\sigma\sqrt{(\tilde{x}'P_{\tilde{z}}\tilde{x})^{-1}}}\right)$$

where $\Phi(s)$ is the cumulative standard normal distribution function evaluated at s. Power therefore increases as the value of σ decreases and/or the value of $\tilde{x}'P_{\tilde{z}}\tilde{x}$ increases. By substituting $\tilde{x}'P_{\tilde{z}}\tilde{x}$ and simplifying, we obtain the following formula for power:

$$Power = \Phi\left(-c_\alpha + \frac{\delta\left(p_Z(p_{XZ} - p_X)\right)\sqrt{n}}{\sigma\sqrt{p_Z(1 - p_Z)}}\right) + \Phi\left(-c_\alpha - \frac{\delta\left(p_Z(p_{XZ} - p_X)\right)\sqrt{n}}{\sigma\sqrt{p_Z(1 - p_Z)}}\right)$$

The formula requires a total of seven parameters to be specified. This includes four parameters that must always be specified–these are the significance level, α; the size of the causal effect, δ; the residual variance, $\sigma^2 = E(U_i^2)$; and the sample size, n. Also three that can be chosen from the following four parameters–these are the frequency of the instrument, $p_Z = P(Z = 1)$; the frequency of exposure, $p_X = P(X = 1)$; the probability of exposure given the instrument $Z = 1$, $p_{XZ} = P(X = 1|Z = 1)$; and the probability of exposure given the instrument $Z = 0$, $p_{XZ} = P(X = 1|Z = 0)$. The chosen parameters must be specified so that the following holds:

$$P(X = 1) = P(X = 1|Z = 0)P(Z = 0) + P(X = 1|Z = 1)P(Z = 1)$$

The formula for power is available for use via an online calculator [https://venexia.shinyapps.io/PharmIV/] and packages for R and Stata can be downloaded from GitHub [https://github.com/venexia/PharmIV].

Note that the frequency of exposure in an instrumental variable analysis of this type is likely to be higher than in a general population study because a drug is compared against one or more other drugs in a population of people with the indication for these treatments. General population studies, on the other hand, tend to compare a population who received the drug of interest with a population who did not receive it, and consequently the frequency of
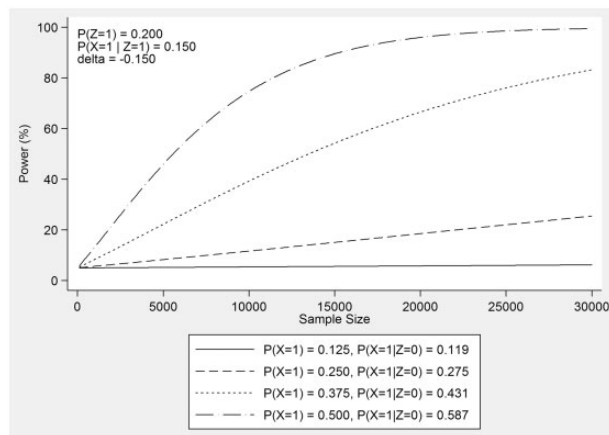
Figure 1. Power curves for several values of the frequency of exposure $p_X = P(X = 1)$ that show the effect on the power of a study to detect a causal effect of $\delta = -0.150$ using an instrument with a frequency of $p_Z = 0.200$, a residual variance of $\sigma^2 = 1$ and a sample size of up to 30 000 participants.

exposure is generally much lower. The effect of varying the parameters within the formula on a study's power is best presented graphically. Figure 1 illustrates an example of the effect of the frequency of the exposure $p_X = P(X = 1)$ on the power of a study to detect a causal effect of $\delta = -0.150$ using an instrument with a frequency of $p_Z = 0.200$, a residual variance of $\sigma^2 = 1$ and a sample size of up to 30 000 participants. Both increasing the frequency of exposure up to 50% and increasing the sample size results in increased power for this study.

Formula validation

To validate the power formula, we conducted a simulation. We simulated the data by defining the three variables necessary to conduct instrumental variable analysis with a single instrumental variable as follows:

$$\text{Instrument}: \quad Z_i \sim \text{Binomial}(1, p_Z)$$

$$\text{Exposure}: \quad X_i \sim \begin{cases} 0, \text{if } c_0 + Z_i(c_1 - c_0) + V_i \leq 0 \\ 1, \text{if } c_0 + Z_i(c_1 - c_0) + V_i > 0 \end{cases}$$

$$\text{Outcome}: Y_i \sim \delta X_i + U_i$$

where $p_Z = P(Z = 1)$ is the frequency of the instrument, $c_j = \Phi^{-1}\left(P(X = 1 | Z = j)\right)$ for $j = 0, 1$ are the inverse cumulative standard normal distribution, or quantile, functions of the conditional probabilities of exposure given the instrument, δ is the causal effect, and U_i and V_i are standard normally distributed error terms with covariance ρ.

The formula uses a binary instrument, binary exposure and continuous outcome and so the above variables were simulated to recreate data of this form. The instrument Z

is modelled by a binomial distribution parameterized by its frequency $p_Z = P(Z = 1)$. This ensures a binary variable with the correct probability of success. The exposure X is also binary but is modelled using a threshold model. The variability in the equation for the exposure comes from the normally distributed error term V_i. The use of the model equation allows the exposure X to be associated with the instrument Z. The outcome Y is modelled by its model equation $Y_i = \delta X_i + U_i$. In the model, the instrument is valid as the outcome Y is only associated with the exposure X, as dictated by the causal effect δ, and is not associated with the instrument Z other than through the exposure X.

Using the generated data, we performed an instrumental variable analysis using the command IVREG2 in Stata.[9] From this analysis, we recorded the coefficient of the exposure X with the 95% confidence interval. We then counted the number of simulations for which the confidence interval excluded the null, and divided this by the total number of simulations to determine the power. By running the simulation and calculating the formula using the same parameters, we are able to validate the formula against the simulation.

We present the power calculated from both the simulation and the formula for several parameter combinations in Table 1. The table contains 27 different simulations and each was repeated 10 000 times. The simulations consider each combination of three values of the frequency of exposure, $p_X = 0.100, 0.250, 0.500$; three values of the probability of exposure given the instrument $Z = 1$, $p_{XZ} = 0.150, 0.300, 0.450$; and three values of the sample size, $N = 10000, 20000, 30000$. We set the frequency of the instrument, $p_Z = 0.200$; the causal effect, $\delta = -0.150$; the residual variance, $\sigma^2 = 1$; and calculated $P(X = 1 | Z = 0)$ according to the following equation:

$$P(X = 1 | Z = 0) = \frac{P(X = 1) - P(X = 1 | Z = 1)P(Z = 1)}{1 - P(Z = 1)}$$
$$= \frac{p_X - p_{XZ} p_Z}{1 - p_Z}$$

The effect of confounding was removed as a parameter because the power was insensitive to its value in the simulation setting. Details of the simulations conducted to test this can be found in Supplementary File 1, available as Supplementary data at *IJE* online. The Stata code for this paper, including that used to create the simulation, is available from GitHub [https://github.com/venexia/PharmIV].

Simulation results

The formula and the simulation consistently provide similar results, with an absolute mean difference of 0.4% for

Table 1. A comparison of the power calculated from the formula and a validation simulation for an instrumental variable analysis where the causal effect $\delta = -0.150$, the frequency of the instrument $p_Z = 0.200$ and the residual variance $\sigma^2 = 1$

p_X	p_{XZ}	10 000 patients		20 000 patients		30 000 patients	
		Formula	Simulation	Formula	Simulation	Formula	Simulation
0.100	0.150	6.6%	6.1%	8.3%	7.9%	10.0%	9.8%
	0.300	32.3%	33.3%	56.4%	55.5%	73.8%	73.9%
	0.450	74.7%	75.6%	96.0%	95.9%	99.5%	99.5%
0.250	0.150	11.7%	11.4%	18.6%	18.3%	25.5%	25.5%
	0.300	6.6%	5.4%	8.3%	7.9%	10.0%	9.8%
	0.450	32.3%	32.8%	56.4%	56.1%	73.8%	73.7%
0.500	0.150	74.7%	74.2%	96.0%	95.9%	99.5%	99.6%
	0.300	32.3%	32.5%	56.4%	57.1%	73.8%	73.7%
	0.450	6.6%	5.0%	8.3%	7.2%	10.0%	10.1%

the parameter combinations presented in Table 1. There is also no discernible pattern in the differences, suggesting systematic bias is not present. Further to this, the power is consistent with its behaviour in other established power calculations. For example, increasing sample size universally improves power for all parameter combinations.

Discussion

In this paper, we have derived the formula necessary to calculate power for instrumental variable analysis with a single binary instrument, binary exposure and continuous outcome in the context of pharmacoepidemiology. The formula has been shown to be valid by comparison against a simulation study, which concluded that the formula provided near true values across a range of realistic parameters.

We acknowledge that there is some overlap of this calculator with existing calculators such as that proposed by Burgess for Mendelian randomization.[7] Although both calculators ultimately have a shared aim, namely to calculate the power of a study using an instrumental variable analysis, their application makes the calculators distinct. This is evident from the choice of parameterization of the power formula. The Mendelian randomization calculator opts for the coefficient of determination, R^2. This is a natural choice for the application, as it summarizes the proportion of the variance expected to be explained by genetic factors. In contrast, we have opted to parameterize our calculator for use in pharmacoepidemiological studies in terms of the frequency of the instrument, the frequency of the exposure and the conditional probabilities that relate them. This is more intuitive to a pharmacoepidemiological audience who will typically use the proportion of patients exposed, i.e. the frequency of exposure. In addition to this, the instruments typically used in this framework–for example, physicians' prescribing preference–do not necessarily fit as naturally to the notion of variance explained and are

summarized much more easily by their frequency and their relationship with exposure.

A concern for any instrumental variable analysis, whether in the context of pharmacoepidemiology or not, is the strength of the instrument. Instruments are termed weak when the correlation between the instruments and the exposure is low.[10] A commonly cited threshold is a partial F statistic of the association between the instrument and the exposure of less than 10.[8,11] Weak instruments will result in low power to detect a causal effect.[12–14] They are also known to induce bias, as such instruments may explain only a small proportion of the association between the exposure and outcome. Therefore, although pharmacoepidemiological studies are likely to have stronger instruments than other forms of instrumental variable analysis such as Mendelian randomization, researchers should remain mindful of their choice of instrument and whether it is appropriate for the research question they wish to study.

As for any power formula, the formula presented here is limited by its parameters, which simplify the dataset being considered. Power calculated from such formulae cannot account for dataset characteristics outside these parameters. For example, the formula makes no allowance for the presence of missing data–a known limiting factor in the power of a study. By allowing for missing data in the anticipated sample size, conservative estimates for the power of a study can be obtained using the formula presented. Further work is needed in order to establish the formula for power in other scenarios that use instrumental variable analysis within a pharmacoepidemiology context. This includes analyses with binary outcomes and analyses that involve multiple instrumental variables.

As the use of instrumental variable analysis in pharmacoepidemiology becomes more commonplace, there is an increasing need to provide power calculations for studies using this type of analysis. To provide such information, accessible and accurate power formulae need to be made

available. By using the formula presented here, either directly or via the online tool and packages in R and Stata, it is hoped that pharmacoepidemiologists can calculate the power of instrumental variable analysis studies with a single binary instrument, binary exposure and continuous outcome with ease.

Supplementary Data

Supplementary data area available at *IJE* online.

References

1. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology* 2006;**17**:360–72.
2. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;**29**:722–29.
3. Davies NM, Gunnell D, Thomas KH, Metcalfe C, Windmeijer F, Martin RM. Physicians' prescribing preferences were a potential instrument for patients' actual prescriptions of antidepressants. *J Clin Epidemiol* 2013;**66**:1386–96.
4. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf* 2010;**19**:537–54.
5. Brookhart MA, Wang P, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology* 2006;**17**:268–75.
6. Davies NM, Davey Smith G, Windmeijer F, Martin RM. COX-2 selective nonsteroidal anti-inflammatory drugs and risk of gastrointestinal tract complications and myocardial infarction: an instrumental variable analysis. *Epidemiology* 2013; **24**:352–62.
7. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol* 2014;**43**:922–29.
8. Freeman G, Cowling BJ, Schooling CM. Power and sample size calculations for Mendelian randomization studies using one genetic instrument. *Int J Epidemiol* 2013;**42**:1157–63.
9. StataCorp. *Stata Statistical Software*. College Station, TX: StataCorp LP, 2015.
10. Bound J, Jaeger DA, Baker RM. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J Am Stat Assoc* 1995;**90**:443–50.
11. Hernán M, Robins J. *Causal Inference*. Boca Raton, FL: Chapman & Hall/CRC, forthcoming.
12. Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *J Clin Epidemiol* 2009;**62**:1226–32.
13. Rassen JA, Brookhart MA, Glynn RJ, Mittleman 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**:1233–41.
14. Davies NM, von Hinke Kessler Scholder S, Farbmacher H, Burgess S, Windmeijer F, Davey Smith G. The many weak instruments problem and Mendelian randomization. *Stat Med* 2015; **34**:454–68.