

**What is effect size?**

Effect size is the magnitude of the difference between two intervention groups. *Absolute* effect size is the raw difference between average outcomes of groups and does not take into account variability in results. Calculated effect size indices are useful when the measurements have no intrinsic meaning, such as numbers on a Likert scale; studies have used different measurement scales; or effect size is examined in the context of variability in subject responses.

**Why should I report effect size?**

The effect size is the most important finding in a quantitative study. It helps the reader determine whether the effort, time, and cost of an intervention are justified by the magnitude of the effect. Effect size should be reported in the Abstract and Results sections.

**Why is the *P* value insufficient?**

Statistical significance is the likelihood that the difference between two groups is due to chance, a sampling accident (type I error). With a sufficiently large sample, a statistical test will always demonstrate a significant difference unless there is no effect whatsoever; that is, when the effect size is exactly zero. Statistical significance does not predict effect size.

**How do I calculate effect size?**

Choice of calculated effect size depends on the analysis being performed. In a comparison of mean outcomes between two independent groups, Cohen *d* can be measured, which is the difference between the means divided by the standard deviation of one of the groups. Statistical experts at your site and online calculators can help you.

**What is the statistical power of my study?**

Statistical power is the probability that your study will find a difference between interventions when an actual difference does exist. If statistical power is high, the likelihood of deciding there is no effect, when one does exist, is reduced. Power is  $1 - \beta$ , where  $\beta$  is the probability of concluding there

is no effect when one actually exists. This type of error is termed type II error and is usually assigned a value of 0.20.

**When should I examine the power of my study?**

The calculation must be done *prior* to your study. Because statistical power is the probability that a test will find a difference when one exists (or will correctly reject the null hypothesis), it comes into play only when the null is false. Calculation done after the fact (post hoc analysis) assumes that the null hypothesis is false, when it actually could be true or false. Thus, the calculation is unreliable.

Researchers who omit this essential first step sometimes perform a post hoc calculation of power. If so, the calculation must be labeled as post hoc and may be of limited value.

**How do I calculate statistical power?**

Statistical software and online programs can perform these calculations.<sup>1,2</sup> Using prior studies or pilot work, the expected difference (or minimum difference considered relevant) between the two outcomes can be estimated (eg, using means and SDs from pilots or similar studies in the literature). If you know any three of the following, you can calculate the fourth:

- Effect size
- Sample size (*n*)
- $\alpha$  error (*P*), often set at .05
- Power, or  $1 - \beta$ , often set at 0.8

**How can I increase power?**

To increase the power of your study, use more potent interventions that have bigger effects; increase the size of the sample/subjects; reduce measurement error (use highly valid outcome measures); and relax the  $\alpha$  level, if making a type I error is highly unlikely.

**What is the relationship between effect size and power?**

The statistical power of the study depends on the effect size and sample size. As the effect of the intervention decreases, more subjects will be needed to have adequate power; that is, to decide that “no difference” is a true finding. As the effect size increases (eg, a more effective intervention), a smaller number of subjects will be needed to have the same power to conclude that “no difference” is a true finding.

**Where can I find out more about effect size?**

These references are a good start.

Gail M. Sullivan, MD, MPH, is Editor-in-Chief, *Journal of Graduate Medical Education*.

Corresponding author: Gail M. Sullivan, MD, MPH, University of Connecticut, 253 Farmington Avenue, Farmington, CT 06030-5215, 860.679.3863, gsullivan@nso1.uhc.edu

DOI: <http://dx.doi.org/10.4300/JGME-D-12-00162.1>

**References**

- 1 Ellis PD. Effect size calculators. <http://www.polyu.edu.hk/mm/sizefaq/calculator/calculator.html>. Accessed April 15, 2012.
- 2 Soper D. Statistics calculators. <http://danielsoper.com/statcalc3/default.aspx>. Accessed April 16, 2012.

**Resources**

- Ellis PD. *The Essential Guide to Effect Size*. Cambridge University Press, Cambridge, UK; 2010.
- Field AP, Wright DB. A bluffer's guide to effect sizes. *PsyPAG Quarterly*. 2006;58:9–23.
- Vaughn RD. The importance of meaning. *Am J Public Health*. 2007;97(4):592–593.