JRH
www.jrh.gmu.ac.ir

Effect size: an overlooked issue in published articles in Journal of Research & Health

Abdoljavad Khajavi[1]

Journal of Research & Health
Social Development & Health Promotion Research Center
Vol. 4, No.4, Winter 2014
Pages: 861-862
Letter to Editor

1. **Correspondence to**: MD, PhD, Assistant Professor of Health Services Management
Tel/Fax: +98 51 57229025
Email: abjkhajavi@gmu.ac.ir

Received: 31 Aug 2014
Accepted: 16 Sep 2014

How to cite this article: Khajavi A. Effect size: an overlooked issue in published articles in Journal of Research & Health. *J Research Health*2014; 4(4): 861-862.

Up to now, from fall 2011 to summer 2014, eighty seven articles have been published in nine issues of the Research & Health Journal. In reviewing these articles, it is clear that although 61 out of the 87 articles (70%) used inferential statistics, none of them reported or interpreted the effect sizes (ESs) along with P-values for statistically significant or non-significant results. ESs, however, allow researchers to move away from the simple expression of statistical significance and toward a more generally interpretable, quantitative description of the size of an effect [1,2]. Despite its potential to illuminate research findings, health researchers seem reluctant to include ES measures in their work [3]. Health professionals, on the other hand, are all increasingly expected to base their practices within evidence-based frameworks. They are not only coming to depend on research for improving their practices, they also rely on researchers appropriately interpreting the results they produce [3]. A pertinent issue in this regard is whether researchers provide important information, such as ESs, to facilitate accurate interpretations of findings. Given the growing trend in scientific journals toward reporting of ESs [2,3] and to invite researchers to go beyond the mere reporting of P-values in their works, the issue is briefly discussed here.

Recall that in the null hypothesis testing, researchers use statistical tests to ask the question, 'If the null hypothesis is true, what is the probability of obtaining the relationship that was found in my sample?' That is, the significance test tells us whether the result was likely obtained by chance or variability in the sample, but does not convey information about the practical or theoretical importance of the difference (ES) [1]. A large sample may enable an effect to reach statistical significance, but the effect may be of negligible importance. Alternatively, a small sample may fail to reach statistical significance, although the result may be practically important in the real world [4]. A significance test thus is properly only one among many criteria by which a finding is assessed. One way that researchers can assess the importance of their finding is to calculate the ES (strength of association) [1, 4]. Providing both a P-value and an E-value (ES) as a measure of the 'statistical' and the 'practical' significance of a result are indeed two sides of a coin—they complement each other but do not substitute for one another [5]. ES can be broadly defined as any statistic that quantifies the degree to which sample results diverge from the expectations described in the null hypothesis [3]. It is, more specifically, a set of statistics that indicates the relative magnitude of the differences between means, or the amount of the total variance in the dependent variable that is predictable from knowledge of the levels of the independent variable [1]. Although there are over 40 different measures of ESs, their families have been categorized into two broad groups: measures of mean differences (the d family) and measures of strength of relations (the r family) [2,3]. The d family is based on the standardized group mean difference and represented, for example, by Cohen's d and Glass's g. The r family is based on

Archive of SID

www.SID.ir

the proportion of variance accounted for or correlation between the independent variable and the dependent variable and represented, for example, by eta squared (η^2) and partial eta squared (ηP^2). The most commonly used ES statistics to compare groups are η^2, ηP^2 and Cohen's d (1). IBM SPSS calculates ηP2 as part of the output from some techniques (eg, analysis of variance), but does not provide η^2 or Cohen's d values, for example, for t-tests [4]. You can, however, use the information provided in the SPSS to calculate whichever ES statistic you need (eg, for t-tests or non-parametric techniques) [4]. For a simple, practical guide to the calculation of ESs, see Pallant [4]. For a more detailed discussion, see Fritz et al [5] and Tabachnick & Fidell [1]. Cohen's d is also reported as the ES statistic in medical literature. This ES describes the difference between groups in terms of standard deviation units, and not on the percentage of variance used in η^2 [1,4]. To calculate Cohen's d you can go to http://www.uccs.edu/lbecker/index.html. for a free online calculator.

A frequent question, in the context of ES, is 'Do we have or expect to find a large effect?' [2] There is no straightforward answer to the question, it depends partly on the research area and type of the study [1,2]. There is also no single criteria for interpreting different ES measures, though Cohen has presented some useful guidelines for 'small', 'medium' and 'large' effects [1]. It should be noted that the Cohen's guidelines also use different criteria for interpreting the different ES measures, see Pallant [4] for more details. In addition to reporting and interpreting the ES estimates, researchers should use the 'Effect Size' terminology in their writing in all instances of reporting strength of association or mean differences [1,2]. The language alerts the reader that an estimate of practical significance (ES) is being reported. Otherwise, some readers may not identify particular statistics as being an estimates of ES [2,5].

In brief, it is questioned that researchers interpret their empirical findings solely through the lens of null hypothesis significance testing [1,3,5]. To address the problem, it seems promising to use both a P-value and an E-value of a result. I hope, therefore, that the researchers who wish to publish in the Journal of Research & Health, are more interested in reporting and interpreting the ES estimates of their results.

References

1. Tabachnick BG, Fidell LS. Using multivariate statistics. 6th international edition. Ed. Boston, [Mass.]; London: Pearson; 2013.
2. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. Journal of Experimental Psychology: General. 2012; 141(1): 2-18.
3. Gaskin CJ, Happell B. Power, effects, confidence, and significance: An investigation of statistical practices in nursing research. *International Journal of Nursing Studies*2014; 51(5): 795-806.
4. Pallant J. SPSS survival manual: a step by step guide to data analysis using IBM SPSS. Maidenhead: McGraw-Hill. 2013.
5. Fritz A, Scherndl T, Kühberger A. A comprehensive review of reporting practices in psychological journals: Are effect sizes really enough? *Theory & Psychology*2012; 23(1): 98-122.

Archive of SID

www.SID.ir