



Editorial

The importance of reporting the effect size in quantitative studiesNasrin Nikpeyma¹, Saman Maroufizadeh², Maryam Esmaeili^{3*}¹Department of Community Health and Geriatric Nursing, School of Nursing and Midwifery, Tehran University of Medical Sciences, Tehran, Iran²Department of Biostatistics, School of Nursing and Midwifery, Guilan University of Medical Sciences, Rasht, Iran³Nursing and Midwifery Care Research Center, School of Nursing and Midwifery, Tehran University of Medical Sciences, Tehran, Iran

In 2001, the American Psychology Association advised researchers to provide readers with information about statistical significance and, even more importantly, to provide information about the magnitude of the observed effect size and the confidence interval (1). A review of published papers in the fields of behavioral sciences, psychology, medical sciences, and nursing shows that only a few papers have reported on the calculation of the effect size or related interpretations (2). The null hypothesis significance testing (NHST) and the reports on the statistical significance value have long been considered as an incomplete measure for analysis (3). The statistical significance of the null hypothesis is made up of several factors: the actual effect, the size of the parameter in the population, the sample size, and the selected α level (4). The statistical level of significance or P-value refers to the probability of rejecting the null hypothesis provided that it is approved based on the observed data. In other words, the observed difference is not random (there is zero chance involved), so that the lower this value, the more one can trust the accuracy of the observed difference (5).

One of the problems with tests of statistical significance is that they can only indicate that the observed relationship is not by chance. However, these tests cannot prevent the probability of the significance of minimal differences in large populations. Thus, one of the limitations of the level of significance is that it depends on the sample size in a way that increasing the sample size

can lead to a decrease in the P-value. In fact, tests of statistical significance are used to reduce the type of I error. Meanwhile, such tests cannot protect the significance level of the research findings against the possibility of the type 2 error. The statistical significance (P-value) shows that there are some effects in the research, but it does not specify the value of the effects (6). Therefore, many critics argue that the indexes of effect size and confidence interval (CI) should also be considered in the reporting of the findings (7).

The confidence interval provides an estimate of the range of the outcome measurement for a particular population. Research studies often use a 95% confidence level, which indicates that the researcher can strongly claim that the average score of the population is within the reported range. In general, the researcher determines a range and can make sure that the average score will be within that range, given 95% confidence (8).

The effect size is defined as an objective and standard index which estimates the effect of correlation between two or more variables (7). Unlike NHST, the effect size is resistant against sample size, and therefore it can provide a more accurate measurement of the effect of an intervention or the extent of the relationship. By examining the magnitude of the effect size, the effect of the independent variable on the dependent variable can be measured. Accordingly, the effect size is the main finding in quantitative research. For example, based on Beck's questionnaire, the scores of depression in two different imaginary groups were (3,5,7,8,6,4) and (5,7,11,9,7,8), respectively.

* Corresponding Author: Maryam Esmaeili, Nursing and Midwifery Care Research Center, School of Nursing and Midwifery, Tehran University of Medical Sciences, Tehran, Iran. Email: esmaeili_m@tums.ac.ir

DOI: <https://doi.org/10.18502/npt.v8i1.4486>

The t value is not statistically significant for the independent groups in these two groups ($t=2$) at the 95% confidence level ($p>0.05$). However, there is a relatively noticeable clinical difference in these scores. Hence, it is recommended to report the effect size even in studies with small samples (9). Cohen (1988) provided a guideline for interpreting the amount of effect size in statistical analysis, which is considered as one of the most well-known indexes of effect size. In applied researches, the effect size indicates the clinical significance of the variable's effect, which is beyond statistical significance (10). In general, the effect size refers to the mean difference or risk difference between the two independent and dependent groups divided by the standard deviation and is reported by Cohen's index (d). Accordingly, the effect size can be reported as small ($d=0.2$), average ($d=0.5$), or large ($d>0.8$) (9), but it should be noted that the research subject and its different dimensions will determine the interpretation of the results of the effect size (11). Standard and global web-based calculators can be used to estimate the effect size value.

Having the appropriate rationale to determine the sample size is one of the key components while designing clinical trial studies. However, according to the NHST framework, interpretation of the power requires determining the effect size of the study population in order to be able to calculate the necessary sample size to achieve the desired power. In the new perspectives, though, the calculation of sample size depends on determining the effect size (12).

Researchers should know that reporting both the level of significance and effect size can provide valuable and distinct information. Hence, if researchers report only one of these values, it can be a source of error (8). In this regard, Fan (2001) used the analogy between the two sides of a coin and pointed out that the P-value and the effect size are complementary but not interchangeable. Therefore, researchers should report both indexes in their quantitative studies (13).

Reporting the effect size can reduce the likelihood of a type II error in the studies.

In addition, providing indexes of the confidence interval and effect size can obviate the limitations of sample size and can increase the applicability of clinical trial studies in meta-analysis. The studies which are observed in the meta-analysis are different in terms of sample size, measurement scales, and intervention design. Besides, there are some differences in estimates and statistical analysis (including structural parameters, variance, covariance, and residual variances). The effect size can integrate these contradictory estimates and facilitate the comparison and combination of findings (10). Accordingly, the authors should pay more attention to reporting statistical significance values and effect sizes in quantitative studies, especially intervention-based studies. Moreover, journal editorial policies should focus on reviewing quantitative articles in terms of reports on these important statistical indexes. It seems that reporting on the effect size has improved in recent years, but it is still necessary to be observed in nursing articles. Thus, journal editors and researchers should implement efficient reporting methods to achieve more valuable and reliable data.

References

1. Hinchliffe LJ, Brown D. American Psychological Association. Publication manual of the American psychological association (5th edition). Boston: McGraw hill; 2001.P:186.
2. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General*. 2012 Feb;141(1):2.
3. Wei R, Hu Y, Xiong J. Effect size reporting practices in applied linguistics research: A study of one major journal. *Sage Open*. 2019 May;9(2):1-11.
4. Ferguson, C. J. An effect size primer: A guide for clinicians and researchers. *Professional Psychology. Research and Practice*. 2009; 40(5): 532-8. DOI: <https://doi.org/10.1037/a0015808>
5. Osborne JW. Sweating the small stuff in educational psychology: how effect size and power reporting failed to change from 1969 to 1999, and what that means for the future of

Importance of effect size

changing practices. *Educational Psychology*. 2008;28(2):151-60.

6. Field, A. P. *Discovering statistics using SPSS: (and sex and drugs and rock "n" roll)* (3rd ed.). Los Angeles, CA: Sage.2009.

7. Farahani, H, Rahiminejad HR, Ghomrani AA. Common statistical error in behavioral sciences (Persian). *Seminary & University Journal of Methodology of Social Science*.2008;14(55):85-94.

8. Lee DK. Alternatives to P value: confidence interval and effect size. *Korean journal of anesthesiology*. 2016 Dec;69(6):555-62.

9. Akbari Zarradkhaneh S. Necessity of Measuring and Reporting Effect Size Measures in Psychological Studies. *Rooyesh-e-Ravanshenasi Journal (RRJ)*. 2016 Jun 10;5(1):29-46.

10. Pogrow S. How effect size (practical significance) misleads clinical practice: The case for switching to practical benefit to assess applied research findings. *The American Statistician*. 2019 Mar 29;73(sup1):223-34.

11. Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. *Journal of graduate medical education*. 2012 Sep;4(3):279-82.

12. Gignac GE, Szodorai ET. Effect size guidelines for individual differences researchers. *Personality and individual differences*. 2016 Nov 1;102:74-8.

13. Fan X. Statistical significance and effect size in education research: Two sides of a coin. *The Journal of Educational Research*. 2001 May 1;94(5):275-82.