A New Standardised Effect Size, *e*

Sir,

Effect size has been reported in scientific literature using several different measures. This variation makes it difficult to compare studies and conduct meta-analyses. A new standardised effect size is proposed that will make it easier to compare results and pool outcomes from multiple studies. This new effect size, designated *e*, ranges from -1 to +1 (0 indicating no effect, analogous to the correlation coefficient). The interpretation of *e* is set as: ±0.2, small effect; ±0.5, moderate; and ±0.8, large effect. This interpretation is independent of practical importance which has to be evaluated subjectively, based on considerations such as usefulness, time and cost. The value of *e* can be obtained from any measure for effect size such as the odds ratio by transforming to a scale of -1 to +1. A standardised effect size measure may foster more widespread reporting of effect size over significance testing and a move towards meta-analytic thinking. In addition to mentioning confidence intervals and p-values, empirical research should report the *e*-value of all outcome measures, whether statistically significant or not.

Effect size is a measure of the magnitude of an observed result. In other words, effect size is the beacon of signal within a sea of noise. While several dozen effect size measures are available, "different effect size measures have different properties, different ranges of values, and different interpretations",[1] which makes it challenging for readers to comprehend them. Hence, there is a need for an effect size measure with a standardised range of values and a generally accepted interpretation.

The solution proposed in this letter is a new universal format for reporting effect size, the standardised effect size, designated as *e*. This standardised effect size can be derived from any effect size by transforming to a range of -1 to +1. Interpretation of values is uniformly set as: ±0.2, small effect; ±0.5, moderate effect; and ±0.8, large effect (Figure 1). This interpretation is independent of the practical significance, which has to be assessed
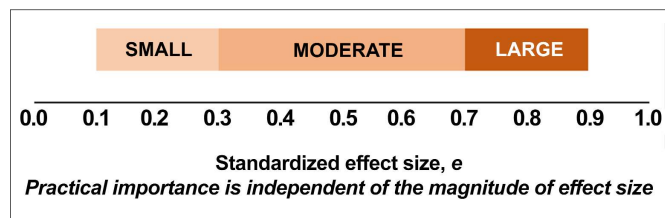

Figure 1: A visual representation of the standardised effect size, *e*.

subjectively, based on factors such as social preferences, perceived usefulness of the outcome, and cost.

Readers of empirical studies will find it easier to interpret effect size when presented in a universal standardised format. Akin to the p-value of significance testing, the *e*-value will be interpretable without additional explanation. Just as standardised p-values are derived from a multitude of statistical tests, the *e*-value can be reported for various outcome measures. Indeed, many current readers of scientific literature would be at a loss to identify partial η^2 as an effect size, let alone accurately interpret its reported values. Consider that the following values are all interpreted as a 'large' effect size: 0.4 for Cohen's *f*, 0.5 for Pearson's *r*, 0.14 for eta-squared, 0.8 for Cohen's *d*, and 4 for odds ratio.[2] Odds ratios, despite widespread use, tend to be misleading as "a measure of effect size, [as] their range is not standardised between 0 and 1, making their interpretation difficult".[3] Note that Cohen's *d* changes somewhat anomalously from fractional values to greater than 1.0 at a corresponding *r* of 0.45 and odds ratio of 6.1. Moreover, these discrepancies have contributed to the confusion between effect size interpretation and practical importance.

Meta-analysis of results from published studies involves pooling of effect size (weighted by sample size and other factors). When effect sizes are reported using different measures such as correlations, coefficients and odds ratios, pooling may be challenging. Although formulae exist for converting one effect size measure to another, this requires access to additional statistical data.[4] Researchers conducting meta-analyses routinely face such challenges and often have to reluctantly exclude studies from their analyses due to these reasons. A standardised effect size measure can facilitate meta-analyses. Effect size estimation is essential in planning the sample size of a study. A standardised effect size will help researchers plan studies to detect effects substantial enough to be practically worthwhile. This line of thinking emphasises pooling and replication of results from multiple studies to discern the overall direction and magnitude of an intervention. In contrast, null hypothesis significance testing emphasises a false dichotomy of whether or not there is any difference. The latter is well known to be influenced by the sample size chosen by the researcher. By streamlining the reporting and interpretation with a standardised effect size, researchers can focus on the emerging picture of the pooled trend in cumulative studies.

A critical issue with effect sizes is the interpretation of values. In a widely cited article, one expert noted the lack of "agreement on what magnitude of effect size is necessary to establish practical significance."[4] Taking a

different stance, we recommend decoupling the magnitude of effect size and its practical significance. While an effect size can be small, moderate or large, the practical significance has to be subjectively evaluated based on issues such as impact, cost and value to society. A small effect size ($e = 0.2$) may be valuable if it results in saving lives while a moderate effect size ($e = 0.5$) may not be worthwhile, if the benefits are transient and costly.

Transforming various effect size measures to a range of -1 to +1 is hereby proposed. Formulae for converting between effect size measures are well known in statistical literature.[5] An easy-to-use online conversion tool is available at https://www.psychometrica.de/effect_size.html.[6] While some have argued against transforming one effect size to another, others have conceded that these "conversions are conceptually helpful and also an essential technique for meta-analysis when integrating the results of studies which have employed different methods (*e.g.* a correlational approach and a two-group design)".[6]

Potential limitations of the standardised effect size *e*, include abstraction from raw data, arbitrary choice of underlying effect size measure and the hint of causality implied with the word 'effect.' Researchers conducting meta-analyses need to be aware of heterogeneous study designs and different outcome measures when pooling effect sizes.

REFERENCES

1. Sun S, Pan W, Wang LL. A comprehensive review of effect size reporting and interpreting practices in academic journals in education and psychology. *J Educ Psychol* 2010; **102**:989-1004.

2. Maher JM, Markey JC, Ebert-May D. The other half of the story: Effect size analysis in quantitative research. *CBE-Life Sci Educ* 2013; **12**:345-51.

3. Livingston EH, Elliot AC, Hynan LS. Clinical importance of effect size in randomized controlled trials-reply. *Arch Surg* 2010; **145**:401.

4. Ferguson CJ. An effect size primer: A guide for clinicians and researchers. *Prof Psychol Res Pract* 2009; **40**:532-8.

5. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* 2007; **82**:591-605.

6. Lenhard W, Lenhard A. Calculation of effect sizes [Internet]. Dettelbach (Germany): Psychometrica. DOI: 10.13140/RG.2.1.3478.4245; 2016. Available from: https://www.psychometrica.de/effect_size.html

Muhammad Jawad Hashim

Department of Family Medicine, College of Medicine, UAE University, Al-Ain, UAE

Correspondence: Dr. Muhammad Jawad Hashim, Department of Family Medicine, College of Medicine, UAE University, Al-Ain, UAE
E-mail: physicianthinker@gmail.com

••••☆••••