

# A Standard Error: Distinguishing Standard Deviation From Standard Error

Rickey E. Carter

**A** recent Perspective in *Nature* issued a call for more transparency in the reporting of preclinical research (1). Although this article focused primarily on experimental design, it emphasized the need for improved reporting in the scientific literature. Within the context of preclinical studies, there have been discussions regarding the appropriate reporting of standard error (SE) and standard deviation (SD) (2–5); however, despite the recommendations, opportunities remain to improve upon the reporting of these statistics in the literature.

To first set the stage for the distinction between SD and SE, we start with the similarities. Both SD and SE measure variability or, informally, “spread.” As such, both statistics give a numerical summary of variability. Given this, how does one distinguish SE from SD?

The distinction is that one summarizes the variability of data and the other describes the variability of an estimated quantity. Let us consider the latter scenario first. It is readily apparent that if an experiment were to be replicated, then slightly different values would be observed. On average, one would expect to have similar numerical summaries (e.g., approximately the same mean across samples), but some sample-to-sample variation would occur. This type of variation is summarized by SE. Specifically, SE quantifies variability in sample estimates, which are oftentimes means but also can be estimated regression parameters, correlation coefficients, or another value. What is important to consider, however, is that SE can be estimated from a single sample. For example, the SE of the mean formula is commonly known to be  $SD/\sqrt{n}$ . This introduces the concept of what is SD.

SD is a measure of spread in data about the mean, as opposed to the variability in an estimated summary of data. When one wants to summarize the variability in data, whether it is sample characteristics or response patterns, SD should be used. When one seeks to summarize the variability, or precision, in an estimated quantity, such as the mean response for a particular experimental condition, SE or a function of the SE should be used.

The standard practice of reporting mean  $\pm$  SE is problematic from several statistical and conceptual perspectives. When these summary measures are used in the standard bar chart with error “whisker,” the presentation is actually consistent with a 68% CI. Practically, this

misrepresents the precision of the estimated mean because the 68% CI is at least half the width of the more generally accepted 95% CI. Furthermore, reporting mean  $\pm$  SE does not allow for ad hoc comparisons between groups because the confidence coefficient (the multiplier of the SE used for the creation of the CI) varies based on sample size and distributional assumptions. Therefore, it is recommended that an estimated summary be accompanied by a 95% CI in text and graphical displays when one wants to describe the precision of the estimate. Of course, when there is a priori (planned) justification for a level of significance other than  $\alpha=0.05$  (6), the prespecified level of significance can be used for the CI. Rarely would one expect  $\alpha=0.32$  (i.e.,  $P < 0.32$ ) to be used for such a level of significance.

In summary, when reporting the characteristics of a sample to express the variability in the observed values, the SD should be used. SE should be reported only when reporting the variation of estimated quantities. Even then, it has been suggested that SE should be multiplied by a confidence coefficient to produce a CI to allow for more robust statistical comparison of the reported data. The combination of SDs and CIs should be the preferred statistics reported in the literature to give the reader a clear impression of the variability of the observed values and precision of estimated summaries of the data, respectively.

## ACKNOWLEDGMENTS

This work was supported by grant number UL1 TR000135 from the National Center for Advancing Translational Sciences.

No potential conflicts of interest relevant to this article were reported.

## REFERENCES

1. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 2012;490:187–191
2. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. *Circulation* 1980;61:1–7
3. Altman DG, Bland JM. Standard deviations and standard errors. *BMJ* 2005; 331:903
4. Nagele P. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. *Br J Anaesth* 2003;90:514–516
5. Herxheimer A. Misuse of the standard error of the mean. *Br J Clin Pharmacol* 1988;26:197
6. Moore CG, Carter RE, Nietert PJ, Stewart PW. Recommendations for planning pilot studies in clinical and translational research. *Clin Transl Sci* 2011;4:332–337

From the Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota.

Corresponding author: Rickey E. Carter, carter.rickey@mayo.edu.

The contents of this letter are solely the responsibility of the author and do not necessarily represent the official views of the National Institutes of Health. DOI: 10.2337/db13-0692

© 2013 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.