



Contents lists available at ScienceDirect

Taiwanese Journal of Obstetrics & Gynecology

journal homepage: www.tjog-online.com

Original Article

Inappropriate use of standard error of the mean when reporting variability of study samples: A critical evaluation of four selected journals of obstetrics and gynecology

Wen-Ru Ko^{a,b}, Wei-Te Hung^{a,b}, Hui-Chin Chang^{c,d}, Long-Yau Lin^{a,e,*}^a Institute of Medicine, Chung-Shan Medical University, Taichung City, Taiwan^b Department of Anesthesiology, Chung-Shan Medical University and Chung-Shan Medical University Hospital, Taichung City, Taiwan^c School of Public Health, Chung Shan Medical University, Taichung City, Taiwan^d Library and Evidence Based Medicine Center, Chung Shan Medical University Hospital, Taichung City, Taiwan^e Department of Obstetrics and Gynecology, Chung-Shan Medical University and Chung-Shan Medical University Hospital, Taichung City, Taiwan

ARTICLE INFO

Article history:

Accepted 19 April 2013

Keywords:

impact factor
prescriptive statistics
standard deviation
standard error of the mean
variability of study sample

ABSTRACT

Objective: The study was designed to investigate the frequency of misusing standard error of the mean (SEM) in place of standard deviation (SD) to describe study samples in four selected journals published in 2011. Citation counts of articles and the relationship between the misuse rate and impact factor, immediacy index, or cited half-life were also evaluated.

Materials and methods: All original articles in the four selected journals published in 2011 were searched for descriptive statistics reporting with either mean \pm SD or mean \pm SEM. The impact factor, immediacy index, and cited half-life of the journals were gathered from *Journal Citation Reports Science* edition 2011. *Scopus* was used to search for citations of individual articles. The difference in citation counts between the SD group and SEM group was tested by the Mann–Whitney *U* test. The relationship between the misuse rate and impact factor, immediacy index, or cited half-life was also evaluated.

Results: The frequency of inappropriate reporting of SEM was 13.60% for all four journals. For individual journals, the misuse rate was from 2.9% in *Acta Obstetrica et Gynecologica Scandinavica* to 22.68% in *American Journal of Obstetrics & Gynecology*. Articles using SEM were cited more frequently than those using SD ($p = 0.025$). An approximate positive correlation between the misuse rate and cited half-life was observed.

Conclusion: Inappropriate reporting of SEM is common in medical journals. Authors of biomedical papers should be responsible for maintaining an integrated statistical presentation because valuable articles are in danger of being wasted through the misuse of statistics.

Copyright © 2014, Taiwan Association of Obstetrics & Gynecology. Published by Elsevier Taiwan LLC. All rights reserved.

Introduction

Researchers have observed a high rate of statistical errors in many journal articles including those in leading journals since the 1960s [1–4]. Inappropriate statistical reporting has gained more attention within the context of evidence-based medicine (EBM) because EBM critically depends on the quality of published research.

Descriptive statistics are often used to describe the study sample in medical research articles. Generally, continuous data must be

summarized by two indices [5,6] to be meaningful, an index of central tendency and an index of dispersion. If data are normally distributed, the sample should be described using the mean and standard deviation (SD).

Authors sometimes use the standard error of the mean (SEM) to describe the variability of study samples. Lang [7] has pointed out that using the SEM as a descriptive statistic is a common statistical error found in biomedical research articles. The SEM is not a descriptive statistic. It indicates the probability of the population mean falling around the range of the sample mean but not variability within the sample. The value of SEM is always smaller than SD so if a sample is described as mean \pm SEM, it might lead readers to underestimate the variability within the sample [8,9].

Conflicts of interest: The authors have no conflicts of interest to declare.

* Corresponding author. Institute of Medicine, Chung-Shan Medical University, Number 110, Section 1, Jianguo North Road, Taichung City 40201, Taiwan.

E-mail address: viola5563@gmail.com (L.-Y. Lin).

<http://dx.doi.org/10.1016/j.tjog.2013.04.035>

1028-4559/Copyright © 2014, Taiwan Association of Obstetrics & Gynecology. Published by Elsevier Taiwan LLC. All rights reserved.

Nagele [10] evaluated four anesthesia journals and found the rate of incorrect use of SEM to be 11.5–27.7%. Nagele's study [10] evoked lively discussion and comments about inappropriate statistical presentation. Since 2003, Nagele's paper [10] has been cited in 26 articles. In addition, Jaykaran and Yadav's [11] study noted misuse of mean \pm SEM to be the most common reason for inappropriate descriptive statistics in the *Indian Journal of Pharmacology* and the *Indian Journal of Physiology and Pharmacology*. Saurabh [12] revealed that inappropriate reporting of SEM was common in articles published in basic science journals. The issue of using SD or SEM had also been discussed in correspondence or letters to the editors of several medical journals [13–16].

The current analysis evaluated the frequency of inappropriate use of the SEM in published articles in four selected obstetrics and gynecology journals in 2011. *Obstetrics & Gynecology* and *American Journal of Obstetrics & Gynecology* are two of the leading journals in this field. Most obstetricians and gynecologists use these journals regularly, and not only in Western countries. *BJOG: An International Journal of Obstetrics & Gynaecology* and *Acta Obstetrica et Gynecologica Scandinavica* are European medical journals published in English. Articles from these journals are used most frequently in EBM learning courses or journal clubs by residents at Chung-Shan Medical University Hospital, Taichung City, Taiwan.

The relationship between misuse rate and impact factor, immediacy index, or cited half-life was evaluated. Citation counts were also compared between articles using SD and misusing SEM.

Materials and methods

All original articles published in *Obstetrics & Gynecology*, *American Journal of Obstetrics & Gynecology*, *BJOG: An International Journal of Obstetrics & Gynaecology*, or *Acta Obstetrica et Gynecologica Scandinavica* in 2011 were searched for descriptive statistics reporting using either mean \pm SD [or mean (SD)] or mean \pm SEM [or mean (SEM)]. SEM is an inferential statistic and it should not be used as a descriptive statistic. If SEM is used to describe the variability of the study sample, it will mislead readers to underestimate the variability of the study sample. Therefore, in the current study, reporting the SEM in the text, tables, and figures was considered inappropriate when the SEM was used to describe the variability of the study sample.

Articles reporting median and range and articles using only inferential statistics were excluded. Case reports, short research reports, short communications, review articles, and systemic reviews were also excluded. Conference abstracts were not evaluated.

The impact factor, immediacy index, and cited half-life of the four journals were gathered from *Journal Citation Reports Science edition 2011* in *ISI Web of Knowledge*, published by Thomson Reuters. *Scopus*, provided by Elsevier, was used to search citations of individual articles. All searches were concluded on September 30, 2012.

Frequency of misuse of SEM in these four journals was reported as a number and percentage. The difference in citation counts between articles using SD and misusing SEM was tested by the Mann–Whitney *U* test. The relationship between SEM misuse rate and impact factor, immediacy index, or cited half-life was also evaluated.

A review and certification of exemption for this research was granted by the Institutional Review Board of Chung-Shan Medical University Hospital.

Results

A total of 456 articles met the inclusion criteria, 121 in *Obstetrics & Gynecology*, 194 in *American Journal of Obstetrics & Gynecology*,

72 in *BJOG: An International Journal of Obstetrics & Gynaecology*, and 69 in *Acta Obstetrica et Gynecologica Scandinavica*.

Three separate articles in *American Journal of Obstetrics & Gynecology*, *BJOG: An International Journal of Obstetrics & Gynaecology*, and *Acta Obstetrica et Gynecologica Scandinavica* that reported median \pm SD were not evaluated.

Twenty-two articles, 12 in *American Journal of Obstetrics & Gynecology*, three in *BJOG: An International Journal of Obstetrics & Gynaecology*, and seven in *Acta Obstetrica et Gynecologica Scandinavica*, only reported mean \pm figure without stating what parameter was used. These 22 articles were not included because neither SD nor SEM was noted in the text, tables, or figures.

Sixteen articles used both the SD and the SEM to describe their study samples, one in *Obstetrics & Gynecology*, 11 in *American Journal of Obstetrics & Gynecology*, two in *BJOG: An International Journal of Obstetrics & Gynaecology*, and two in *Acta Obstetrica et Gynecologica Scandinavica*. In these articles, the authors usually used SD to describe the demographic data and SEM to describe laboratory or test results. We also noticed that the use of SEM was more common in laboratory or basic science studies (41 in 62) than in clinical studies (21 in 394).

The frequency of inappropriate reporting of SEM was 13.6% (62 in 456) in total. For individual journals, the misuse rate ranged from 2.9% in *Acta Obstetrica et Gynecologica Scandinavica* to 22.68% in *American Journal of Obstetrics & Gynecology*. Impact factor, immediacy index, and cited half-life of each journal in 2011 are presented in Table 1. The relationships between the misuse rate and impact factor, immediacy index, and cited half-life are presented in Figs. 1–3 respectively.

Citation counts were different between articles using SD and misusing SEM. Surprisingly, articles using SEM were cited more frequently than those using SD (Table 2, mean rank 262.89 vs. 223.09, $p = 0.025$).

Discussion

The role of descriptive statistics is to describe a given study sample without regard to the whole population. SD shows variability of the mean within the sample. SEM is an inferential estimate of the stability of the mean of the study. The term “error” is used to indicate the fact that due to sampling error, each sample mean is likely to deviate somewhat from the true population mean. Thus, the SEM is a measure of precision for an estimated population mean but not a measure of data variability of the mean of a sample.

In general, readers are interested in variability within the sample not the proximity of the sample mean to the population mean. The SEM is calculated as SD divided by the square root of the sample size, so it is always less than the SD. Authors summarize their data with SEM as it makes data seem less variable and more representative. Readers may falsely conclude that the variability of

Table 1
Frequency of using standard error of the mean (SEM) and/or standard deviation (SD) in four selected journals listed in order of decreasing percentage of misusing SEM.

	Misuse of SEM, n (%)	Correct use of SD, n (%)	Total n	Impact factor ^a (2011)	Immediacy index ^a	Cited half-life ^a (y)
AJOG	44 (22.68)	150 (77.32)	194	3.468	0.782	9.9
BJOG	7 (9.27)	65 (90.28)	72	3.407	1.345	8.3
O&G	9 (7.44)	112 (92.56)	121	4.730	1.134	8.7
AOGS	2 (2.90)	67 (97.10)	69	1.771	0.294	8.4
Total	62 (13.6)	394 (86.40)	456	—	—	—

AJOG = American Journal of Obstetrics & Gynecology; AOGS = Acta Obstetrica et Gynecologica Scandinavica; BJOG = BJOG: An International Journal of Obstetrics & Gynaecology; O&G = Obstetrics & Gynecology.

^a Data from *Journal Citation Reports Science edition 2011*.

Download English Version:

<https://daneshyari.com/en/article/3975389>

Download Persian Version:

<https://daneshyari.com/article/3975389>

[Daneshyari.com](https://daneshyari.com)