Why the P-Value Alone is Not Enough: The Impact of Reporting Confidence Intervals in the Plastic Surgery Literature

Mona T. Al-Taha, MD Candidate, BBA; Osama A. Samargandi, MD, MHSc; Kit Moran, MD Candidate; Sarah A. Al Youha, MD, PhD; Michael Bezuhly, MD, MSc, SM
Division of Plastic and Reconstructive Surgery, Faculty of Medicine, Dalhousie University, Halifax, NS, Canada

PURPOSE: The p-value is one of the most utilized descriptors in statistical analysis; however, when reported in isolation, it does not convey the effect size of a treatment[1,2]. The reporting of confidence intervals is an essential adjunct to determine the clinical value of a treatment, as it permits an assessment of the effect size[1,3]. The objective of this study was to assess the reporting of confidence intervals in clinical trials within the plastic surgery literature.

METHODS: The seven highest impact plastic surgery journals of different domains were screened using MEDLINE for clinical trials in the years 2006, 2009, 2012, and 2015. Studies were randomized based on a predetermined sample size. Various characteristics including the Jadad quality score, statistical significance of the study findings, year of publication, journal impact factor, and participation of a methodologist were documented and their influence on the use of confidence intervals was examined.

RESULTS: Two independent reviewers analyzed 135 articles. There was substantial inter-rater agreement (kappa=0.78). Although, 86% of the studies reported a p-value, only 27% reported the confidence intervals. The quality of the studies had a median Jadad score of 2 out of 5 (IQR 0-3.75). Bivariate analysis revealed that a higher Jadad score (p=0.023) and inclusion of a research methodologist (p=0.002) were associated with the reporting of confidence intervals. Multivariate analysis revealed similar findings, while journal impact factor, year of publication and statistical significance were not correlated with confidence interval reporting.

CONCLUSION: Confidence intervals are under-reported in the plastic surgery literature. The main reason for reporting confidence intervals is that they focus on effect size and statistical significance of results, whereas p-values do not reveal all the information needed to interpret study findings. To improve the reporting quality of clinical trials, results should always include the confidence intervals to avoid the misinterpretation of the effect size of a statistically significant result.

REFERENCES:

1. Fethney J. Statistical and clinical significance, and how to use confidence intervals to help interpret both. *Aust Crit Care*. 2010;23(2):93-97. doi:10.1016/j.aucc.2010.03.001.

2. Davies HT, Crombie IK. What are confidence intervals and p-values? *Stat What is ? Ser*. 2009;(April):1-8. www.whatisseries.co.uk.

3. Vavken P, Heinrich KM, Koppelhuber C, Rois S, Dorotka R. The use of confidence intervals in reporting orthopaedic research findings. *Clin Orthop Relat Res*. 2009;467(12):3334-3339. doi:10.1007/s11999-009-0817-7.