



With Adequate Power Comes Great Responsibility

Dmitry Tumin¹

Received: 4 November 2022 / Accepted: 16 November 2022 / Published online: 22 November 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Dear Editors,

Herrera-Carrillo et al. highlight concerns regarding statistical power in pediatric cardiology randomized controlled trials (RCTs) [1]. The authors conducted a systematic review of RCTs with a primary endpoint, and checked if the sample size was adequate to attain 80% power for rejecting the null hypothesis at $p < 0.05$, given the results presented in each article. This post hoc power analysis should be differentiated from an a priori power analysis that might be conducted during trial design, which aims to determine whether a study will be adequately powered for an expected effect size, regardless of what the data ultimately show. On post hoc power analysis, only 45% of pediatric cardiology RCTs attained a power of 80%, leading the authors to call for reducing target power to 60%, relaxing the p value threshold for statistical significance, or both [1].

Recently, Bababekov et al. performed a similar review of studies in the surgical literature, finding a median post hoc power of just 16% [2]. Bababekov et al. advanced a similar critique of 80% as an “unreasonable power threshold,” [2] but their approach and conclusion were widely criticized, most notably by Althouse [3]. The root of this criticism is the distinction between a priori and post hoc power analysis. Whereas a priori power analyses require an additional assumption about an expected or reasonable effect size, post hoc power analyses do not include such information, and are composed of the same ingredients as the statistical tests that are already reported in RCTs. For example, group subsample sizes, means, and standard deviations can be used to calculate, at the reader’s discretion, the p value of an independent t test; the 95% confidence interval around the difference in means; or the post hoc power.

Apart from duplicating information already presented in other forms, post hoc power calculations distort the intended meaning of “power,” which properly refers to the chance that a null hypothesis would be rejected in the future. A reader considering the purchase of a lottery ticket might well ask what are the chances of winning (a priori power), but a lottery winner has no need to speculate on the chances that they have won (post hoc power). While a priori power analysis is important during RCT design [4], standards for trial methodology should not be swayed by analyses of post hoc power.

Author Contributions DT conceptualized and drafted the manuscript.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Herrera-Carrillo FE, Patel R, Flores S, Villarreal EG, Farias JS, Loomba RS (2022) Randomized controlled trials in pediatric cardiology: a power struggle? *Pediatr Cardiol*. <https://doi.org/10.1007/s00246-022-03039-z>
2. Bababekov YJ, Hung YC, Hsu YT, Udelsman BV, Mueller JL, Lin HY, Stapleton SM, Chang DC (2019) Is the power threshold of 0.8 applicable to surgical science?-Empowering the underpowered study. *J Surg Res* 241:235–239
3. Althouse AD (2021) Post hoc power: not empowering, just misleading. *J Surg Res* 259:A3–A6
4. Mascha EJ, Vetter TR (2018) Significance, errors, power, and sample size: the blocking and tackling of statistics. *Anesth Analg* 126(2):691–698

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

✉ Dmitry Tumin
tumind18@ecu.edu

¹ Department of Pediatrics, Brody School of Medicine at East Carolina University, 600 Moye Boulevard, Greenville, NC 27834, USA