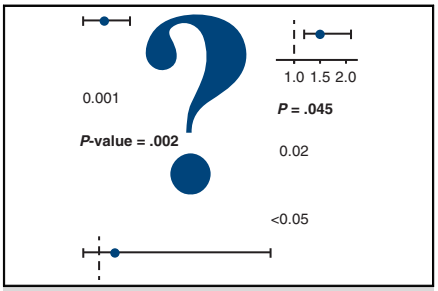
See Article page 1373 in the April 2021 issue.

Commentary: Too much ado about *P* value

Umberto Benedetto, MD, PhD, and
Arnaldo Dimagli, MD



P values are far from being the source of thorough information that we believe.

CENTRAL MESSAGE

P values do not provide clinically meaningful interpretation.

For decades, research results have been strictly dichotomized into statistically significant or not significant; that is, whether or not the *P* value was below .05. This method has been commonly used as a mark of the soundness of a study and used often as the major criterion on which the decision to submit research for publication is based.

An evolving discussion of new statistical strategies to fill the power vacuum created by the extensive overuse of the *P* value is being debated.[1,2] A lot of effort is still needed to establish a new tool—or a plethora of new statistical strategies each suited to a different research question. Hitherto, only a few journals, like the *New England Journal of Medicine*, have been imposing new restricting rules about reporting the *P* value.[3]

Visintainer[4] offers valuable guidance to move beyond the robustly rooted misconduct of using statistical significance as a benchmark of the presence of an effect or an association. *P* values are far from being the source of thorough information that we believe. For example, they do not relate to the extent of the clinical effect and relay nothing about the confidence we can have about the results. Confidence intervals can help interpret results from a clinical perspective and provide more information regarding the precision of the estimate. Opposite to *P* values, which cannot be compared, confidence intervals can be compared regardless of their statistical significance, providing information on the consistency of treatment effects across studies and supporting or disproving original hypotheses.

However, confidence intervals may be mistakenly used if they are intended as an instrument to dichotomize conclusions and therefore used as proxy of *P* value. Moreover, confidence intervals rely on the same statistical assumptions applied to *P* values, and therefore can provide distorted results in case of violation.

Whether or not the relevance of a statistical association should only be based on *P* value is clearly not resolved. Cardiovascular research needs to rely upon robust and common rules of interpretability. Visintainer provides guidance that should be taken into account by researchers aiming to successfully submit their research for publication.

References

1. Halsey LG. The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? *Biol Lett*. 2019;15:20190174.
2. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. *Am Statistician*. 2016;70:129-33.
3. Harrington D, D'Agostino RB Sr, Gatsonis C, Hogan JW, Hunter DJ, Normand ST, et al. New guidelines for statistical reporting in the *Journal*. *N Engl J Med*. 2019; 381:285-6.
4. Visintainer P. Moving beyond significance testing: confidence intervals in clinical research. *J Thorac Cardiovasc Surg*. 2021;161:1373-6.

From the Bristol Heart Institute, University of Bristol, Bristol, United Kingdom.
Disclosure: Authors have nothing to disclose with regard to commercial support.
Received for publication Jan 30, 2020; accepted for publication Feb 3, 2020; available ahead of print March 13, 2020.
Address for reprints: Umberto Benedetto, MD, PhD, Bristol Royal Infirmary, Office Room 84, Level 7, Upper Maudlin St, BS2 8HW, Bristol, United Kingdom (E-mail: umberto.benedetto@bristol.ac.uk).
J Thorac Cardiovasc Surg 2022;164:e39
0022-5223/$36.00

https://doi.org/10.1016/j.jtcvs.2020.02.027