Review of: "Comments on “The roles, challenges, and merits of the p value” by Chén et al. (Patterns, 2023, 4(12), 100878)"

Jordi Cortés[1]

1 Statistics and Operational Research, Universidad Politécnica de Cataluna, Spain

Potential competing interests: No potential competing interests to declare.

Hening Huang's commentary critiques the article *"The roles, challenges, and merits of the p value"* by Chén et al. Huang argues that the original article fails to adequately capture the true meaning of the p value. He acknowledges some misconceptions about p values and contends that p values derived from null hypothesis significance testing (NHST) are not reliable probabilistic measures and are prone to misuse, such as "N-chasing". Huang further emphasizes that p values reflect differences at the sample mean level rather than measure differences at the individual level, and proposes Exceedance Probability (EP) as a more robust alternative that is independent of the sample size.

The commentary provides an insightful analysis of the limitations of p values, particularly in real-world applications. Some constructive suggestions to improve the commentary could be:

1) Probabilistic decision-making systems versus risk management. In the conclusion, the author mentions the need for probabilistic measures in “decision-making systems”. While this is a valid point, it would be beneficial to expand the statement to emphasize the importance of managing the risks associated with incorrect decisions. In the context of the Neyman-Pearson framework for hypothesis testing, decisions are made by controlling for two types of errors (Type I and Type II errors) to minimize the risk of false positives or false negatives.

2) Confidence intervals. It could be suggested to include confidence intervals in addition to p-values in research papers, as recommended by the CONSORT guideline [1]. Confidence intervals provide valuable information on the magnitude of the effect, as well as the uncertainty around that estimate. This addition could serve as a compromise for those authors who wish to retain p values.

3) ***Significance***. The expression "*has no practical significance*" could be ambiguous in the context that it is mentioned. It would be clearer to replace "*significance*" with "*relevance*," making the sentence less prone to confusion, as "significance" could be misinterpreted in a statistical sense.

[1] Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG; CONSORT. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2012;10(1):28-55. doi: 10.1016/j.ijsu.2011.10.001. Epub 2011 Oct 12. PMID: 22036893.