

A call for greater power in an era of publishing negative results

Anna L Oberhofer¹, Robert P Lennon²

¹Department of Family Medicine, Naval Hospital Jacksonville, Jacksonville FL, USA, ²Department of Family Medicine Naval Hospital Okinawa, Okinawa, Japan

Corresponding author:

Robert P Lennon
ATTN Bush Clinic
Naval Hospital Okinawa
PSC 482, FPO AP 96362

Robert.lennon@med.navy.mil
Tel.: + 080 4881 0766

Historically, medical journals published only positive results from experimentation. With the growth of electronic media, many journals are now able to publish negative results as well. Because most medical experiments are designed to have different thresholds of significance for positive and negative results, this can lead readers to misunderstand the level of significance of a published negative result. We propose a technical shift – setting α equal to β – to avoid this potential for misunderstanding.

Key words: Statistics, Beta statistic, Power negative results.

To the Editor,

When searching for truths in nature, we risk committing two types of error. A Type I error is rejecting a hypothesis that should have been accepted. A Type II error is accepting a hypothesis that should have been rejected (1, 2). Alpha is the probability of making a Type I error while beta is the probability of making a Type II error. The medical community generally accepts as true those results with a five percent or less chance of being random – our threshold for claiming statistical significance. That is, after the probability of the null hypothesis is calculated, we accept as “true” results in which $\alpha \leq 0.05$. We accept an error rate for significant negative outcomes of 20%. That is, $\beta \leq 0.2$. The statement, “an observed difference if $\alpha > 0.05$ fails to show significance” means that we are unable to assert that there is a 95% chance that an observed difference is due to intervention and not chance. We are tempted to publish a negative result indicating that the interven-

tion does not work. However, although the observed difference is statistically a lack of difference, the appropriate evaluation of the negative relies on the β statistic. The lack of difference if $\beta \leq 0.2$ means that one can only assert that there is an 80% chance that the lack of difference is not due to chance.

A more rigorous threshold of acceptance of a positive value than a rejection of a negative value was acceptable when only studies with positive results were published. In an age of publishing negative results, this is problematic. To reject a hypothesis that fails to meet 95% certainty in the positive is to accept the opposite hypothesis with only 80% certainty – a historically unacceptable level of surety. To avoid rejecting true hypotheses using current conventions, the negative result must be interpreted differently than the positive result. There is no compelling reason for this.

One could evaluate both positive and negative results to 95% certainty by setting both α and β to ≤ 0.05 . Only then is one's re-

jection of the “not-different” (which ought to be measured by the β statistic) as sure as one’s acceptance of the different (which is measured by the α statistic). This will increase the difficulty of performing quality studies; as β is lowered from 20% to 5%, power decreases. This may be offset by increasing enrollment or identifying interventions with greater effect size, but may limit researchers’ ability to identify significant differences in situations in which attaining such power is not feasible. However, this will also increase the value of our studies. When α equals β , we are equally sure of our results regardless of outcome, and we can accept negative results with the confidence with which we accept positive results.

The views presented are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the United States Government.

References

1. Neyman J, Pearson ES. The testing of statistical hypotheses in relation to probabilities a priori. Joint Statistical Papers. Cambridge: University Press; 1967.
2. Pearson ES, Neyman J. On the Problem of Two Samples. Joint Statistical Papers. Cambridge: University Press; 1967.

Authors’ contributions: Conception and design: ALO, RPL; Acquisition, analysis and interpretation of data’s: ALO, RPL, Drafting the article: ALO, RPL; Revising it critically for important intellectual content: ALO, RPL.

Conflict of interest: The authors declare that they have no conflict of interest.

Received: 3 June 2014; Accepted: 3 December 2014

Copyright © 2014 by Academy of Sciences and Arts of Bosnia and Herzegovina. E-mail for permission to publish: amabih@anubih.ba

Reprints will not be available from the author.