Wasserstein, Schirm, and Lazar argue against the use of the phrase "statistically significant" and its variants because they believe it has led to misinterpretation and misuse in scientific research. Specifically, they highlight that the binary thinking (significant/non-significant) oversimplifies the complexity of statistical evidence, encourages dichotomous thinking, and often leads to the dismissal of potentially important findings that do not meet the arbitrary p < 0.05 threshold. This practice can result in the loss of nuanced understanding of data and can foster publication bias towards results that achieve statistical significance.

Instead, they suggest several alternatives to improve statistical practice:

1. \*\*Reporting Effect Sizes\*\*: They advocate for the reporting of effect sizes, which provide a measure of the magnitude of an effect, offering a more informative view of the results than a binary significance test.

2. \*\*Confidence Intervals\*\*: Using confidence intervals to convey the precision of an estimate. This approach helps in understanding the range within which the true effect size might lie, promoting a more comprehensive interpretation of the data.

3. \*\*Emphasizing the Context of the Data\*\*: They recommend that researchers should contextualize their findings within the broader scope of the study, considering prior evidence, study design, and practical significance. This holistic approach helps in interpreting results beyond mere statistical thresholds.

These suggestions aim to foster a more nuanced and transparent reporting of statistical results, moving away from the rigid reliance on p-values.