



Statistical significance should be abandoned, say scientists

Harriet Pike

Cambridge

Scientists writing in *Nature* have called for the concept of statistical significance to be abandoned, a move backed by the American Statistical Association.

Valentin Amrhein, a professor of zoology at the University of Basel, Switzerland, and colleagues argued that researchers often apply P values incorrectly and use them to draw unhelpful conclusions. “The false belief that crossing the threshold of statistical significance is enough to show that a result is ‘real’ has led scientists and journal editors to privilege such results, thereby distorting the literature,” they wrote in *Nature*.¹

The authors said that researchers should never conclude that there is “no difference” or “no association” between two results just because a P value is larger than a threshold such as 0.05 or because a confidence interval includes zero. They also want to see an end to the common misconception that two studies are in conflict because one has a statistically significant result and the other does not. “These errors waste research efforts and misinform policy decisions,” they said.

The authors, who gathered more than 800 signatories to their piece, emphasised that they were not calling for a ban on P values. “We are calling for a stop to the use of P values in the conventional, dichotomous way—to decide whether a result refutes or supports a scientific hypothesis,” they wrote.

In an accompanying editorial, *Nature* editors acknowledged that misuse of statistical measures could mean that “some analyses are biased, some false positives are overhyped and some genuine effects are overlooked.”² But the publishing giant is not ready to ditch statistical significance. “There are reasonable viewpoints on all sides,” the editorial said. “*Nature* is not seeking to change how it considers statistical analysis in evaluation of papers at this time.”

The *Nature* articles coincide with a special issue of *American Statistician*, an official journal of the American Statistical Association, that presents 43 papers on the topic. The papers set out how researchers should use statistics in their work—with an emphasis on accepting uncertainty and being thoughtful, open, and modest.

The authors of an editorial introducing the special issue said that accepting uncertainty would prompt researchers to seek better measures, more sensitive designs, and larger samples, increasing the rigor of research.³ Thoughtfulness, in turn, would

help researchers look ahead to prospective outcomes and consider multiple data analysis techniques.

The ASA series advocates openness in the development and presentation of research, including through adherence to good practice (for example, public preregistration of methods, transparency, and completeness in reporting). The authors called on researchers to be modest “about the role of statistical inference in scientific inference” and by encouraging others to reproduce their work.

David Colquhoun, a professor of pharmacology at University College London, was one of the more than 800 *Nature* signatories and the author of one of the 43 papers published in *American Statistician*.

“Statisticians have been talking about the deficiencies of P values as evidence for at least 70 years, but nobody listens,” he told *The BMJ*.

Colquhoun offered two reasons for this. “Both authors and editors are rewarded for publishing false positives—that’s why we have a reproducibility crisis, and statisticians can’t agree on what to do about it,” he said.

In his paper Colquhoun proposed that P values be supplemented with a number that indicates the risk that a positive result is a false positive.⁴ “Several approaches suggest that if you observe a P value of just below 0.05, this false positive risk is at least 20-30%,” he said. “No wonder that there’s a crisis.”

Not all scientists are ready to abandon the concept of statistical significance, however. Deborah Mayo, a philosopher of science at Virginia Tech, told *The BMJ*, “Banning the word ‘significance’ may well free researchers from being held accountable when they downplay negative results and search the data for impressive looking subgroups.”

- 1 Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature* 2019;567:305-7. 10.1038/d41586-019-00857-9 30894741
- 2 Nature. It’s time to talk about ditching statistical significance. *Nature* 2019;567:283. 10.1038/d41586-019-00874-8 30894740
- 3 Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond “P<0.05”. *Am Stat* 2019;73:1-19. 10.1080/00031305.2019.1583913.
- 4 Colquhoun D. The false positive risk: a proposal concerning what to do About p-values. *Am Stat* 2019;73:192-201. 10.1080/00031305.2018.1529622.

Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to <http://group.bmj.com/group/rights-licensing/permissions>