I think the enemy here is the arbitrarily chosen threshold. An easy solution is to report the p-value in results. In the case in which a p-value calculation and thresholding is an intermediate step, generalize downstream analysis wherever possible to take into account the minimum significance level required for a traditional statistical analysis to conclude significance (AKA p-value). Or make a non-arbirary choice of thresholding: perhaps one that aligns with some estimation of the signal to noise estimation in a given context.

The comment in Nature made by Leek and Peng in *Pvalues are just the tip of the iceberg* makes the point that p-values very often in practice are not the weakest link in the chain. They point out that drastic downstream changes in results can more easily be achieved by changing upstream experimental designs. To that I rebut. Sure, significance can be toggled by the specific choices made during data cleaning or the permutations of confounding factors for which one adjusts. Garbage in garbage out; I agree. But, the fact that there are potentially other causes to the increasingly arbitrarily results published in recent years is simply a red herring to the very specific dialogue surrounding the utility of the p-value. Just because there are broken cogs in our machine doesn’t mean we should ignore the cog that has been building up rust over the years. We need to dust off the p-value just are readily as we should standardize the handling of batch effects. Because a single loose bolt and we can be assured that our machine will produce garbage no matter the input.

With other points made by the co-authors, I can agree. (In spite of the slippery slop argument made when they claim, “deregulating statistical significance opens the door to even more ways to game statistics.”). I don’t think deregulation is necessarily the best approach. And I do think more rigorous statistical education can make a difference. The authors are a bit alarmist with their claim that “people need to stop arguing about p-values, and prevent the [upstream factors in data analysis pipelines] from sinking science”. I am, however, intrigued about their suggestions for an analogue of evidence-based medicine—‘evidence-based data analysis’.

Error:

Technical (pcr bias+ gc bias + indicator function for threshold of missing values)+ Biological

Signal:

Organism + Tissue(environment) + cell type + cell state