American Psychologist
2016, Vol. 71, No. 3, 238–239

0003-066X/16/$12.00 http://dx.doi.org/10.1037/a0040195

COMMENT

Preregister If You Want To

Don A. Moore
University of California, Berkeley

Prespecification of confirmatory hypothesis tests is a useful tool that makes our statistical tests informative. On the other hand, selectively reporting studies, measures, or statistical tests renders the probability of false positives higher than the *p* values would imply. The bad news is that it is usually difficult to tell how much higher the probability is. Fortunately, there are enormous opportunities to improve the quality of our science by preregistering our research plans. Preregistration is a highly distinctive strength that should increase our faith in the veracity and replicability of a research result.

Keywords: hypothesis testing, selective reporting, pre-registration, open science, false positives

In 2004, the news broke that pharmaceutical manufacturer GlaxoSmithKline (GSK) had hidden evidence about risks of its drugs. GSK had conducted experimental drug trials that it had chosen not to publish or publicize. The resulting scandal produced a $3 billion penalty from the U.S. Food and Drug Administration and underscored the risk of selective reporting. In an attempt to make selective reporting visible, the International Committee of Medical Journal Editors began requiring that clinical drug trials be preregistered starting July 1, 2005.

Yet despite the widely acknowledged problem of selective reporting (Franco, Malhotra, & Simonovits, 2014; Rosenthal, 1979; Simonsohn, Nelson, & Simmons, 2013), many scientists resist preregistering their own studies. This resistance is largely driven by three concerns. First, skeptics fear preregistration would constrain their creativity and the serendipity of scientific discovery. In this, they are partly right. Sometimes after the data come in, you think of better ways to test your theory. But if you choose the right way to test your theory after you see the data, you should worry about being misled by chance. Preregistration helps protect you from being fooled by false positives. Do you want to run an exploratory test? You can register it as such; or do not bother registering it at all. Take what you learn from that test, design a better study, and preregister that.

Second, skeptics worry that preregistration brings added scrutiny to their research reporting. In this, they fail to consider the increased credibility that preregistration brings. If you preregister your plan to analyze your data using, say, a negative binomial regression, you will eliminate concerns that you selected it post hoc. Your scholarly colleagues will reward you by lending greater credence to your research results if you preregister your predictions.

Third, preregistration adds another effortful step in the research process already overburdened with requirements from institutional review boards and grant funders. Part of this resistance is driven by an overestimation of the administrative burden of preregistration. Medical journals require preregistrations to include 20 components.[1] Most of these are not particularly laborious, but you are not even bound to these 20. You can preregister whatever you want. What should you preregister? Preregister anything you want to get credit for having been able to anticipate—any confirmatory test.[2]

Would preregistration rule out the possibility of reporting other tests? Certainly not. But it clarifies the distinction between confirmatory and exploratory tests. Exploratory tests are useful. By all means, try different ways of testing your theory or specifying your analysis. But selectively reporting those results that are most sensible, clear, and statistically significant increases the risk of false positives. The number of tests you run (or could run) should reduce the statistical significance threshold. If you conduct more

Correspondence concerning this article should be addressed to Don A. Moore, Haas School of Business, 2220 Piedmont Avenue, University of California, Berkeley, Berkeley, CA 94720. E-mail: dm@berkeley.edu

[1] Read about them here: https://prsinfo.clinicaltrials.gov/trainTrainer/WHO-ICMJE-ClinTrialsgov-Cross-Ref.pdf.

[2] It's easy! You can do it through the Open Science Framework (http://osf.io/), AsPredicted (http://aspredicted.org/), the American Economic Association (https://www.socialscienceregistry.org/), or even the U.S. National Institutes of Health (https://clinicaltrials.gov/).

statistical tests than you report, the cutoff for statistical significance should be lower than .05, but it is often hard to tell how much more stringent it should be.

The value of preregistration is nicely illustrated by the study that set out to document the benefits of chocolate (Bohannon, Koch, Homm, & Driehaus, 2015). A total of 15 participants were assigned to one of three diets for 3 weeks: a low-carb diet, a low-carb diet plus a daily 1.5-ounce bar of dark chocolate, or no dietary changes (the control group). When the results came in, the researchers sifted through the 18 dependent measures to find any result that could plausibly be interpreted as favoring the chocolate group. Lo and behold, the chocolate group lost weight. The results were speedily published in a for-profit journal and eagerly reported in press outlets around the world.

If the researchers had preregistered this entire plan, their selective reporting would have been obvious. Without preregistration, it is difficult to establish whether $p < .05$ is the right threshold for statistical significance. If you are only conducting one confirmatory hypothesis test that you had prespecified, then $p < .05$ keeps the probability of a false positive at 5%. But the chance of false positives rises quickly with the number of possible tests. With 18 tests to choose from, the probability of finding at least one false positive result rises to 60%. And if there are a couple of ways to operationalize each variable, the probability of a false positive would be closer to 80%.

Would it be possible to "game" the system by preregistering everything? Yes, but preregistering 18 tests and publishing one ought not be impressive. Specifying the number of tests ex ante makes it easy to tell how impressed we should be by the result ex post. The timing is obviously crucial here. Could researchers simply do a little hypothesizing after the results are known (Kerr, 1998) and pretend they preregistered that? That would be fraud. There are already penalties in place to punish scientific fraud, but that is not why it is rare. It is rare because scientists share the goal of the scientific enterprise: the truth. And preregistration serves this goal.

In 2014, *Psychological Science* started offering badges for preregistration, posting data, and posting materials. The preregistration badge has been used least frequently—in about 1% of papers published. That is too bad, because preregistration ought to contribute enormously to our confidence in a published result. Indeed, it is essential to interpreting the results of statistical tests. Preregistration makes *p* values useful for their intended purpose—estimating the probability of false positives. Without preregistration, the risk of a false positive is higher than the reported *p* value, but it is difficult to tell how much higher.

References

Bohannon, J., Koch, D., Homm, P., & Driehaus, A. (2015). Chocolate with high cocoa content as a weight-loss accelerator. *International Archives of Medicine, 8,* 1–8. http://dx.doi.org/10.3823/1654

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science, 345,* 1502–1505. http://dx.doi.org/10.1126/science.1255484

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2,* 196–217. http://dx.doi.org/10.1207/s15327957pspr0203_4

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2013). P-Curve: A key to the file drawer. *Journal of Experimental Psychology: General, 143,* 534–547. http://dx.doi.org/10.1037/a0033242

Received December 2, 2015
Accepted December 31, 2015 ■