



COVID-19 is an emerging, rapidly evolving situation.

- Get the latest public health information from CDC »
- Get the latest research information from NIH »
- NIH staff guidance on coronavirus (NIH Only) »

October 15, 2018

Tips for Communicating Statistical Significance



Regina Nuzzo, Ph.D.

By Regina Nuzzo, Ph.D.

Professor, Department of Science, Technology, and Mathematics, Gallaudet University

“So what exactly is a p-value?”

That’s the number-one question I get from non-scientists when they hear my specialty is communicating p-values and other statistics. (The second question is usually something like, “Who cares?” followed in short order by, “You know, I really hated stats in school.”)

P-values are the gatekeepers of statistical significance. When a press release or article claims results to be statistically significant, it usually means researchers used a certain procedure to calculate a p-value, and that number was smaller than 0.05.

This post isn’t about teaching Stats 101. It’s about communicating statistical significance, p-values, and their accompanying results to a non-statistician audience. It’s about making sure that your communication — press release, blog post, journalism article, and the like — is as clear, accurate, helpful, and engaging to readers as can be.

Basics to remember

What’s most important to keep in mind? That we use p-values to alert us to surprising data results, not to give a final answer on anything. (Or at least that’s what we *should* be doing). And that results can get flagged as “statistically surprising” with a small p-value for a number of reasons:

1. **There was a fluke.** Something unusual happened in the data just by chance.
2. **Something was violated.** By this I mean there was a mismatch between *what was actually done* in the data analysis and *what needed to be done* for the p-value to be a valid indicator. One little-known requirement, for example, is that the data analysis be planned before looking at the data. Another is that all analyses and results be presented, no matter the outcome. Yes, these seem like strange, nit-picky rules, but they’re part of the deal when using p-values. A small p-value might simply be a sign that data analysis rules were broken.
3. **There was a real but tiny relationship**, so tiny that we shouldn’t really care about it. A large trial can detect a true effect that is too small to matter at all, but a p-value will still flag it as being surprising.
4. **There was a relationship that is worth more study.** There’s more to be done. Can the result be replicated? Is the effect big enough to matter? How does it relate to other studies?

Or any combination of the above.

Most of the time we jump to #4 as an explanation for a surprising result. That’s because we’re human and we like solid explanations for things. But #2 is the dark horse that we can’t ignore. In fact, many failures to replicate can be traced back to this one. And #1 is always there, more often than we think. Remember, meeting someone who shares your birthday is a rare fluke, and yet it happens all the time.

What to do

Report an effect size in whatever form makes sense, but be sure to [put it into context](#). Remember, statistical significance alone doesn’t tell us how big or how important an effect is.

Consider banishing “significant” and “confidence” from your vocabulary when writing for a general audience. Yes, statisticians use these words in their line of work. But their precise technical meanings don’t necessarily line up with the colloquial. Better to consider these terms to be jargon and treat them like any other specialized technical terms.

Take the time to explain why the claimed relationship may be the real deal (explanation #4 above) and not a rare fluke, trifle, or misfire in the bunch of statistical requirements (#1, #2, and #3). For example, do these results agree with those in animal models and different human populations?

What to avoid

Trying to explain to a general audience what a p-value means is fraught with peril. Many have tried and failed. An accurate summary isn't very understandable, and most understandable summaries are just plain wrong! It's not necessarily the best use of either your word count or energy.

- Avoid: A p-value of 0.05 means the probability that these results were due to random chance is five percent.
- Avoid: The chance of this result being a false positive is five percent.
- Avoid: The probability that the researchers' data validated their hypothesis just by chance is five percent.
- Better, but probably not worth it: If only random chance were at work here, and if many technical details about the data collection and analysis are accurate, then results at least as extreme as those found in this study would happen only one time out of 20. (Not very satisfying, is it? I told you.)

Bottom line

For the lay public, statistical significance is more about scientific-community behavior than about the math. Getting a small enough p-value allows researchers in our current scientific system to publicize their results with credibility even while allowing for the possibility of error. But most of the useful information for your audience is contained *behind* the statistical significance. It's your job as a communicator to pull that forward.

Finally, one way I like to think of statistical significance is relating it to the online system of up votes and down votes for sorting web comments. It can definitely be a useful way to direct our attention by flagging potentially interesting and important information. But by no means should it be the final arbiter of either truth or substance.

Subscribe

NIH...Turning Discovery Into Health®

National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland 20892

U.S. Department of Health and Human Services