

10: What Have We Wrought?

I've painted a grim picture. But anyone can pick out small details in published studies and produce a tremendous list of errors. Do these problems matter?

Well, yes. I wouldn't have written this otherwise.

John Ioannidis's famous article "Why Most Published Research Findings are False"³¹ was grounded in mathematical concerns rather than an empirical test of research results. If most research articles have poor statistical power – and [they do](#) – while researchers have the freedom to choose among multitudes of analyses methods to get favorable results – and [they do](#) – when most tested hypotheses are false and most true hypotheses correspond to very small effects, we are mathematically determined to get a multitude of false positives.

But if you want empiricism, you can have it, courtesy of John Ioannidis and Jonathan Schoenfeld. They studied the question "Is everything we eat associated with cancer?"^{51[1]} After choosing fifty common ingredients out of a cookbook, they set out to find studies linking them to cancer rates – and found 216 studies on forty different ingredients. Of course, most of the studies disagreed with each other. Most ingredients had multiple studies claiming they increased *and* decreased the risk of getting cancer. Most of the statistical evidence was weak, and meta-analyses usually showed much smaller effects on cancer rates than the original studies.

Of course, being contradicted by follow-up studies and meta-analyses doesn't prevent a paper from being cited as though it were true. Even effects which have been contradicted by massive follow-up trials with unequivocal results are frequently cited five or ten years later, with scientists apparently not noticing that the results are false.⁵⁵ Of course, new findings get widely publicized in the press, while contradictions and corrections are hardly ever mentioned.²³ You can hardly blame the scientists for not keeping up.

Let's not forget the merely biased results. Poor reporting standards in medical journals mean studies testing new treatments for schizophrenia can neglect to include the scale they used to evaluate symptoms – a handy source of bias, as trials using unpublished scales tend to produce better results than those using previously validated tests.⁴⁰ Other medical studies simply [omit particular results](#) if they're not favorable or interesting, biasing subsequent meta-analyses to only include positive results. A third of meta-analyses are estimated to suffer from this problem.³⁴

Another review compared meta-analyses to subsequent large randomized controlled trials, considered the gold standard in medicine. In over a third of cases, the randomized trial's outcome did not correspond well to the meta-analysis.³⁹ Other comparisons of meta-analyses to subsequent research found that most results were inflated, with perhaps a fifth representing false positives.⁴⁵

Let's not forget the multitude of physical science papers which misuse confidence intervals.³⁷ Or the peer-reviewed psychology paper allegedly providing evidence for psychic powers, on the basis of uncontrolled multiple comparisons in exploratory studies.⁵⁸ Unsurprisingly, results failed to be replicated – by scientists who appear not to have calculated the statistical power of their tests.²⁰

We have a problem. Let's work on fixing it.

Footnotes

[1] An important part of the ongoing [Oncological Ontology](#) project to categorize everything into two categories: that which cures cancer and that which causes it.

This page titled [10: What Have We Wrought?](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Alex Reinhart](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.