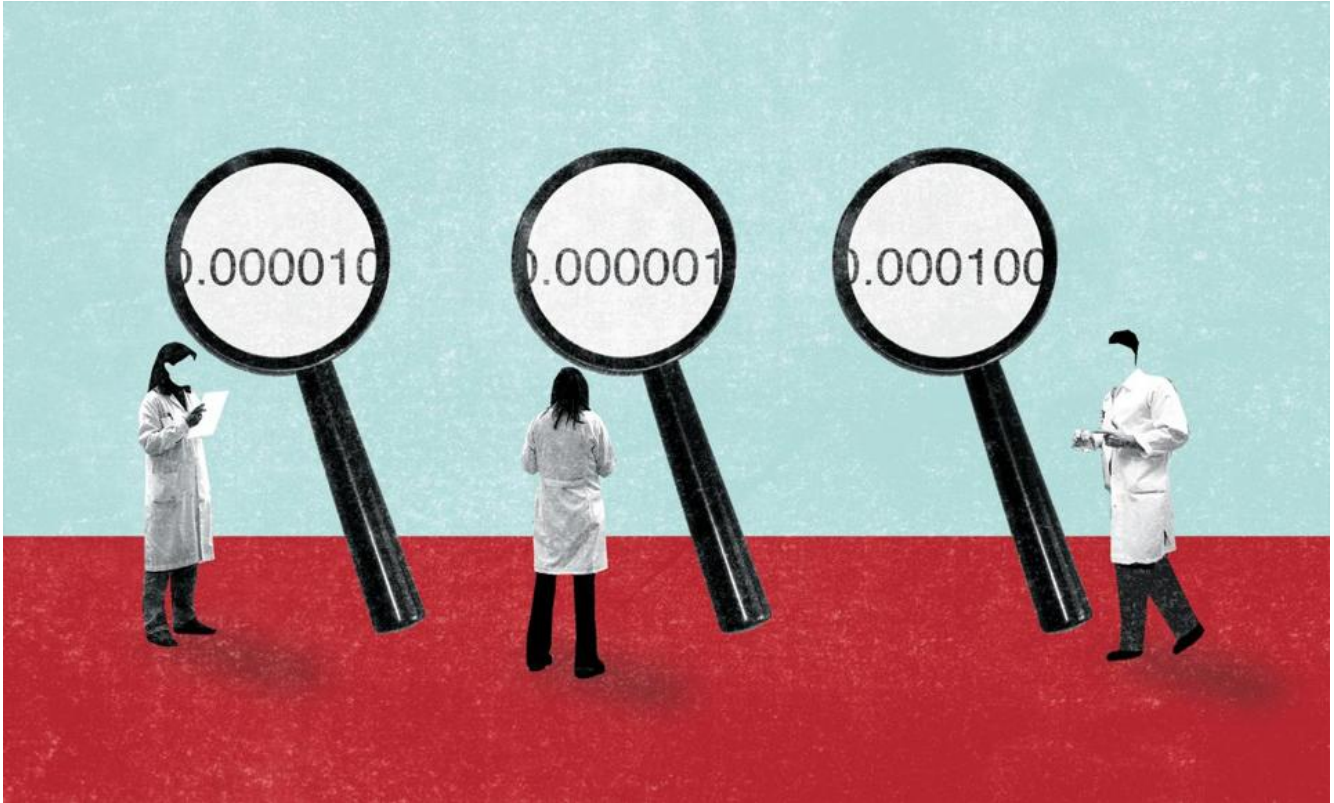


In science, irreproducible research is a quiet crisis



PATRIC SANDRI FOR THE BOSTON GLOBE

By Carolyn Johnson | GLOBE STAFF MARCH 19, 2015

EVEN WHEN NO one's done anything obviously wrong, scientific experiments sometimes yield results that turn out to be incorrect. When Doug Melton's team at Harvard University discovered betatrophin, a hormone that could trigger the pancreas to make beta cells lost in diabetes, their 2013 paper was touted as a breakthrough. But when they redid the experiment and increased the number of animals, the original result didn't quite hold up. The hormone's effect was far weaker than first reported.

As so often happens, the biology at work was more complex than it originally seemed. Melton is continuing a long list of experiments to understand how betatrophin works. He vows to publish the results, whether they point to a diabetes therapy or not.

CONTINUE READING BELOW ▼

To many in the scientific community, this was an example of how science self-corrects. It was Melton's lab, along with an outside group, that identified the problems in the earlier work. Yet the case also exemplifies a broader problem in the research world. The rush to celebrate "eureka" moments often overshadows a rather mundane activity on which science depends: repetition. Any finding needs to be "reproducible" — confirmed in other labs — if it is to matter.

But talk to a scientist long enough, and you'll probably hear a story like this: An intriguing new discovery was reported in a research journal. Maybe it was a biologist describing a new Achilles' heel in cancer cells, a psychologist's profound insight into human behavior, or an astronomer's finding about the first moments of the universe. The scientist read about the finding and tried to confirm it in her own lab, but the experiment just didn't come out the same.

Evidence of a quiet crisis in science is mounting. A growing chorus of researchers worry that far too many findings in the top research journals can't be replicated. "There's a whole groundswell of awareness that a lot of biomedical research is not as strongly predictive as you think it would be," said Dr. Kevin Staley, an epilepsy researcher at Massachusetts General Hospital. "People eventually become aware because there's a wake of silence after a false positive result," he added. The same is true in every field of science, from neuroscience to stem cells.

CONTINUE READING IT BELOW ▼



Evan Horowitz: Studies show many studies are false

There are good scientific reasons to think that lots of published research is actually false.

Ideally, science builds on and corrects itself. In practice, the incentives facing scientists can hamper the process. It's more exciting and advantageous to publish a new therapeutic approach for a disease than to revisit a past discovery. Yet unless researchers point out the limitations of one another's work, the scientific literature can end up cluttered with results that are partially or, in some cases, not at all true.

Recently, researchers and the US government alike have sought to assess how much research is irreproducible — and why — and are looking for systematic ways to retest experiments that make headlines but yield no further progress.

THOUGH ACADEMIC FRAUD exists, most irreproducible findings simply result from the trial and error inherent in the scientific process. Experiments can be finicky. A “discovery” can turn out to be a statistical aberration; an exciting result might later be explained by something researchers hadn't considered. Experiments may not be properly set up. Scientists who genuinely want to solve problems may be blinded by subtle bias. Experiments may depend on animals or ingredients that have major limitations.



PATRIC SANDRI FOR THE BOSTON GLOBE

Sometimes, by random chance, an experiment will give an exciting positive result, the way a lucky golfer occasionally hits a hole in one.

Attempts to quantify how much research is wrong have been sobering. Scientists at Bayer HealthCare reported that when they tried to reproduce 67 published discoveries in oncology, women's health, and cardiovascular disease, only a quarter of the time were their in-house results completely consistent with what had been reported. A scientist at Amgen disclosed that, despite concerted effort, the company had successfully repeated only six of 53 major cancer findings. At the ALS Therapy Development Institute in Cambridge, researchers found that clinical trials of drugs that failed could have been avoided, saving time and money, if earlier experiments on animals had been carefully repeated. Last year, neuroscience researchers tried to repeat five studies that showed 17 links between certain brain structures and behaviors — for example, a finding that people with more gray matter in particular brain regions also have more Facebook friends. The researchers could not replicate any of the connections and found evidence for what scientists call the “null hypothesis” — the idea that there was no relationship.

Scientists move ahead in their careers by publishing papers in top journals. But when they conduct experiments that don't seem to show anything — when a technique fails or a hypothesis is not confirmed — they often publish nothing at all, even though the failure may be deeply informative. For their part, the journals that referee top research will rarely publish papers on experiments that didn't work.

Some researchers believe that as the size of the biomedical research enterprise has grown and competition for limited federal funding and slots in leading journals has intensified, the pressures have increased to draw sweeping conclusions from evidence that may not fully support it.

“The waste, and the thinking that ‘I have to get my paper into these [top] journals,’ just corrupts and poisons the way people do science itself,” said Michael Eisen, a biologist at the University of California at Berkeley. “It leads to people making their work sound sexy even at the expense of its veracity.”

Drug companies trying to build off research insights to create drugs often gather a good deal of information about which research bears out. In theory, their scientists could share failures, so everyone would know to avoid dead ends. But their knowledge often ends up in a file drawer. When Dr. William Sellers, global head of oncology at Novartis

Institutes for BioMedical Research in Cambridge, came to the pharmaceutical industry from a lab at Dana-Farber Cancer Institute in 2005, he was idealistic: His researchers would find what didn't work and let people know.

That's not, for the most part, what happened. "It just turns out there's not an appetite to publish data that are invalidating to an original paper," Sellers said. "It turns out to be a lot of work — and the work will not be published in as high a profile journal as the other work was."

And even correcting problems by publishing papers is no guarantee that the record will be set straight. Sellers found that a set of 60 iconic cancer cell lines kept by the federal government for research had several labeling problems. There were duplicates in the set. A commonly used breast cancer cell line turned out to be the deadly skin cancer melanoma instead.

A 2007 paper by another group described the issue, declaring it "a loss for breast cancer, but a boon for melanoma research," but Sellers continued to see papers published incorrectly describing them as breast cancer cells. That experience has led Sellers to focus on the part of irreproducibility that seems most tractable: trying to make sure that the ingredients that scientists use in their experiments are vetted enough that they can be confident about what they are.

SELLERS HAS BECOME one of a group of experts trying to talk about the problem publicly and brainstorm about fixes. He imagines a crowd-sourced website like TripAdvisor, where scientists could rate cell lines. In a recent paper in the journal Nature, other researchers called for standardization of the antibodies that are used to track proteins in biomedical research. In the United States alone, the authors estimated, \$350 million is lost per year due to antibodies that don't work as advertised.

The irreproducibility problem is being recognized at the highest levels; the White House's Office of Science and Technology Policy mentioned it last summer in a request for public comments on innovation strategy. "Given recent evidence of the irreproducibility of a surprising number of published scientific findings, how can the federal government leverage its role as a significant funder of scientific research to most effectively address the problem?" the document said.

Some proposals focus on the journals that publish research, the gatekeepers of new scientific knowledge. Dozens of journals signed a pledge with the National Institutes of Health to adopt measures that will help enable scientists to repeat experiments easily — allowing them to determine more easily which findings are reproducible and which are not.

But for now, each field is coming to terms with the crisis in its own way. Staley, the MGH epilepsy researcher, and some of his colleagues have proposed an epilepsy-focused journal that would publish attempts to repeat experiments — documenting failures as carefully as successes. Christophe Bernard, editor of a new journal called eNeuro, said part of the journal's mission is to publish negative results. When a study that can't be replicated remains in the literature, he argues, it can lead an entire field astray.

Whether the new focus on irreproducibility will solve the problem is unknown; Harvard cell biologist Bjorn Olsen, has seen attention to the issue rise and fall. As a junior researcher, Olsen spent a year trying to repeat an experiment he had seen published in a prestigious journal. The technique in question supposedly could be used to track molecules within cells. At wits' end, he contacted the scientist who had led the research to ask what he was doing wrong.

As Olsen recalled, the scientist seemed surprised to hear from him and said that the technique was very difficult and that he would not recommend that people do it. It became clear to Olsen that it was something that the original research team had tried and failed to do many times, and then had gone and published a paper based on the few times it worked.

“When papers are written and data are presented in public it looks like everything is just perfect, and that is not what science is,” Olsen said. “Science is an imperfect human activity that we try to do as best we can.”

Olsen and a colleague started the Journal of Negative Results in BioMedicine nearly 15 years ago. It's a small journal that publishes about two papers a month. In the long run, the goal is to bring this conversation into the open — and ultimately put his own journal out of business.

Carolyn Y. Johnson can be reached at cjohnson@globe.com. Follow her on Twitter [@carolynjohnson](https://twitter.com/carolynjohnson).

Related:

- [Evan Horowitz: Studies show many studies are false](#)
- [Q&A: Russian science is amazing. So why hasn't it taken over the world?](#)
- [What does Cambridge sewage say about residents? MIT plans to find out](#)
- [What 'urban physics' could tell us about how cities work](#)
- [2011: How crowdsourcing is changing science](#)

Correction: An earlier version of this story incorrectly stated the name of the ALS Therapy Development Institute.

Get Today's Headlines from the Globe in your inbox:

Enter your email address

SIGN UP

[Privacy Policy](#)

SHOW 10 COMMENTS