

News Buzz Life Quizzes Videos More ▾



Scientists Are Hoarding Data And It's Ruining Medical Research

Major flaws in two massive trials of deworming pills show the importance of sharing data — which most scientists don't do.

posted on Jul. 22, 2015, at 4:18 p.m.



Ben Goldacre
BuzzFeed Contributor



A girl cries while receiving a deworming pill in Managua, Nicaragua. *Hector Retamal / AFP / Getty Images / Via gettyimages.com*

We like to imagine that science is a world of clean answers, with priestly personnel in white coats, emitting perfect outputs, from glass and metal buildings full of blinking lights.

The reality is a mess. A [collection of papers published](#) on Wednesday — on one of the most commonly used medical treatments in the world — show just how bad things have become. But they also give hope.

The papers are about deworming pills that kill parasites in the gut, at extremely low cost. In developing countries, battles over the usefulness of these drugs have become so contentious that some people call them “The Worm Wars.”

Every year hundreds of millions of children in the developing world are given deworming tablets, whether they have worms or not. It's easy to see why this intervention is so appealing: In principle, children's health, survival, and school performance can improve with a simple pill, given just once a year, costing only 2 cents.

This approach was endorsed by the World Health Organization. A meeting of eminent researchers, including four Nobel Prize winners, listed the medications in the top four most cost-effective interventions worldwide. As part of a publicity stunt at the Davos Summit in 2008, Cherie Blair, wife of British Prime Minister Tony Blair, [reportedly](#) chased world leaders around the room pretending to be a giant intestinal worm.

Nobody doubts that treating people who have worms is a good idea. More problematic is the idea of treating whole populations of schoolchildren to improve health and school performance.

This “deworm everybody” approach has been driven by a single, hugely influential [trial](#) published in 2004 by two economists, Edward Miguel and Michael Kremer. This trial, done in Kenya, found that deworming whole schools improved children's health, school performance, and school attendance. What's more, these benefits apparently extended to children in schools several miles away, even when those children didn't get any deworming tablets (presumably, people assumed, by interrupting worm transmission from one child to the next).

A decade later, in 2013, these two economists did something that very few researchers have ever done. They handed over their entire dataset to independent researchers on the other side of the world, so that their analyses could be checked in public. What happened next has every right to kick through a revolution in science and medicine.



Hookworms, *Ancylostoma caninum*, attached to the intestinal mucosa. CDC / Via en.wikipedia.org

The independent researchers — epidemiologists from the London School of Hygiene and Tropical Medicine, where I've worked, on and off, since 2009 — got a huge pile of files: the original data describing the trial's 30,000 participants, but also explanatory notes and, most crucially, the original computer programs used to analyse the data.

The researchers wrote their own new programs from scratch to check that they always got the same results for individual numbers in the results tables.

They discovered, first off, that a phenomenal amount of information was simply missing from the original data: For 21% of children, there was no age recorded anywhere (and for more than 10%, no information on their gender). This can happen, sure, but the extent of these gaps wasn't covered explicitly in the original paper.

When the replication team tried to rerun the original analyses, things got worse. The original paper had 10 tables giving the various results of the trial. Most of these had errors. Some were trivial — eight tables had rounding errors, where 0.745 was incorrectly truncated to 0.74 rather than 0.75, and so on.

But other errors were fundamental. When analyzing data, scientists generally run a test for something called "statistical significance." This offers some insight into how likely it is that a difference between two groups could have occurred by

chance.

Traditionally, findings are considered interesting if they meet a statistical significance of $p < .05$. Translated very roughly into everyday language, that means: If you ran this same trial over and over again, taking your participants from the same population, you'd always expect to see one group doing a bit better than another, simply from the play of chance, because of other factors in their lives. But you'd only see a difference as big as this one here, purely by chance, on about 1 occasion in every 20.

In the deworming paper, several findings that were labelled as statistically significant actually were not. One observation — that the pills reduced anemia — was labelled as having a p value of $< .05$, when in fact it was only 0.194.

This is no small error. And it was just 1 of 11 findings that had their level of statistical significance mislabelled.

Then things get worse. When the replication team began to check the economists' original code, they found that there were frank errors in the instructions to the statistics package. The wrong commands had been typed into the program, and because of this, the wrong answers had come out.

One of the trial's biggest findings was the supposed benefit for nearby schools, as well as the single treated school. But when the reanalysis team tried to reproduce this result, they found that the lines in the program to calculate which schools fell into this "deworming nearby" category had erroneously excluded the majority of schools. Once that key error was corrected, the benefit for neighboring schools — one of the key messages of the trial, something trumpeted at meetings and by fundraisers around the world — effectively disappeared.

And so it went on. The improvement in school attendance ceased to be statistically significant, depending on how the data was examined. There was insufficient information to replicate the analyses.

If your eyes are glazing over, then that is perhaps fair enough. Maybe you're thinking, quite understandably, "The details are someone else's problem. I get it. The research was crap."

And here, you'd be wrong.



Students of Loltulelei primary school in Samburu, Kenya. AFP / Tony Karumba / Getty Images / Via gettyimages.com

The deworming study was an excellent piece of research. Or at least, there is no reason to think that it was any better or worse than the average clinical trial. It was expensive, it was difficult, and it produced some useful results.

It also deserves something of a free pass, because it had hugely beneficial impact beyond just one research question. This trial was pivotal, in that it helped create an [entire movement](#) of people doing proper randomized trials — the most “fair test” of whether an intervention works — throughout the entire community of development work. Before that, the field was in a kind of [dark ages](#), blown on the winds of expert opinion and whim.

So some of the results of this individual trial shifted, under closer examination, and that is definitely problematic. But fundamentally there is only one thing different about this deworming trial and the rest of social science and medicine: Miguel and Kremer had the decency, generosity, strength of character, and intellectual confidence to let someone else peer under the bonnet.

This kind of statistical replication is almost vanishingly rare. A recent [study](#) set out to find all well-documented cases in which the raw data from a randomized trial had been reanalysed. It found just 37, out of many thousands. What’s more, only five were conducted by entirely independent researchers, people not involved in the original trial.

These reanalyses were more than mere academic fun and games. The ultimate outcomes of the trials changed, with terrifying frequency: One-third of them were so different that the take-home message of the trial shifted.

This matters. Medical trials aren't conducted out of an abstract philosophical interest, for the intellectual benefit of some rarefied class in ivory towers. Researchers do trials as a [service](#), to find out what works, because they intend to act on the results. It matters that trials get an answer that is not just accurate, but also reliable.

So here we have an odd situation. Independent reanalysis can improve the results of clinical trials, and help us not go down blind alleys, or give the wrong treatment to the wrong people. It's pretty cheap, compared to the phenomenal administrative cost of conducting a trial. And it spots problems at an alarmingly high rate.

And yet, this kind of independent check is almost never done. Why not? Partly, it's resources. But more than that, when people do request raw data, all too often the original researchers duck, dive, or simply ignore requests.

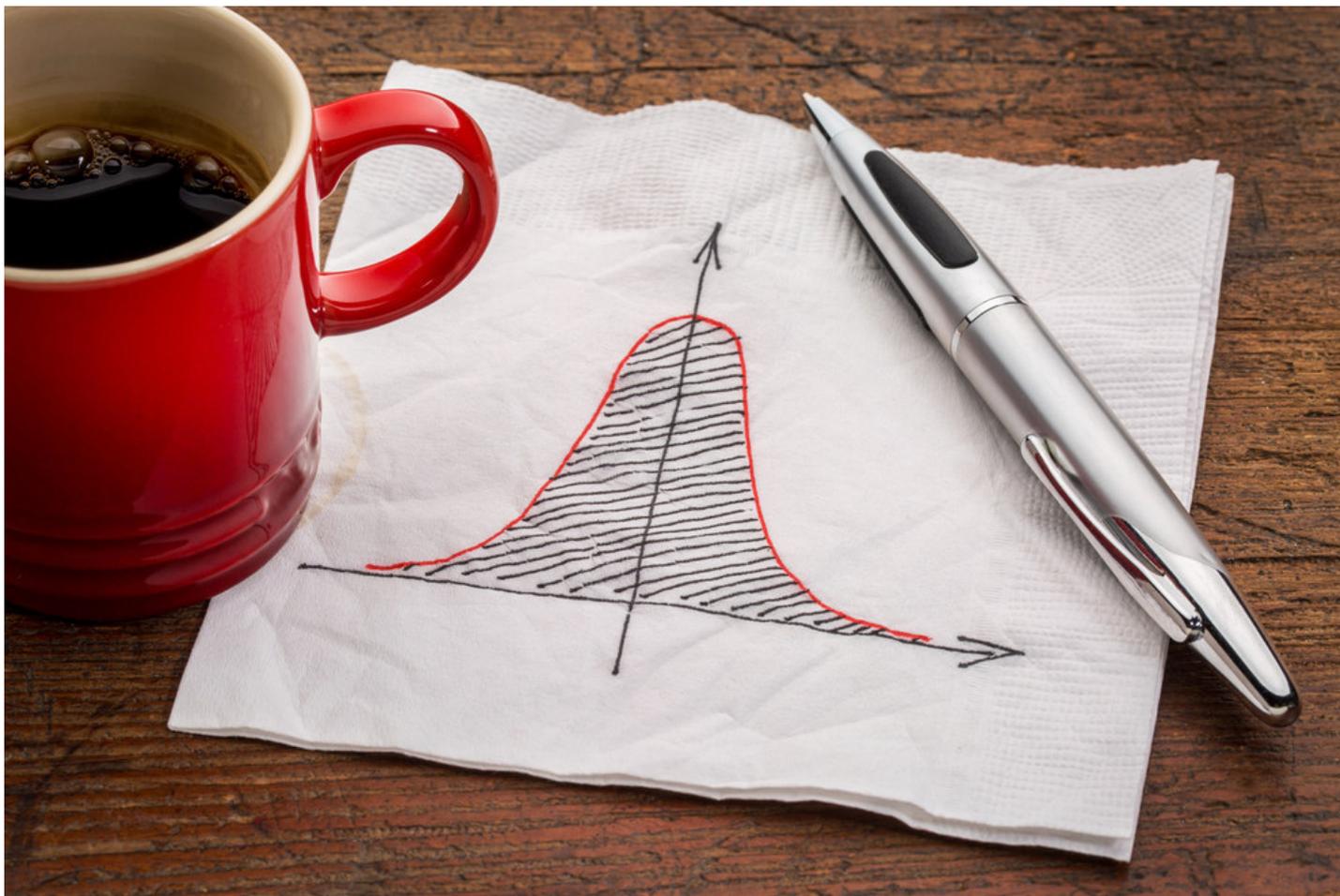
Take statins, for example: the single most commonly prescribed class of drugs in the developed world. That makes them a very good indicator for all the problems in medicine, because if we can't get things [right on statins](#), there's little hope for the more obscure treatments. In general, looking at all the evidence collected so far, it's very likely that statins do [more good than harm](#). Still, there are legitimate concerns about the extent of that benefit, as well as the frequency of side effects.

The best way to assess the evidence of statins is to combine the raw data from all trials. This week the *British Medical Journal* published an [editorial](#) explaining that they have requested the original patient data from 32 major statins trials to do just that. Despite follow-up calls and emails, only seven teams have deigned to respond.

Conducting a trial, and then refusing to let anyone see the data, is like claiming you've flown a spaceship to Pluto, but refusing to let anyone see the photos.

That would be laughable. But the justifications for secrecy from drug companies and researchers are hardly any more plausible. Sometimes, they play on fear and authority. What about the idiots, they say? The anti-vaccination conspiracy theorists, and the journalists who love them: Won't they use this information to create mischief, by picking endless holes in perfectly good data?

Well, this week, NASA flew a spaceship past Pluto, and the [truthers](#) appeared right on cue to say it was all a fake. Idiots gave them coverage. And then...the sky did not fall in. NASA was not defunded. The claims were debunked, by bloggers who enjoyed the sport. And everyone tweeted that [clip](#) of astronaut Buzz Aldrin punching a conspiracy theorist in the face.



Marekulasz / Getty Images / Via thinkstockphotos.com

That's the point. Science thrives on public debate (ideally without punching). Any attempt to preserve the authority of medicine or science by hiding information is doomed to fail, for one reason: The bizarre secrecy that we've come to accept in medicine is, in reality, the polar opposite of science.

The Royal Society is one of the oldest scientific institutions in the world. Across the door of its headquarters in London, its motto is carved in stone: *nullius in verba*. This is Latin, inevitably, meaning "on the word of no one."

Because scientists don't care if you wear a white coat, or how many letters you have after your name. We want to know what you did in your experiment, in detail. We want to know what you measured, and what the results were. Then we want to pore over details, to see if we can find any holes, and decide if we agree with your conclusions. If scientists have any legitimate authority in the world, it flows entirely from this transparency about the methods and results of our experiments.

And that's where things start to break down even more for the deworming movement.

The Cochrane Collaboration is a global nonprofit organization of 37,000 academics producing systematic reviews on pretty much every topic in medicine. In 2012, they [did one on deworming](#).

Looking at 41 deworming trials, the Cochrane review found little to no evidence that these pills boost cognition, school performance, anemia, nutritional status, and so on.

Is that interesting? I've no position on the Worm Wars. But from my perspective, as a researcher who is obsessively committed to good methods, there is one very interesting feature in this Cochrane review: missing data.

The single biggest randomized trial ever conducted, in the entire history of medicine, was on deworming. The trial ran in India, and it recruited 2 million participants. This was an epic, vast piece of work. It ended in 2006, and then... nothing. The full results were finally [published](#) in 2013 (too late for the last Cochrane review). For seven years, aid workers, economists, and development workers were left in the dark, carrying on spending on deworming medications without the full data they needed to make informed decisions.

Amazingly, this is business as usual. Two years ago I published a [book](#) on problems in medicine. Front and center in this howl was "publication bias," the problem of clinical trial results being routinely and legally withheld from doctors, researchers, and patients. The best available evidence — from [dozens of studies](#) chasing results for completed trials — shows that [around half](#) of all clinical trials fail to report their results. The same is true of industry trials, and academic trials. What's more, trials with positive results are about twice as likely to post results, so we see a biased half of the literature.

This is a cancer at the core of evidence-based medicine. When half the evidence is withheld, doctors and patients [cannot make informed decisions](#) about which treatment is best. When I wrote about this, various people from the pharmaceutical industry cropped up to claim that the problem was all in the past. So I befriended some campaigners, we assembled a group of senior academics, and started the [AllTrials.net](#) campaign with one clear message: "All trials must be registered, with their full methods and results reported."

Dozens of academic studies had been published on the issue, and that alone clearly wasn't enough. So we started collecting signatures, and we now have more than 85,000 supporters. At the same time we sought out institutional support. Eighty patient groups signed up in the first month, with hundreds more since then. Some of the biggest research funders, and even government bodies, have now signed up.

This week we're [announcing](#) support from a group of 85 pension funds and asset managers, representing more than 3.5 trillion euros in funds, who will be asking the pharma companies they invest in to make plans to ensure that all trials — past, present, and future — report their results properly. Next week, after two years of activity in Europe, we launch our campaign in the U.S.

So, as a single-issue obsessive, preoccupied with missing data, do I regard the team behind this huge deworming trial in India — conducted on 2 million people, but then withheld from decision makers for seven years — to be malicious, corrupt, and incompetent?

No.



A Bangladeshi medical assistant administers a vitamin A capsule to a child at a health center in Dhaka. AFP / Munir Uz Zaman / Getty Images / Via gettyimages.com

In its conception, this vast deworming trial was everything a randomized controlled trial should be.

First, it was largely free from the biases and design flaws that can sometimes make trials unfair tests of which treatment is best. These biases wouldn't matter if the benefit you were trying to detect was huge. But in the case of deworming, where the benefit is modest, that true signal could easily be either drowned out, or exaggerated, by design flaws.

The second huge advantage of this trial is that it was very big. The smaller the benefit you're trying to detect, the larger the number of participants you need. A clinical trial on whether [parachutes saved lives](#) would reach pretty firm conclusions after just a few participants had hit the ground. This deworming trial had 2 million participants, to detect a tiny but important potential benefit on death rates.

But perhaps most importantly, this trial was incredibly cheap. In the 1990s, Oxford researchers Richard Doll and Richard Peto won a few hundred thousand pounds in a prize awarded to them — not unreasonably — for their work on whether smoking causes cancer.

The award was cash for their own pockets, but Doll and Peto decided instead to see if they could do something outrageously impactful with such a small sum of money. With Shally Awasthi, an Indian researcher, they decided to run the biggest trial ever conducted, on deworming, vitamin A, and death.

Peto is a social acquaintance of mine, and overall, I think this idea was pretty impressive. Why? Because the trial cost about 20 cents per participant. For context, the pharma industry typically quotes the cost of a randomized trial as being around \$10,000 per participant, and trials routinely cost tens of millions of dollars.

So this deworming trial was huge, cheap, and could answer important questions. Why did it take so long to publish?

The trial actually looked at two different interventions: vitamin A and deworming. In 2007, when the vitamin A results had been analysed, the team organized a meeting in Oxford to present the results.

The World Health Organization had been claiming that this treatment would reduce mortality by 25%, an enormous amount. The results of this 2 million participant trial found some benefit from vitamin A, but couldn't support that huge claim. Just like deworming, the benefits of vitamin A are a bitterly polarized topic: The seductiveness of simple pills, as a solution to complex problems in the developing world, is perhaps overwhelming.

Participants at the meeting lashed out at the Oxford researchers, saying their data was suspect.

"We were afraid," Peto [told the *BMJ*](#) in 2013, "that if any trivial defects were found in the data, they would be misused to undermine the credibility of the study. So we did a lot more data checks to try to weed out any duplicated records. We didn't want to be in a position where people could pick holes."

This data checking and analysis would ultimately take more than a year of person-time. But with the trial run on such an extraordinarily tight budget, they didn't have the spare resources to do it right away. It took seven years.

It's hard to see this spectacular publication delay in the same light as a trial withheld by a zealot, or by a company with money to lose from transparency. It still has a grave impact on the practice and reputation of medicine. But as with all problems, it pays to understand them before you try to fix them.



Heap Of Drugs *Wanjajacob / Getty Images / Via thinkstockphotos.com*

And here is where I think the threads come together. The press releases on the reanalysis of the Miguel and Kremer deworming trial in Kenya will go live this week. Somewhere, I'm sure, people will attack or mock them for their errors. One way or another, I can't believe they won't feel bruised by the reanalysis. And that is where we have gone wrong. It's not just naive to expect that all research will be perfectly free from errors, it's actively harmful.

There is a replication crisis throughout research. When subjected to independent scrutiny, results routinely fail to stand up. We are starting to accept that there will always be glitches and flaws. Slowly, as a consequence, the culture of science is shifting beneath everyone's feet to recognise this reality, work with it, and create structural changes or funding models to improve it.

The reanalysis of the Miguel and Kremer deworming trial was funded by 3ie, the International Initiative for Impact Evaluation, a huge funding agency for development research. But it is only the second of a [great many more](#) similar reanalyses that they are funding, as a new matter of principle.

Meanwhile, the [Reproducibility Project](#) has done independent replications of 100 studies in psychology (and preliminary results suggest that only 39 of the 100 key findings could be replicated by independent researchers). Some drug companies have [come together](#) to share some of their raw clinical trial data, on request, to independent researchers; and the Institute of Medicine has given the idea a [positive kick](#) forward. The *BMJ* have said they [will only publish](#) trials that commit to sharing data on request.

This is the beginning of a massive, and long overdue, culture shift. Trials, in particular, have acted like imperial city-states, sucking up resources from all around them. One trial can cost tens of millions of dollars, but nobody has ever

received such a vast sum to work on the structural problems surrounding all trials: to run an [audit and name](#) the researchers, institutions, and sponsors who have most frequently failed to report trial results; to throw modest funding at trials that have finished, but run out of resources before finalizing their analysis; to coordinate a series of cheap reanalyses to check the results of dozens, or hundreds of trials; and so on.

There is no doubt that such work is needed, and would be hugely cost-effective. There is no doubt that there will be a long period of discomfort, and the whole culture of science will have to shift. But it is only through this transparent, open, networked approach that we can finally realise the fundamental principles of science.

That's why the saga of these two deworming trials should be regarded as a pivotal point in history. These core problems in science and medicine — missing data, and the need for reproducibility checks — are now instantiated by the single biggest trial ever conducted, on one of the most commonly used treatments in the world; and by Miguel and Kremer's deworming study, the pivotal trial for an entire movement.

It is time to change. *Nullius in verba*. On the word of nobody. Show me the data.

Tagged:science news, clinical trials, deworming, statistics

[Advertise](#) [Jobs](#) [Mobile](#) [Newsletter](#)

 [US Edition](#) 

[About](#) [Press](#) [RSS](#) [Privacy](#) [User Terms](#) [Ad Choices](#) [Help](#) [Contact](#)

© 2015 BuzzFeed, Inc