

(/)

Follow me on: (https://www.twitter.com/cblatts)



(https://www.facebook.com/chrisblattmanblog)



(https://plus.google.com/102532515971215413968)

(/subscribe-by-email/)

(/feed/)

RESEARCH (HTTP://CHRISBLATTMAN.COM/RESEARCH/) BLOG (/)

**TEACHING** 

(HTTP://CHRISBLATTMAN.COM/TEACHING/)

ABOUT (HTTP://CHRISBLATTMAN.COM/ABOUT/)

## ← Previous post

(http://chrisblattman.com/2015/07/22/statistician-

neal-beck-just-justified-my-longstanding-hatred-

and-loathing-of-logit/)

## Dear journalists and policymakers: What you need to know about the **Worm Wars**

23 JUL 2015

One of my favorite science writers, Ben Goldacre, enters the so-called Worm Wars (http://www.buzzfeed.com/bengoldacre/deworming-trials). He's not alone, with a flurry of new articles today (https://www.google.com/search?g=google+news&ie=utf-8&oe=utf-8#tbm=nws&g=deworming). The question is simple: is a deworming pill that costs just a few cents one of the most potent anti-poverty interventions of our time?

Below is the picture from Goldacre's post. I assume Buzzfeed editors chose it. It's a nice illustration that nothing you will read in this debate is dispassionate. Everyone wants one thing: your clicks (and retweets, and likes, and citations). Most writers sincerely want the truth too. Sadly the two are not always compatible.



In brief: Ted Miguel and Michael Kremer are Berkeley and Harvard economists who ran the original deworming study

(http://cega.berkeley.edu/assets/cega\_research\_projects/1/Identifying-Impacts-on-Education-and-Health-in-the-Presence-of-Treatment-Externalities.pdf) that showed big effects of the medicine on school attendance in Kenya—one of the few to attempt to measure such impacts. That study ignited the impact evaluation movement in international development, especially through their students (like me). It also ignited a movement to deworm the world (http://www.evidenceaction.org/dewormtheworld). This is a big claim, worth investigating. Calum Davey led the team who did a replication (http://ije.oxfordjournals.org/content/early/2015/07/21/ije.dyv128).

I know this study. In fact, as a first year graduate student I spent a summer working for Miguel and Kremer designing their long term follow up survey. Relationships are incestuous on all sides of the deworming debate, so you can hardly call me an impartial judge. Nonetheless, bear with me as I try.

I haven't paid much attention to the deworming world for more than a decade. So I spent last night and this morning reading as much as I could. There's an overwhelming amount to process, but I've drawn a few early conclusions.

The bottom line is this: both sides exaggerate, but the errors and issues with the replication seem so great that it looks to me more like attention-seeking than dispassionate science. I was never convinced that we should deworm the world. There are clearly serious problems with the Miguel-Kremer study. But, to be quite frank, you have throw so much crazy sh\*t at Miguel-Kremer to make the result go away that I believe the result even more than when I started.

Backing up, we should remember that most scientific studies don't stand up to scrutiny very well. Most are utterly wrong.

This was Ben Goldacre's overarching point, and I couldn't agree more. But reading the details, I think he may have accidentally chosen an example of the opposite problem: a love of the witch hunt. A bias among authors to have a flashy result, and a bias among journals to publish it.

If you throw enough at a study, the results will eventually get imprecise enough that you can't draw a strong conclusion. This is why, ultimately, the body of evidence matters. In a single study, the amount of assault a result can take is a good indication of its quality.

By this metric, the Miguel-Kremer deworming result is actually impressive. Davey and company find some real errors, but most don't change the results. They have to throw a whole lot at the results to put them just beyond the realm of statistical significance.

From what I can tell, you have to do three or four things at once:

- 1. You have to divide it into two smaller experiments. The medicine was phased in over time. Some people received medicine in year 1 and some in year two year 2. If you split years 1 and 2 into two separate experiments, the precision goes down. Naturally. But the rationale for doing that is completely weird. I've never seen any study do this before.
- 2. You also have to ignore the fact that the disease could pass to other people. If you give medicine to person A, it can affect the health of person B nearby is it means you don't pass on the infection. That means if you compare A and B, you're biased towards underestimating the effect of the medicine. Many health studies (amazingly) make this mistake. So does the replication.
- 3. You have to care about school rather than pupil effects. Some schools are small (50 pupils) and some are large (1300 pupils). As it happens, the medicine has a bigger effect in bigger schools, probably because the medicine stops the disease from spreading. If you ignore this, and take school averages rather than pupil averages, you will get a lower estimate of how well the medicine works. Again I'm not sure why you'd do that.

4. You have to recode the groups to put people who didn't get the medicine into the group that was supposed to get the mechanism. This looks to me like a mistaken understanding of the experiment by Davey and team.

Reading Miguel-Kremer's response in the same journal (<a href="http://ije.oxfordjournals.org/content/early/2015/07/21/ije.dyv129.full">http://ije.oxfordjournals.org/content/early/2015/07/21/ije.dyv129.full</a>) (none of the journalists I've read seem to have read or cite this) here's the amazing thing: Just doing two or three of these things is not enough to make the result go away. It looks to me like you have to do three or all four. In particular, if you do everything except split the experiment in two, everything holds. Most of the "debunking" rides on splitting the experiment in two.

[Read World Bank economist Berk Ozler's similar points in greater depth.
(http://blogs.worldbank.org/impactevaluations/worm-wars-review-reanalysis-miguel-and-kremer-s-deworming-study)]

There was a lot to absorb, so I invite other views. But my quick read is this: Davey and team's 1 through 4 are useful checks on the data, but rather weird choices. A reasonable scientist might choose one of them. Maybe, and probably erroneously. But all four? Something is amiss.

This is all rather technical, so it's not surprising that the journalists writing on the debate don't understand. But if you're not a statistician, here's what should make you suspicious: Every single choice ticks in the direction of making the effects of the medicine less impressive. This is either correct, coincidental, convenient, or conniving.

Since I see no argument for correct, someone should interrogate the other three. Instead most of the journalism has accepted the article at face value.

To me there's a simple and sad explanation why: Whether it's a sensational photo, a sensational result, or a sensational take down of a seminal paper, everyone has incentives to exaggerate. This whole episode strikes me as a sorry day for science.

## Share:

- **■** Email (http://chrisblattman.com/2015/07/23/dear-journalists-and-policymakers-what-you-need-to-know-about-the-worm-wars/?share=email&nb=1)
- Twitter 112 (http://chrisblattman.com/2015/07/23/dear-journalists-and-policymakers-what-you-need-to-know-about-the-worm-wars/?share=twitter&nb=1)
- **f** Facebook (http://chrisblattman.com/2015/07/23/dear-journalists-and-policymakers-what-you-need-to-know-about-the-worm-wars/?share=facebook&nb=1)
- © Reddit (http://chrisblattman.com/2015/07/23/dear-journalists-and-policymakers-what-you-need-to-know-about-the-worm-wars/?share=reddit&nb=1)