A conversation with Alex Aiken, Calum Davey, Paul Garner, and David Taylor-Robinson, February 26, 2016

Participants

London School of Hygiene and Tropical Medicine

- Alexander Aiken – Clinical Lecturer, Department of Infectious Disease Epidemiology
- Calum Davey – Research Fellow, Department of Social and Environmental Health Research

Liverpool School of Tropical Medicine

- Paul Garner – Coordinating Editor, Cochrane Infectious Diseases Group; Professor, Liverpool School of Tropical Medicine

University of Liverpool

- David Taylor-Robinson – Senior Clinical Lecturer in Public Health, Department of Public Health and Policy, University of Liverpool

GiveWell

- Alexander Berger – Program Officer in U.S. Policy, Open Philanthropy Project
- Josh Rosenberg – Senior Research Analyst, GiveWell

Note: These notes were compiled by GiveWell and give an overview of the major points made by Dr. Aiken, Mr. Davey, Dr. Garner, and Dr. Taylor-Robinson.

Summary

GiveWell spoke with Dr. Aiken and Mr. Davey (who worked on the Kremer replication); and Dr. Garner and Dr. Taylor-Robinson (who are responsible for the Cochrane review) about the evidence base for mass deworming. Topics included the recent replications of Miguel & Kremer 2004 (a key study on deworming), the updated Cochrane review on mass deworming, the costs and counterfactual for mass deworming, and relevant future research work.

Recent replications of Miguel & Kremer 2004

Questions about the attendance effect

Out of the many outcomes measured by Miguel & Kremer 2004, the one that has been most used as rationale for deworming, is a difference in school attendance in the intervention and control groups. In its blog post (http://blog.givewell.org/2015/07/24/new-deworming-reanalyses-and-cochrane-review/) on replication studies led by Dr. Aiken and Mr. Davey (http://ije.oxfordjournals.org/content/44/5/1572, http://ije.oxfordjournals.org/content/44/5/1581), GiveWell interprets the replications as not challenging this attendance effect of the complex intervention

that includes deworming. In fact, Dr. Aiken and Mr. Davey believe that the issues described below do cast doubt on the existence of such an effect.

Lack of intermediate explanatory outcomes

No evidence was found in the replication studies of Miguel & Kremer 2004 for any short-term health benefits for anaemia or nutritional status from deworming. In addition, the most recent Cochrane review on mass deworming noted reasonably strong evidence that deworming has no effect on a number of health and educational indicators in follow up periods for up to and after 1 year including nutritional status, average hemoglobin level, cognitive and school performance. This was observed in several studies. The lack of an observed intermediate stage of improvement makes the existence of a school attendance effect less plausible.

Missing data

The replication noted that a number of planned visits to schools to check attendance during Miguel & Kremer 2004 were not conducted, and a relatively high proportion of these missed visits were for children who were supposed to begin receiving deworming during the second year of the trial. This increases the risk of bias of results for that treatment arm.

Correlation between attendance rates and frequency of attendance measurements

In each control group in both years of the study, the replication noted a negative correlation between the amount of data each school provided and that school's attendance rates: schools with higher attendance rates provided fewer observations. In two of the three intervention groups, by contrast, there was a positive correlation: schools with higher attendance rates provided more observations. These unexpected correlations raise the concern about risk of bias in the study possibly accounting for observed differences.

Analysis of correlated outcomes

In Miguel & Kremer 2004, the intervention was rolled out to the treatment group in two stages: one group of schoolchildren began receiving the intervention in the first year of the study, and another group began receiving the intervention in the second year. This type of study design is called a "stepped-wedge trial."

It may be methodologically inappropriate to use simple fixed-effect analysis to combine results from both years in stepped-wedge trials as was done in Miguel & Kremer 2004. To provide a check on using simple analysis methods, the replication calculated the effect of the intervention in the first and the second years of the trial individually, which amount to standard cluster-randomized trials. For this exercise, point estimates were more important than statistical significance or confidence intervals.

The first-year data gives an attendance effect odds ratio of about 1.2 (for an odds ratio, the further away from 1, the stronger the effect). The second-year data gives

an odds ratio of about 1.4. But the analysis combining the two years gives an odds ratio of about 1.9, a much stronger effect than was observed in either of the two years. This discrepancy suggests that the main combined-year analysis may be overestimating the magnitude of any observed attendance effect. This could be because the combining of these results uses both the randomized comparison and secular, non-random comparison over time. Preliminary follow-up research into this question indicates that standard analysis techniques can often give inaccurate results for stepped-wedge trials, and that new techniques may be needed for more reliable analysis.

Potential for a confounding behavior-change intervention

The intervention studied in Miguel & Kremer 2004 included a sanitation & hygiene education program in addition to mass deworming, meaning that any observed effect could potentially be due to this other component of the program. Some factors suggest that this could be the case:

- Aiken et al. 2015 finds essentially no difference in attendance rates between dewormed children versus children who were not dewormed but attended schools that received the intervention (and so still may have received the behavior change intervention). If the deworming were responsible for the attendance effect, some difference in these attendance rates would be expected, as it is unlikely that spillovers would be as large as the direct benefit of deworming.
- Macartan Humphreys's replication in his blog (http://www.columbia.edu/~mh2245/w/worms.html) noted essentially no effect for children in the second year of deworming. This suggests that the observed effect could be due to something that lasts for about a year, which is consistent with generally-observed effects of behavior change interventions.
- The Hawthorne effect occurs when study participants change their behavior in response to their awareness of being observed. It is possible that this could be responsible for some or all of the observed attendance effect, especially as treatment group participants may have interacted more frequently with researchers due to the behavior change component of the intervention.

Questions regarding "Worms at Work"

The Cochrane review of mass deworming for soil-transmitted helminths (http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD000371.pub6/abstract) excludes some additional surveys carried out on the Kenya and Uganda studies. These surveys examine differences in health and economic outcomes 5 to 10 years after the deworming was commenced, and they examine whether starting deworming a few years earlier in one group gave long term health benefits compared to children who started 2-3 years later. To address this concern, the Cochrane group has begun a separate review of long-term follow-up studies,

including “Worms at Work” (http://files.givewell.org/files/DWDA%202009/Interventions/Deworming/Worms-at-Work_2015-11-11-Final.pdf), the follow-up study to Miguel & Kremer 2004. This critical appraisal of these studies has not been carried out before and has been submitted for publication. Early observations examining these papers discussed during the call were:

- It is unintuitive that an additional 2.5 years of deworming treatment should lead to such a substantial effect on income ten years later.
- The analysis was problematic due to multiple subgroup analysis and multiple significance testing.

The call participants discussed multiple hypothesis testing without preregistration allows for selective reporting of statistically significant outcomes. This raises the risk that any reported outcomes are false results. There is no systematic way of accounting for this possibility; one must simply closely study the results.

GiveWell asked whether the multiple hypothesis testing has to do with confidence intervals: standard errors should be expanded when multiple hypothesis testing occurs. However, the London and Liverpool teams pointed out that normal techniques like Bonferroni correction are only applicable if the testing plan was pre-specified.

Costs and counterfactuals of mass drug administrations

Costs of deworming

Mass drug administrations for deworming can pose significant costs to the governments that organize them: in India, for example, the government must spend significant resources over the course of 9 to 12 months in organizing a mass deworming program for 200 million children, during which its capacity for other work may be limited. These costs may not be fully accounted for in cost-effectiveness estimates of deworming.

Professor Tim Allen of the London School of Economics could be a good source for better understanding the costs of deworming programs; he has experience with Uganda’s deworming efforts. GiveWell could also look at the costs to government of polio campaigns and other vertical health programs.

Counterfactual of mass deworming

Though most schoolchildren do not seem to benefit from mass deworming, some children with particularly heavy worm burdens are symptomatic and so would benefit. However, because these children show worms in their stool, and because deworming medicine is now widely available in developing countries (due to improvements in primary health systems), these children would likely still be treated in the absence of mass deworming programs.

Future relevant work

There are a number of research projects in progress that could be of interest to GiveWell:

- The Campbell Collaboration intends to publish its own systematic review of the mass deworming evidence base. This is completed and has been submitted for publication.
- Tim Allen, mentioned above, has written a history of the "worm wars."
- A couple of randomized controlled trials of deworming infants in Peru were recently published. These studies look for effects on short-term health and cognition and do not find an effect of the intervention.
- As mentioned above, a team (including Prof. Garner) is working on a systematic review of long-term follow-ups of mass deworming.
- Some preliminary research is underway on appropriate methods for analyzing stepped-wedge trials.

All GiveWell conversations are available at http://www.givewell.org/conversations