

ABDUL LATIF JAMEEL

Poverty Action Lab



TRANSLATING RESEARCH INTO ACTION



Case 4: Deworming in Kenya

Addressing threats to experimental integrity

This case study is based on Edward Miguel and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72(1): 159-217, 2004

J-PAL thanks the authors for allowing us to use their paper



Key Vocabulary

- 1. Phase-in Design:** a study design in which groups are individually phased into treatment over a period of time; groups which are scheduled to receive treatment later act as the comparison groups in earlier rounds.
- 2. Equivalence:** groups are identical on all baseline characteristics, both observable and unobservable. Ensured by randomization.
- 3. Attrition:** the process of individuals joining in or dropping out of either the treatment or comparison group over the course of the study.
- 4. Attrition Bias:** statistical bias which occurs when individuals systematically join in or drop out of either the treatment or the comparison group for reasons related to the treatment.
- 5. Partial Compliance:** individuals do not comply with their assignment (to treatment or comparison). Also termed "diffusion" or "contamination."
- 6. Intention to Treat:** the measured impact of a program that includes all data from participants in the groups to which they were randomized, regardless of whether they actually received the treatment. Intention-to-treat analysis prevents bias caused by the loss of participants, which may disrupt the baseline equivalence established by randomization and which may reflect non-adherence to the protocol.
- 7. Treatment on the Treated:** the measured impact of a program that includes only the data for participants who actually received the treatment.
- 8. Externality:** an indirect cost or benefit incurred by individuals who did not directly receive the treatment. Also termed "spillover."

Between 1998 and 2001, the NGO International Child Support Africa implemented a school-based mass deworming program in 75 primary schools in western Kenya. The program treated the 30,000 pupils enrolled at these schools for worms—hookworm, roundworm, whipworm, and schistosomiasis. Schools were phased-in randomly.

Randomization ensures that the treatment and comparison groups are comparable at the beginning, but it cannot ensure that they remain comparable until the end of the program. Nor can it ensure that people comply with the treatment they were assigned. Life also goes on after the randomization: other events besides the program happen between initial randomization and the end-line. These events can reintroduce selection bias; they diminish the validity of the impact estimates and are threats to the integrity of the experiment.

How can common threats to experimental integrity be managed?

Worms—a common problem with a cheap solution

Worm infections account for over 40 percent of the global tropical disease burden. Infections are common in areas with poor sanitation. More than 2 billion people are affected. Children, still learning good sanitary habits, are particularly vulnerable: 400 million school-age children are chronically infected with intestinal worms.

Worms affect more than the health of children. Symptoms include listlessness, diarrhea, abdominal pain, and anemia. Beyond their effects on health and nutrition, heavy worm infections can impair children's physical and mental development and reduce their attendance and performance in school.



Poor sanitation and personal hygiene habits facilitate transmission. Infected people excrete worm eggs in their feces and urine. In areas with poor sanitation, the eggs contaminate the soil or water. Other people are infected when they ingest contaminated food or soil (hookworm, whipworm, and roundworm), or when hatched worm larvae penetrate their skin upon contact with contaminated soil (hookworm) or fresh water (schistosome). School-age children are more likely to spread worms because they have riskier hygiene practices (more likely to swim in contaminated water, more likely to not use the latrine, less likely to wash hands before eating). So treating a child not only reduces her own worm load; it may also reduce disease transmission—and so benefit the community at large.

Treatment kills worms in the body, but does not prevent re-infection. Oral medication that can kill 99 percent of worms in the body is available: albendazole or mebendazole for treating hookworm, roundworm, and whipworm infections; and praziquantel for treating schistosomiasis. These drugs are cheap and safe. A dose of albendazole or mebendazole costs less than 3 US cents while one dose of praziquantel costs less than 20 US cents. The drugs have very few and minor side effects.

Worms colonize the intestines and the urinary tract, but they do not reproduce in the body; their numbers build up only through repeated contact with contaminated soil or water. The WHO recommends presumptive school-based mass deworming in areas with high prevalence. Schools with hookworm, whipworm, and roundworm prevalence over 50 percent should be mass treated with albendazole every 6 months, and schools with schistosomiasis prevalence over 30 percent should be mass treated with praziquantel once a year.

Primary School Deworming Program

International Child Support Africa (ICS) implemented the Primary School Deworming Program (PSDP) in the Busia District in western Kenya, a densely-settled region with high worm prevalence. Treatment followed WHO guidelines. The medicine was administered by public health nurses from the Ministry of Health in the presence of health officers from ICS.

The PSDP was expected to affect health, nutrition, and education. To measure impact, ICS collected data on a series of outcomes: prevalence of worm infection, worm loads (severity of worm infection); self-reported illness; and school participation rates and test scores.

Evaluation design — the experiment as planned

Because of administrative and financial constraints the PSDP could not be implemented in all schools immediately. Instead, the 75 schools were randomly divided into 3 groups of 25 schools and phased-in over 3 years. Group 1 schools were treated starting in both 1998 and 1999, Group 2 schools in 1999, and Group 3 starting in 2001. Group 1 schools were the treatment group in 1998, while schools Group 2 and Group 3 were the comparison. In 1999 Group 1 and Group 2 schools were the treatment and Group 3 schools the comparison.



Figure 1: The planned experiment: the PSDP treatment timeline showing experimental groups in 1998 and 1999

	1998	1999	2001
Group 1	Treatment	Treatment	Treatment
Group 2	Comparison	Treatment	Treatment
Group 3	Comparison	Comparison	Treatment

Threats to integrity of the planned experiment

Discussion Topic 1: Threats to experimental integrity

Randomization ensures that the groups are equivalent, and therefore comparable, at the beginning of the program. The impact is then estimated as the difference in the average outcome of the treatment group and the average outcome of the comparison group, both at the end of the program. To be able to say that the program caused the impact, you need to be able to say that the program was the only difference between the treatment and comparison groups over the course of the evaluation.

1. What does it mean to say that the groups are equivalent at the start of the program?
2. Can you check if the groups are equivalent at the beginning of the program? How?
3. Other than the program’s direct and indirect impacts, what can happen over the course of the evaluation (after conducting the random assignment) to make the groups non-equivalent?
4. How does non-equivalence at the end threaten the integrity of the experiment?



Managing attrition—when the groups do not remain equivalent

Attrition is when people join or drop out of the sample—both treatment and comparison groups—over the course of the experiment. One common example in clinical trials is when people die; so common indeed that attrition is sometimes called experimental mortality.

Discussion Topic 2: Managing Attrition

You are looking at the health effects of deworming. In particular you are looking at the worm load (severity of worm infection). Worm loads are scaled as follows:

Heavy worm infections = score of 3
 Medium worm infections = score of 2
 Light infections = score of 1

There are 30,000 children: 15,000 in treatment schools and 15,000 in comparison schools. After you randomize, the treatment and comparison groups are equivalent, meaning children from each of the three categories are equally represented in both groups.

Suppose protocol compliance is 100 percent: all children who are in the treatment get treated and none of the children in the comparison are treated. Children that were dewormed at the beginning of the school year (that is, children in the treatment group) end up with a worm load of 1 at the end of the year because of re-infection. Children who have a worm load of 3 only attend half the time and drop out of school if they are not treated. The number of children in each worm-load category is shown for both the pretest and posttest.

Worm Load	Pretest		Posttest	
	Treatment	Comparison	Treatment	Comparison
3	5,000	5,000	0	Dropped out
2	5,000	5,000	0	5,000
1	5,000	5,000	15,000	5,000
Total children tested at school	15,000	15,000	15,000	10,000

1.
 - a. At posttest, what is the average worm load for the treatment group?
 - b. At posttest, what is the average worm load for the comparison group?
 - c. What is the difference?
 - d. Is this outcome difference an accurate estimate of the impact of the program? Why or why not?
 - e. If it is not accurate, does it overestimate or underestimate the impact?
 - f. How can we get a better estimate of the program's impact?
2. Besides worm load, the PSDP also looked at outcome measures such as school attendance rates and test scores.
 - a. Would differential attrition (i.e. differences in drop-outs between treatment and comparison groups) bias either of these outcomes? How?
 - b. Would the impacts on these final outcome measures be underestimated or overestimated?
3. In Case 1, you learned about other methods to estimate program impact, such as pre-post, simple difference, differences in differences, and multivariate regression.
 - a. Does the threat of attrition only present itself in randomized evaluations?



Managing partial compliance—when the treatment does not actually get treated or the comparison gets treated

Some people assigned to the treatment may in the end not actually get treated. In an after-school tutoring program, for example, some children assigned to receive tutoring may simply not show up for tutoring. And the others assigned to the comparison may obtain access to the treatment, either from the program or from another provider. Or comparison group children may get extra help from the teachers or acquire program materials and methods from their classmates. In any of these scenarios, people are not complying with their assignment in the planned experiment. This is called “partial compliance” or “diffusion” or, less benignly, “contamination.” In contrast to carefully-controlled lab experiments, diffusion is ubiquitous in social programs. After all, life goes on, people will be people, and you have no control over what they decide to do over the course of the experiment. All you can do is plan your experiment and offer them treatments. How, then, can you deal with the complications that arise from partial compliance?

Discussion Topic 3: Managing partial compliance

Suppose none of the children from the poorest families have shoes and so they have worm loads of 3. Though their parents had not paid the school fees, the children were allowed to stay in school during the year. Parental consent was required for treatment, and to give consent, the parents had to come to the school and sign a consent form in the headmaster’s office. However, because they had not paid school fees, the poorest parents were reluctant to come to the school. Consequently, none of the children with worm loads of 3 were actually dewormed. Their worm load scores remained 3 at the end of the year. No one assigned to comparison was treated. All the children in the sample at the beginning of the year were followed up, if not at school then at home.

<i>Worm Load</i>	<i>Pretest</i>		<i>Posttest</i>	
	Treatment	Comparison	Treatment	Comparison
3	5,000	5,000	5,000	5,000
2	5,000	5,000	0	5,000
1	5,000	5,000	10,000	5,000
Total children tested at school	15,000	15,000	15,000	15,000

1. Calculate the impact estimate based on the original group assignments.
 - a. This is an unbiased measure of the effect of the program, but in what ways is it useful and in what ways is it not as useful?

You are interested in learning the effect of treatment on those actually treated (“treatment on the treated” (TOT) estimate).

2. Five of your colleagues are passing by your desk; they all agree that you should calculate the effect of the treatment using only the 10,000 children who were treated.
 - a. Is this advice sound? Why or why not?

3. Another colleague says that it’s not a good idea to drop the untreated entirely; you should use them but consider them as part of the comparison.
 - a. Is this advice sound? Why or why not?

4. Another colleague suggests that you use the compliance rates, the proportion of people in each group that did or did not comply with their treatment assignment. You should divide the “intention to treat” estimate by the difference in the treatment ratios (i.e. proportions of each experimental group that received the treatment).
 - a. Is this advice sound? Why or why not?



Managing spillovers—when the comparison, itself untreated, benefits from the treatment being treated

People assigned to the control group may benefit indirectly from those receiving treatment. For example, a program that distributes insecticide-treated nets may reduce malaria transmission in the community, indirectly benefiting those who themselves do not sleep under a net. Such effects are called externalities or spillovers.

Discussion Topic 4: Managing spillovers

In the deworming program, randomization was at the school level. However, while all boys at a given treatment school were treated, only girls younger than thirteen received the deworming pill. This was due to the fact that the World Health Organization (WHO) had not tested (and thus not yet approved) the deworming pill for pregnant women. Because it was difficult to determine which girls were at risk of getting pregnant, the program decided to not administer the medication to any girl thirteen or older. (Postscript: since the deworming evaluation was implemented, the WHO has approved the deworming medication for pregnant women).

Thus at a given treatment school, there was a distinct group of students that was never treated but lived in very close proximity to a group that was treated.

Suppose protocol compliance is 100 percent: all boys and girls under thirteen in treatment schools get treated and all girls thirteen and over in treatment schools as well as all children in comparison schools do not get treated.

You can assume that due to proper randomization, the distribution of worm load across the three groups of students is equivalent between treatment and control schools prior to the intervention.

Worm Load	Posttest					
	Treatment			Comparison		
	All boys	Girls <13 yrs	Girls >= 13 yrs	All boys	Girls <13 yrs	Girls >= 13 yrs
3						
2						
1						0
Total children tested at school	20000			20000		

1.
 - a. If there are any spillovers, where would you expect them to show up?
 - b. Is it possible for you to capture these potential spillover effects? How?
2.
 - a. What is the treatment effect for boys in treatment v. comparison schools?
 - b. What is the treatment effect for girls under thirteen in treatment v. comparison schools?
 - c. What is the direct treatment effect among those who were treated?
 - d. What is the treatment effect for girls thirteen and older in treatment v. comparison schools?
 - e. What is the indirect treatment effect due to spillovers?
 - f. What is the total program effect?



References:

- Crompton, D.W.T. 1999. "How Much Helminthiasis Is There in the World?" *Journal of Parasitology* 85: 397 – 403.
- Kremer, Michael and Edward Miguel. 2007. "The Illusion of Sustainability," *Quarterly Journal of Economics* 122(3)
- Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72(1): 159-217.
- Shadish, William R, Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company
- World Bank. 2003. "School Deworming at a Glance," *Public Health at a Glance Series*. <http://www.worldbank.org/hnp>
- WHO. 1999. "The World Health Report 1999," World Health Organization, Geneva.
- WHO. 2004. "Action Against Worms" *Partners for Parasite Control Newsletter*, Issue #1, January 2004, www.who.int/wormcontrol/en/action_against_worms.pdf

MIT OpenCourseWare
<http://ocw.mit.edu>

Resource: Abdul Latif Jameel Poverty Action Lab Executive Training: Evaluating Social Programs
Dr. Rachel Glennerster, Prof. Abhijit Banerjee, Prof. Esther Duflo

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.