EXTERNAL EVALUATION

Combination Deworming Program

URL: <http://www.givewell.org/international/technical/programs/deworming>

GiveWell
REAL CHANGE FOR YOUR DOLLAR

By Pierre Thompson
Georgetown University

Initial Draft: 23 December 2011
Revised Draft: 30 December 2011
Table of Contents

I. Conclusions
   A. Fairness of Summary
   B. Bottom Line

II. Content
   A. Disease and Intervention Background
      1. Burden and Scope of Disease
      2. Description of Intervention
   B. Track Record and Benefits of Intervention
      1. Relevant Evidence Base
      2. Evidence of Past Impact
         a. Evidence of Improved Health Outcomes
         b. Successful Delivery and Usage of Intervention
   C. Cost-Effectiveness of Intervention
      1. Issues with Cost-Effectiveness Estimates
         a. Health Benefits of Deworming
         b. Economic Costs of Deworming
      2. Alternative Cost-Effectiveness Estimates
   D. Room for More Funds

III. Sources
   A. Books
   B. Periodicals
   C. Miscellaneous
PART 1: CONCLUSIONS

A. Fairness of Summary

The summary is exceedingly fair, addressing the structure of GiveWell’s justified beliefs and the limits of its knowledge. In fact, there is a bit of unnecessary hedging, which might undermine a donor’s initial confidence in the review. The reader is told that “the benefits are potentially major, but also debatable” and learns in the next five sentences about “weaker evidence,” “limited information” and “potential problems.” An earlier version of the review even confessed that “GiveWell’s guess [is] not with high confidence.” There must be a less discrediting way to convey the same message.

Perhaps the writing style is emblematic of GiveWell’s dispassionate approach to evaluating charities. Still, the summary should articulate a stronger endorsement for this program. It should describe the intervention as being cost-effective in absolute terms, not just in relative terms which not everyone understands ("less cost-effective than distribution of LLINs, but may be more cost-effective"). Without any further context, a first-time reader of the summary would not easily infer that deworming is one of GiveWell’s highly recommended programs.

B. Bottom Line

I believe that GiveWell has reached a reasonable assessment with this review, but the reasoning process itself could have been more lucid. At the end of the day, this review must make the case that deworming does (or does not) change lives for the better, and it should provide abundant evidence and logic to support that viewpoint. With respect to measuring impact, the evidence focuses largely on demonstrating improved health outcomes through randomized controlled trials; as such, the current evidence base does not balance internal and external validity. Considering a broader range of methodological and disciplinary perspectives, as well as alternative outcomes, could make GiveWell’s conclusions more robust. Donors also need practical information about how the intervention is designed and implemented, including challenges in the field – these details are too important to be implied. Finally, I maintain reasonable confidence in the cost-effectiveness analysis which GiveWell has provided, and do not expect the marginal adjustments I describe to change that assessment.

No matter how strong, the evidence is only as good as the logic that synthesizes it. Much like a mathematical proof, the review should follow a clear line of reasoning and cite relevant evidence at every step. The statement which best describes the rationale for deworming is not found in the review itself, but in a later blog post about it. With great
clarity, Elie writes that “[GiveWell’s] positive view of deworming stems from the fact that … the intervention is so cheap that it is likely to be a relatively good buy and may be a great buy.” Thus, the burden of proof for supporting this particular intervention seems to depend on the “balance of probabilities” rather than “beyond reasonable doubt.” GiveWell would do well to emphasize that distinction. I feel that the “balance of probabilities” approach, which might entail looking at successful delivery of the intervention, has been shorted; this explains why important concepts, such as coverage rate and cure rate, are unfortunately relegated to the footnotes. Finally, there is a strong philosophical case to be made for this intervention because it is extremely cost-effective and the burden of disease is spread across many disenfranchised poor people. GiveWell should dedicate time and space to explain the principles of distributive justice which may have influenced its recommendation of this intervention.
PART 2: CONTENT

A. Disease and Intervention Background

Burden and Scope of Disease

This section provides a fair overview of two infectious diseases commonly targeted by preventive chemotherapy interventions: schistosomiasis and soil-transmitted helminthiasis. It carefully differentiates schistosomiasis from STH infections according to prevalence, pathology and symptoms. (A matrix could probably facilitate this comparison.) I believe more could be done to emphasize the commonalities and to place them in the framework of the neglected tropical diseases. Donors deserve to know that the burden of disease falls disproportionately on poor people in remote and rural areas who do not constitute an effective health lobby – and therefore lack access to essential medicines.

The body of the review should include critical information about the life cycle (transmission) of the parasite and human risk factors, which are buried in the footnotes. To most lay readers, the schistosomiasis description would appear to contain a blatant contradiction. It is asserted that “schistosomes cannot reproduce inside the body” and later in the same paragraph that “the morbidity caused by schistosomiasis arises from the eggs that the parasite lays while it inhabits the human host.” A non-specialist would reasonably associate oviparity with reproduction, and therefore be led to believe that schistosome reproduction takes place inside the human body. According to my understanding, schistosome eggs hatch only after being returned to the water source via defecation or urination; the hatching requirement is an exceedingly technical detail, but probably worth explaining to avert greater confusion.

GiveWell notes that its analysis excludes S. Japonicum because S. Japonicum does not fall within the purview of “the charity for which we have undertaken this review.” But the statement that S. Japonicum is “believed by some scholars to be more dangerous than the other strains of schistosomiasis” seems to draw unwarranted attention. A critical donor might wonder about the omission. Is GiveWell selectively withholding information that might compromise its analysis? Or are human lives in the Asia-Pacific region somehow discounted relative to those in Sub-Saharan Africa? The review should explain

---

1 I disagree with the nomenclature used throughout this review: “soil-transmitted helminthiasis” or “soil-transmitted helminth infection” should refer to the disease, whereas “soil-transmitted helminth” should be reserved for the agent. The World Health Organization also employs this convention.

2 I disagree with GiveWell’s characterization of schistosomiasis prevalence. I have read multiple papers stating that 200 million people are presently infected with schistosomiasis, while an additional 600 million people are at risk of infection (because they live in endemic areas).
that national programs in China and Japan have largely succeeded in controlling the geographically localized species of schistosomiasis, meaning there are probably no extant charities a donor could fund to target *S. Japonicum* (or *S. Mekongi* or *S. Intercalatum*).

**Description of Intervention**

This section provides a very concise description of the target population and treatment regimen – and even fewer details about the program design, implementation and measurement of success. At a minimum, the description should clarify that mass deworming is a community-based control program intended to prevent morbidity and in some cases reduce transmission. Mass deworming is a simple treatment, not an integrated control strategy. The field notes can be misleading because they describe a health education component instructing students in better hygiene, which falls technically outside the domain of mass deworming.

At the moment, it is hard for me to discern how donations to this program would be spent. I would like to know how funding is typically allocated across different program expenses (e.g. administration, procurement, logistics). Natalie and Holden’s field notes contain numerous observations from a demonstration deworming that provide a more holistic picture of the program. We learn, inter alia, about the casual nature of the program, the free provision of treatment and some incidence of public hostility toward the drug. These details may not be generalizable, but excluding them from the body of the review would deprive donors of potentially relevant criteria and practical information on which to assess the effectiveness of the program.

The main drawback with this section is that it describes how mass deworming would work in theory, but not in practice. There could easily be a large discrepancy between the two scenarios. For example, the Disease Control Priorities report asserts that “with support from the local health system, teachers can deliver the [combination deworming] drugs safely.” However, this forces a major assumption about the functionality and quality of the local health system. Abhijit Banerjee and Esther Duflo point out that health workers in poor countries have been known to exhibit high rates of absenteeism, low motivation and plain incompetence. An intelligent donor would question whether the local health system is actually capable of making this intervention work. Finally, complex case management can always occur during implementation.

---

3 See *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*
B. Track Record and Benefits of Intervention

Relevant Evidence Base

The review presents an evidence base that is highly relevant, but homogeneous. The majority of the studies are sourced from literature reviews in respected medical journals (*The Lancet, Acta Tropica, et al*) and systematic reviews of primary research in healthcare (*Cochrane Collaborative*). There is a serious concern that publication bias may have shaped the available literature, to the extent that academic journals are disposed to publish work that shows health or economic interventions having a nonzero effect. For the most part, the literature reviews conveniently summarize the most relevant research from a probably large universe of studies on deworming. Several reviews incorporate meta-analysis, which is a potentially useful statistical method for estimating the overall effect of multiple studies. Finally, GiveWell points to SCI’s internal monitoring results, but since this is a program review I wonder whether data from comparable charities (Carter Center) is available.

I happen to believe the evidence base could have been construed more broadly. The review draws heavily from randomized field evaluations in the health and economics literature. While undoubtedly rigorous, I am not convinced that the current evidence base balances internal and external validity. I would appreciate seeing more large-scale or retrospective studies (similar to Bleakley 2007). Only a small number of studies (King, Dickman and Tisch 2005; Baird 2011) incorporate qualitative research methodologies, such as longitudinal studies and systematically administered surveys. It may be worth surveying the literature from outside disciplines, including medical anthropology and medical sociology, for insights on health-seeking behavior of poor people or how specific institutions function. I do not suppose any of these research strategies would fundamentally change GiveWell’s conclusions, but they could make existing conclusions more robust and may even shift priors.

Evidence of Past Impact

According to GiveWell’s impact criteria, there are two approaches to proving the impact of health interventions. The most intuitive approach is to show “evidence of improved health outcomes [such as] lowered incidence or prevalence of disease; drops in death rates; etc.” But in some cases, improved health outcomes may be too difficult or subtle to measure. An alternative approach is to show successful delivery and usage of the intervention, especially if the treatment has been “thoroughly and rigorously tested” and

---

4 I am not familiar with the preliminary research that went into this review, so it is possible that GiveWell investigated a much broader evidence base yet determined there was no more relevant information.
its efficacy is not in scientific dispute. The review thoroughly assesses evidence related to “improved health outcomes,” but insufficiently addresses evidence related to “successful delivery and usage of the intervention.”

Evidence of Improved Health Outcomes

Evidence of improved health outcomes arising from mass deworming would ideally: (a) demonstrate causality, and (b) ensure the studies are representative. The review submits a large number of randomized controlled trials for consideration; RCTs are excellent at demonstrating causality under very narrowly defined circumstances. The review chooses to discuss the effect of deworming on three main health outcomes: subtle general health impacts, prevention of potentially severe effects, and developmental (lifetime productivity) effects. This is a logical way to cover ground since most studies focus on a single health outcome for either disease. Still, there are multiple areas where GiveWell’s selection or interpretation of the evidence seems less than competent.

Starting with deworming’s effectiveness at curing subtle morbidity: the review looks for positive change in hemoglobin concentration as evidence of reduced anemia at the population level. GiveWell asserts that anemia is the symptom “for which we have, by far, the most and highest-quality evidence.” Because anemia is most strongly associated with hookworm infection, extrapolating that relationship to other nematode and trematode infections seems misleading at best. According to my understanding, morbidity is correlated with the number of eggs the parasite lays in the human host. It seems more appropriate to measure subtle health improvements by the egg reduction rate, adjusted for the baseline in various populations. Two of the Cochrane Reviews cite a large number of studies where health outcome is measured in eggs per gram of human stool using the Kato-Katz method.$^5$

The section discussing potentially severe health effects is not very helpful. It manages to avoid any discussion of deworming’s impact on preventing deaths or clinically acute symptoms. There is obviously a scarcity of data in this area, but I suppose GiveWell could have compared actual and expected mortality rates in a population, using the case fatality rate. The review also notes that death estimates in Sub-Saharan Africa for STH fell from 4,000 deaths in 2001 to 412 deaths in 2004. Such a dramatic drop in mortality in a span of four years deserves further investigation, and might even be submitted as large-scale evidence in support of deworming’s impact (assuming the decrease could be explained by a contemporaneous deworming campaign). Unlike other sections, this one has no “bottom line” on severe health effects.

$^5$ One paper (Black 2009) proposed: $\frac{\text{Egg count before treatment} - \text{Egg count after treatment}}{\text{Egg count before treatment}} \times 100$
Moving on to developmental impacts, I agree with the essence of the argument but personally find the term “developmental impacts” too vague. Nominally, it refers to “subtle, lasting impact on children’s [cognitive and physical] development.” As far as I’m concerned, it more accurately describes a “lifetime productivity” effect, i.e. how subtle health improvements raise economic productivity over a lifetime through the mechanism of schooling or labor.\(^6\) I think it would be helpful to introduce the analytical framework of a “nutrition-based poverty trap” to advance this argument. I am not surprised that the studies testing for effects on schooling and income would face difficulties. It is notoriously hard to measure educational gains based on test scores, especially when the underlying problem might be the education system itself. Labor market outcomes are theoretically based on returns to education, experience and skills (which is largely unmeasurable, sometimes correlated with the error term).

I believe more could be done to emphasize the adverse effect that morbidity has on adult worker productivity, since most poor people work in agriculture or labor-intensive jobs. Similar to children, some adult populations (notably farmers and women) may be at increased risk of infection because of daily water contact.

The review hardly addresses the issue of generalizability (other than point out which studies are not representative), so I will add a few remarks. Combining the results from RCTs carried out under diverse circumstances may achieve a kind of aggregate representativeness. In San Francisco, Natalie suggested that meta-analyses of these studies may provide some degree of external validity. The main concern, I think, is that a large number of clinical trials were pre-screened. The review does not explain how pre-screening for infected individuals is a form of selection bias that exaggerates the intervention’s effect on the general population (which contains both infected and uninfected individuals). In practice, preventive chemotherapy focuses on targeted treatment rather than selective treatment, due to the high marginal cost of screening individuals.

**Successful Delivery and Usage of Intervention**

We know that the prevalence of schistosomiasis or STH in some populations is so high that everyone is presumed to be infected. We also know that deworming has extremely high cure rates, well above 50 percent in most cases. If we choose to demonstrate successful delivery and usage of the intervention, then measuring impact becomes a matter of probabilities. According to GiveWell’s impact criteria, the evidence would have

\(^6\) By this logic, deworming children will always be more impactful than deworming adults simply because children have more life remaining.
to bear that: (a) mass drug administration is carried out appropriately and consistently, (b) health education modules succeed in changing behavior over the long term, and (c) supplies are used appropriately and consistently by beneficiaries. Since the review does not pursue this line of inquiry, I will discuss some potential problems with the body of evidence for this approach.

There is sufficient reason to doubt that combination deworming drugs are administered appropriately or consistently. Natalie and Holden’s field notes suggest that the drug is not administered by trained health professionals; there is significant guesstimation involved in selecting each dosage; feeding programs are not available at all of the schools; some students are absent from the deworming; and the supply of drugs can be depleted. To be sure, I have abstracted details from one specific deworming project, but my intention is simply to point out several areas where program implementation has a considerable margin for error. The weight of this evidence suggests that treatment does not always reach the target population (0 < coverage rate < 1), nor would it prove efficacious in one hundred percent of the cases it does reach (0 < cure rate <1).

The second and third criteria are not really germane to this review. The question of whether health education changes behavior is interesting – at least one study (Midzi 2011) suggests schoolchildren know very little about schistosomiasis and STH prevention – but falls technically outside the domain of this review. As to whether supplies are used appropriately and consistently by beneficiaries, the drug’s efficacy is normally considered beyond the individual’s control after the point of administration.

**Alternative Hypotheses for Empirical Patterns**

The review summarily considers strong alternative hypotheses for nearly every empirical pattern that could influence the weight of the evidence. The studies which could potentially provide the strongest validation of the program’s impact are not coincidentally subjected to the closest scrutiny. For example, Miguel and Kremer 2004 advances the argument that deworming students improves school performance, implying potential developmental and lifetime productivity effects; the paper substantiates what GiveWell considers “the most compelling case for deworming as a cost-effective intervention.” But rather than elevate Miguel and Kremer’s findings, the review proceeds to challenge them in previously unconsidered ways. GiveWell’s readiness to play devil’s advocate is a recurring theme throughout this review, yet the burden of proof is never so exacting as to defy common sense. Therefore, I am convinced that the empirical patterns which the review submits as strong evidence must hold beyond a reasonable doubt.

---

7 The review argues that the study was conditioned on unseasonal flooding and arbitrary school assignment, which are not likely to be replicated in successive mass deworming interventions.
Likelihood of Future Impact

Considering how strenuous it was to demonstrate past impact, it seems ambitious to assess the future impact of this program. Jeffrey Sachs reminds us that two challenges involve scale and sustainability.\(^8\) Whether the program can succeed on a large scale will depend on its ability to adapt to local conditions while still adhering to best practices. This proposition is difficult for me to remark on because the review has not sufficiently established what the ideal program model would even look like. Implementation details, such as whether children are fed prior to mass drug administration, are important enough to consider because they could well determine the margin of impact and overall success of the scale-up. And as the program expands to more countries, special circumstances in those countries may hinder future success. For example, political instability and military conflict in African countries may complicate the intervention.

Buried deep in the field notes is Professor Alan Fenwick’s assessment that “much of the low-hanging fruit for schistosomiasis control programs (in terms of countries where it’s likely to work) has been plucked, though the Kenya program is promising and an Ethiopia or Angola program could go well.”

Aside from scale-up, the other main concern is whether the program can be sustained. The program review hardly mentions the issue of sustainability, but the charity review for SCI makes explicit “it is important that deworming programs are sustained over time, as re-infection is rapid and a one-time treatment may have little long-term effect.” A judicious donor would be interested in knowing whether the program has enough resources to adhere to the schedule of administration recommended by WHO.

Best and Alternative Methods to Assess Empirical Evidence

In most cases, the review uses superior analytical methods to assess empirical evidence. For example, the disability-adjusted life year metric is used not only to evaluate the cost-effectiveness of deworming against other interventions, but also to compare the death and non-death burden of schistosomiasis. In fact, GiveWell could have also used DALYs to describe the burden of chronic parasitic diseases relative to other global diseases. By some accounts, the disease burden of schistosomiasis exceeds 70 million DALYs (Gray 2011). In theory, this should bolster the case that schistosomiasis and STH are worthy of our attention because they exact a huge burden which happens to be distributed across many poor people.

\(^8\) See Common Wealth: Economics for a Crowded Planet
At some point, it would be helpful for GiveWell to explain the principles of distributive justice that inform its value judgments. An intervention such as this one raises many interesting philosophical questions: Is the goal of this intervention to maximize welfare? to ensure equal access to health? to benefit the neediest people? I am aware that GiveWell describes itself as having “global humanitarian” values, but it is unclear whether this approach to charity implies a utilitarian or egalitarian or justice as fairness interpretation.

Finally, I beseech GiveWell to describe its own research methodology for this review and the heuristics it used to determine which papers found their way into the evidence base and which were excluded from it. Many systematic reviews quantify how many papers are screened at each stage of the review and how many meet each successive criteria. GiveWell would do well to follow their example.

**Offsetting Impacts**

The review mentions some interesting theoretical downsides to the program, but cites a lack of evidence to support any of those claims. Resistance to antischistosomal or anthelmintic treatment is a very real possibility that should not be discounted. Danso-Appiah 2008 expresses “considerable concern” that resistance could develop against praziquantel if it is used exclusively as the antischistosomal drug. Aside from that, I think it may be worthwhile to consider whether serious adverse effects related to deworming are great enough to deter people from any sort of mass drug administration. Poor people often have health-related superstitions and prejudices which may be hard for an outside intervention to overcome.

**C. Cost-Effectiveness of Intervention**

**Issues with Cost-Effectiveness Estimates**

Cost-effectiveness, as the term implies, is based on estimates of the intervention’s effectiveness and cost. When GiveWell estimates the benefits of deworming, it should underline assumptions regarding: (a) the burden of disease in a population and (b) how much of that burden the intervention will alleviate. When GiveWell estimates the costs of deworming, it should factor in all relevant costs regarding: (a) program, (b) administrative and (c) advocacy costs. Finally, there may be additional cost and benefit externalities, as well as dynamic trends which could have been mistaken for static events.

In truth, there are myriad ways in which the cost-effectiveness of deworming could be overstated or understated. Actual costs and benefits may vary widely based on a number
of local factors. My sense is that those who set out looking for evidence to support an incremental upward or downward revision will probably find it. It seems fair to say that most new adjustments are minor and, on balance, probably cancel themselves out. For the sake of illustration, I outline a few below:

*Health Benefits of Deworming*

The review neglects to mention one positive externality of deworming: there is at least one study (Bhunu 2010) which suggests that schistosomiasis infection enhances HIV susceptibility through co-infection. Reducing the disease burden of schistosomiasis may indirectly reduce the disease burden of HIV/AIDS. Thus, we may have underestimated the benefits of deworming.

*Economic Costs of Deworming*

And we may have overestimated the costs of deworming. Cost-effectiveness analysis should account for pharmaceutical trends in the international political economy. The cost-effectiveness calculation assumes a fixed cost of $0.08 per donated praziquantel. Though not a fundamental input into the $0.51-0.68 cost per person dewormed, we can reasonably expect the cost of praziquantel (and other drugs) to fall over time. The WTO's August 2003 declaration on the Agreement on Trade-Related Aspects of Intellectual Property Rights made it easier for developing countries to import cheap generics for public health purposes. Although TRIPS implementation has varied widely across the poor countries, to the extent more countries adopt flexibilities such as compulsory licensing, to that extent market prices will be driven down.

*Alternative Cost-Effectiveness Estimates*

At the moment, I believe GiveWell’s cost-effectiveness estimate is the most parsimonious that can be reached. At around $170 per DALY averted, deworming is competitive with the most cost-effective public health interventions, most notably LLINs. The conservative calculations from the Disease Control Priorities (second edition) scenario estimate an average cost between $5000 and $7000 per life saved, well within the “cost-effective” range of $1000 and $10,000 per life saved. The estimate could be significantly higher or lower, depending on how one evaluates “developmental impacts.” Given the wide margin of error afforded to cost-effectiveness estimates, I doubt that revised estimates will ever find deworming convincingly outside the cost-effectiveness range.
The monetary inputs for the cost-effectiveness estimate were based on the upper amount of funding available to SCI from 2003 to 2010. Since this is a program review and not a charity review, one could argue that SCI’s program is not representative of all deworming programs (SCI is likely more cost-effective due to superior management). I am intrigued by GiveWell’s method of using grant receipts to figure out total costs. It is the perfect approach if we assume that grant amounts are empirically determined, and that treatment capacity is directly correlated with program expenses. I happen to believe that grant making and program spending are not mathematical sciences, and cost-effectiveness estimates based on these inputs will probably be off by a small margin.

Finally, I believe more could be done to emphasize the relative costs of this intervention. The private market for deworming drugs often does not reach poor people in poor countries, making individual procurement prohibitively expensive. Mass deworming can achieve cost-effectiveness through economies of scale, which makes it appealing as a public health intervention. The review should also discuss the economic rationale for combining antischistosomal and anthelmintic treatments, since many areas can be afflicted with both kinds of parasites.

D. Room for More Funds

This section redirects to the charity review of SCI, which offers a concrete plan for how additional donations will be employed. Without reading too much into the details, it is possible to see that SCI plans to fund “unexpected and urgent” projects, prepare for program expansion to new countries and achieve better coverage rates with the existing program. The review notes that the proposed activities will be “less beneficial” at the margin than the deworming activities described in the review. In any case, these seem to be good strategies for long-term schistosomiasis control. I would be interested in knowing whether the countries where this program will expand have lower prevalence rates, or are logistically more challenging to reach, or both. Even supposing the intervention became less cost-effective over time, I cannot see any case where there would be no room for additional funding. As long as more than a quarter of humanity is infected with schistosomiasis or STH, there will always be room for morbidity control.
PART 3: SOURCES

A. Books


B. Periodicals


C. Miscellaneous


