**The impact of mass deworming programmes on schooling and economic development: an appraisal of long-term studies**

**Sophie Jullien, David Sinclair, Paul Garner\***

Author affiliations

Centre for Evidence Synthesis in Global Health, Department of Clinical Sciences, Liverpool School of Tropical Medicine

\*Corresponding author: Centre for Evidence Synthesis in Global Health, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. E-mail: [Paul.Garner@lstmed.ac.uk](mailto:Paul.Garner@lstmed.ac.uk)

4 Sept 2016

**Background**

Documents from advocacy and fund-raising organizations for child mass deworming programmes in low and middle-income countries cite unpublished economic studies claiming long-term effects on health, schooling and economic development.

**Methods**

To summarise and appraise these studies, we searched for and included all long-term follow-up studies, based on cluster-randomized trials included in a 2015 Cochrane review on deworming. We used Cochrane methods to assess risk of bias, and appraised the credibility of the main findings. Where necessary we contacted study authors for clarifications.

**Results**

We identified three studies (Baird 2016, Ozier 2016, and Croke 2014), evaluating effects more than nine years after cluster-randomized trials in Kenya and Uganda. Baird and Croke evaluate short additional exposures to deworming programmes in settings where all children were dewormed multiple times. Ozier evaluates potential spin-off effects to infants living in areas with school-based deworming. None of the studies used pre-planned protocols, or blinded the analysis to treatment allocation.

Baird 2016 has been presented on-line in six iterations. It is at high risk of reporting bias and selective reporting, and there are substantive changes between versions. The main cited effects on secondary school attendance and job sector allocation are from post-hoc sub-group analyses, which the study was not powered to assess. The study finds no evidence of effect on nutritional status, cognitive tests, or school grades achieved, but these are not reported in the abstracts.

Ozier 2016 has been presented on-line in four iterations, without substantive differences between versions. Higher cognitive test scores were associated with deworming but were only beyond the play of chance with inclusion of the non-randomised data. The size of the stated effect seems inconsistent with the short and indirect nature of the exposure to deworming, and a causal pathway for this effect is unclear.

Croke 2014 utilizes a data set unrelated to the base trial to report improvements in English and maths test scores. The analysis is at high risk of attrition bias, due to loss of clusters, and is substantially underpowered to assess these effects.

**Conclusion**

In the context of reliable epidemiological methods, all three studies are at risk of substantial methodological bias. They therefore help in generating hypothesis, but should not be considered reliable evidence of effects.

**Keywords**

Helminths, parasitic worms, children, cluster analyses, bias

**Key messages**

* The long-term societal effects of mass deworming programmes for soil transmitted helminths in low and middle income countries are contested.
* Advocates cite economic studies reporting long-term effects on health, schooling, and economic development. We sought and appraised these studies using health technology assessment methods based on epidemiological principles.
* In the eleven reports from three studies, we found multiple potential sources of bias; in the study methods, analysis, and reporting. Of particular concern are the lack of pre-planned protocols, multiple hypothesis testing followed by selective reporting of favourable results, and post-hoc sub-group analyses.
* Our interpretation is that these trials do not provide credible evidence to support the claims of long-term effects. However, they raise interesting hypotheses that could be considered in further research.

**Introduction**

Soil transmitted helminths remain common in many low and middle-income countries, despite some evidence that global infection intensity may be declining [1]. The worms are unpleasant, cause discomfort, and with heavy infections can undermine nutritional status and lead to serious complications [2][3]. It is therefore obvious that children with symptomatic infection should be treated. It is also obvious that repeated mass treatment of whole communities with effective drugs will reduce the overall worm burden, where helminths are common, at least in the short term [4][5].

What is less obvious is whether mass deworming programmes have any measurable long-term effect on health and nutrition at the community-level. A 2015 Cochrane review of trials administering multiple rounds of deworming treatment found little or no effect on average weight gain or average haemoglobin across 10 trials with more than 38 000 participants [6]. The review authors (which include one of the authors of this paper) interpreted this as reasonable evidence of no effect, but others have claimed the trials were simply too short, or poorly designed for detecting effects once infected children are diluted among uninfected children [7][8].

However, much of the advocacy and fund-raising for mass deworming programmes in children has drawn on studies reporting long-term effects on school attendance and economic development [9][10][11][12]. This advocacy contributed to the decision by India to run the largest national deworming programme in the world (targeting 270 million children in schools and preschools in 2016), and the Cochrane review has been criticised for excluding the studies cited for these effects [10][13].

The objective of this paper is therefore to use health technology assessment methods, based on epidemiological principles, to appraise the methods of these studies, and interpret their findings in the light of this appraisal.

**Methods**

**Inclusion criteria**

All follow-up studies based on randomized or quasi-randomized experimental trials (termed “base trials”) included in the 2015 Cochrane review.

We included all outcome domains identified by the literature in this field as important for decision-making, and included in the Cochrane review (see figure 1): nutritional status (measured by weight, height and haemoglobin); physical well-being (measured by exercise tolerance or self-reported measures); school attendance (measured by days present at school or years of school enrolment); cognition and school performance (measured by formal tests and exam performance) [6]. In addition, we included all economic productivity outcomes the author teams deemed reasonable overall measures.

**Search strategy**

We identified the main unpublished studies being cited, by reviewing the reference lists of prominent papers [9][13], and the webpages of deworming advocates [10][11][14].

We also searched Pubmed for published follow-up studies to all cluster-randomized studies included in the Cochrane review by using the search terms: ((deworm\*[Title/Abstract] OR helminth\*[Title/Abstract]) AND (Alderman[Author] OR Awasthi[Author] OR Hall[Author] OR Stoltzfus[Author] OR Wiria[Author] OR Rousham[Author] OR Miguel[Author])).

Two authors independently screened the search results and applied the inclusion criteria.

**Risk of bias assessment**

For the base trials, we described the study design, setting, population, intervention and control. We assessed the risk of bias using the Cochrane tool for appraising randomized controlled trials and considered the potential for bias to also influence the results of the follow-up studies [15].

For each follow-up study, we described the study design, population, timing, intervention and control group exposure to mass deworming, analytic approach, outcome measurement and reporting, and results. For the risk of bias assessment, we adapted the Cochrane tool for randomized controlled trials to take into account the additional risks posed by cross-sectional sampling from communities, many years after the planned experiment finished:

* Selection bias: we considered the randomization process of the base trial, the methods for selection of a proportional sample, and the balance of potential confounders between groups.
* Reporting and detection bias: we considered the methods used to blind those collecting and analysing the data from the treatment allocation.
* Attrition bias: we considered loss of clusters, exclusion of participants after enrolment, migration in and out of the study area, and the proportion and potential impact of missing outcome data.
* Selective reporting bias: we considered the use of a pre-planned protocol, the number of outcomes assessed and the potential for false positive results, changes in the reporting of outcomes over time, and inclusion of important findings (showing association, or showing a lack of association) in the abstract.

### For each domain, we classified studies as ‘low risk’ when appropriate methods were described to reduce the potential for bias; ‘high risk’ when the methods described were inadequate to negate the potential for bias to influence the results, and ‘unclear risk’ when the impact of any methodological problems was uncertain or there was insufficient information to make a clear judgment. We refined these assessments after contacting the study authors for additional information.

**Outcome credibility assessment**

We first summarised the effect size and 95% confidence interval for all outcomes reported by the studies across the policy-important domains.

As the included studies are outside the scope of what would normally be included in a Cochrane Review, we familiarized ourselves with the study methods and findings, and discussed which factors would be important when appraising the results. We then applied this appraisal systematically across studies.

We then further assessed the credibility of all the main findings reported in the abstracts, by considering: the evidence base for the stated effect from the main text (we considered an effect to be present if P < 0.05); the power of the study to detect this effect; the consistency of the effect across sub-groups; the consistency of the effect across similar or related outcomes; and the robustness of the effect to adjustment for multiple inferences (although statistical adjustment for multiple testing is of limited value without a pre-planned analytic protocol, we considered the effect to be robust if the FDR q-value < 0.05) [16].

We also considered whether intermediate effects were present or absent on plausible causal pathways, and the plausibility of the effect in relation to the intensity of the intervention.

**Results**

We identified three unpublished, long-term follow-up studies [17][18][19], based on two cluster-randomized trials from Kenya and Uganda [20][21] (see table 1). One additional study was excluded, as it was not based on a randomized experiment [22].

All three follow-up studies are economic working papers available on-line but not formally published (Appendix 1). The study by Baird has been presented on-line in six iterations, although we were only able to access five (2011a, 2011b, 2012, 2015, 2016). The study by Ozier has been presented on-line in four iterations (2011, 2014, 2015, 2016).

The search for published studies returned 94 records, of which none were judged relevant to this review: 8 reports corresponded to the base trials in the Cochrane review, 11 were studies older than the base trials, 40 were not relevant to deworming interventions, and the remaining 35 were not long-term follow-up studies.

**Kenya trial (Miguel & Kremer 2004)**

The base trial for the first two studies was conducted in Busia District, in Western Kenya, by Miguel and Kremer. The intervention comprised deworming drugs administered every six months, plus regular worm prevention education through public health lectures, wall charts, and training of teachers. Seventy-five schools, with 32,565 pupils aged between 6 and 18 years, were allocated sequentially to one of three treatment groups. Group 1 received the intervention from 1998, group 2 from 1999, and all groups received the intervention from 2001 onwards.

**Risk of bias assessment:** The quasi-randomized design means that there is a small risk of systematic differences between groups, and this risk will probably still be present in the follow-up studies. In addition, any effects observed in the follow-up studies may be attributable to the effects of the public health education activities rather than the anti-helminthic drugs. Although some would argue this is part of the intervention, it is not the main component of most large national deworming initiatives [10]. A complete risk of bias assessment is available in table 2. Of note, an independent replication analysis of this trial was carried out in 2015, which found errors in the analysis of reported effects on haemoglobin and nutritional status,;the authors now acknowledge these effects are not “statistically significant”. In a second replication that used the original authors analytical approach the externalities were also not demonstrable, but the original trial authors have adjusted the parameters, conducted new analyses, and contest this [23][24][25].

**Baird study (reported in a series of papers 2010-16)**

The Baird series analyse a questionnaire survey of 5084 adults, 9 to 11 years after they participated in the Kenyan trial [20]. The analysis compares adults from schools who began receiving the intervention in 1998 and 1999, with adults from schools who did not receive the intervention until 2001 (see Appendix 2).

As all participants eventually received the intervention, this study looks for effects attributable to the intervention group receiving an additional 2.4 years of the deworming intervention compared to the control group (table 1). The paper presents data on nutritional, health, schooling and labour market outcomes.

**Risk of bias assessment:** The survey sampled adults from a complete list of all children who attended the schools quasi-randomized in the base trial, using computerized randomization and stratified by school, grade and gender (see table 3). Baseline data were presented for age and academic performance prior to the base trial, and although these appear balanced this is probably insufficient to exclude the possibility of confounding due to the quasi-randomized design of the base trial. The analysis did not follow a pre-planned protocol, and those analysing the data were not blinded to treatment allocation.

The five versions available on-line to mid-2016 contain substantially different analyses which appear exploratory, and there is a high risk of false positive results with the number of hypotheses tested for statistical significance increasing from 228 in Baird 2011a to 650 in Baird 2016, largely due to the introduction of sub-group analyses (see Appendix 3). This process appears to be at high risk of reporting bias, and a narrative analysis suggests selective reporting:

* Some outcomes reported in early versions were dropped from later versions. It is not clear to the reader why, but it is likely to be due to the failure to demonstrate an effect (for example, cognitive test results reported in 2011a, but absent in 2016; with no apparent effect on Ravens Matrices or English vocabulary);
* Effects are presented for outcomes which appear to be part of a larger undisclosed data set (for example, “self-reported health rated as very good” presented without additional categories; and “Kenyan women who participated as girls have fewer miscarriages” without presenting other health related outcomes);
* Results from post-hoc sub-group analyses are given prominence in the abstract and results (for example, an increase in secondary school attendance in females is stated in the 2016 abstract, but no effect was apparent in the whole sample, and disaggregation by sex only appeared from 2012 onwards);
* The abstract changed substantially between versions, but none reported important findings of no effect (for example, there were no effects apparent on BMI or height but these are not reported in any of the five abstracts; see table 4 and Appendix 4).

To further examine the influence of selective reporting we compared the “statistically significant” findings (P < 0.05) presented in the abstract, with the overall findings presented in Baird 2011a (table 4). In the abstract, Baird reports that physical well-being, school enrolment and attendance, and school performance or cognition are significantly higher in the group receiving earlier deworming. However, in the main text tables, only one of the seven outcomes measuring school performance/cognition is statistically significant. Similarly, for school attendance, an effect was only apparent in one of the three outcomes reported. Economic productivity was more complicated, as there were numerous sub-group analyses and a variety of derivative measures; an effect was apparent in 13 outcomes, with a further 19 reporting no statistical significant effect.

**Credibility assessment:** In table 5 we attempt to provide a balanced presentation of the key results from Baird 2016, stratified by the policy-relevant outcome domains, and in table 6 we present our credibility assessment for the outcomes reported in the abstract of Baird 2016.

In their 2016 abstract, Baird et al state that men stayed “enrolled for more years of primary school”, and women were “approximately one quarter more likely to have attended secondary school”. These statements are supported by “statistically significant” results within the text, but presentation of these two results in isolation could be regarded as misleading, as there is other information that is required for a balanced interpretation: 1) These effects were not present in the whole sample, and are only apparent in post-hoc sub-group analyses which the analysis was not adequately powered to examine; 2) Neither result is robust to the authors’ own adjustments for multiple inferences; 3) These are selected positive findings among a group of results for similar or related outcomes, that either show no effect (there was no evidence of an increase in the number of school grades attained in either sex), or provide an alternative explanation for these effects (those in the intervention group were actually more likely to have repeated a grade).

The abstract then uses these selected measures of educational effects to explain apparent shifts in the labour market, which are presented as beneficial. However, it is not clear to us which of these shifts represent a genuine economic improvement. For example, the number of hours women worked in agriculture appears lower in the intervention group and is presented as a benefit, but the number of hours worked by men appears higher. In reality, an effect in either direction could be interpreted as a benefit due to the alternative explanations of better health (enabling longer hours in manual work), or better education (enabling a move to higher skilled work). It is perhaps more useful to note that there was no evidence of an increase in hours worked in waged employment, and no evidence of an increase in non-agricultural earnings (waged earnings plus self-employed profits).

The authors clarified that the sample size was calculated to detect a 15% relative increase in secondary school attendance in the whole sample. The analysis was therefore not powered to look for sub-group effects. Furthermore, the sample size calculation does not seem to have been adjusted for the cluster design.

**Ozier study (reported in a series of papers 2011-16)**

The Ozier series report a field survey of 21 309 children attending the Kenyan trial 11 to 12 years earlier [20]. These children were too young at the time of the original trial to have received deworming treatment through the school-based programme. The analysis compares outcomes within each birth cohort from 1995 to 2001. Children aged less than 1 year living in communities where the deworming intervention had started are classified as the intervention group, and those living in communities where deworming had not yet started are classified as controls. The difference between these two groups is only that the children in the intervention group may have benefitted from decreased worm prevalence among older siblings and the community during the first year of life, while the children in the control group did not.

**Risk of bias assessment:** The field survey conducted cognitive tests on a computer generated random sample representing approximately 12 % of the eligible population. This sample covered seven annual school cohorts from 1995 to 2001. Only the 1998 and 1999 cohorts contain quasi-randomized comparisons relevant to the study question. In the 1995 and 1996 cohorts none of the children lived in areas with active deworming programmes during the first year of life; and in 2001 all the children lived in areas with active deworming programmes. Analyses across the whole sample (seven cohorts) are thus secondary, observational analyses, with unknown secular changes potentially confounding the findings (see table 3). Data collection was appropriately blinded to treatment allocation, but again data analysis was not blinded and was not guided by a pre-planned protocol. Important findings of no apparent effect on height and height-for-age were not reported in the abstract until the 2016 version (despite being one of the main a-priori hypotheses, according to communication with the authors), and although weight data were collected for 21 309 children, they were not part of the analysis and not presented.

**Credibility assessment:** In the 2016 abstract, Ozier states that exposure to the spill over effects of deworming programmes during the first year of life produced “large cognitive effects, comparable to between 0.5 and 0.8 years of schooling”. This statement is based on demonstrable effects on 2 out of 5 cognitive tests (Raven’s matrices and verbal fluency; P < 0.05), and a trend towards benefit on all 5 tests. These positive effects are taken from analyses across the whole sample, which include non-randomized data. There was a trend towards benefit on Ravens Matrices in the 1998 and 1999 annual cohorts. Clarification from the author stated that the analysis was powered to detect an improvement in tests of cognition of 0.2 standard deviations in the whole sample, and is underpowered to evaluate effects within annual cohorts.

The authors themselves explain the lack of effect on height to be related to the low worm load in young children. We consider this observation, along with the very low intensity of the intervention being tested, to question the plausibility of the stated effect.

**Uganda trial (Alderman 2006)**

The base trial for the third study [18] was conducted in Eastern Uganda by Alderman et al [21]. The intervention was implemented through Child Health Days (CHD) and comprised albendazole 400mg every six months. Fifty parishes in five districts were identified as having heavy worm loads and randomly allocated to the intervention and control arms. Over the three-year programme from 2000 to 2003, children in both groups attended 1.74 CHDs in average, with only the intervention group scheduled to be dewormed, but both groups receiving additional health services such as vaccination and health promotion. Participants were pre-school children aged between one and seven years, and deworming became routine and free for all shortly after the end of the study.

**Risk of bias assessment:** The base trial used a truly random method of allocation (a coin toss), but while deworming was the intended difference between intervention and control groups, up to 35% of those in control areas were also dewormed, from private clinics or shops (see table 2).

**Croke study 2014**

Croke utilizes a large-scale questionnaire survey conducted in Uganda seven to eight years after the end of the trial by Alderman [21]. The survey was unrelated to the base trial but covered some of the same parishes, and included 763 children who would have been aged between 1 and 7 years at the time of the base trial, and who therefore might have participated. The study compares children living within the intervention parishes of the base trial with children living in the control parishes. The difference between the two groups (ignoring migration in and out of the area) is therefore likely to be less than two additional doses of albendazole during the 3 years of the programme. The analysis reports on numeracy and literacy test outcomes.

**Risk of bias assessment:** The sampling method is reported as random, but the descriptions of sampling are inadequate to make a clear assessment of the risk of selection bias (see table 3). Data acquired through correspondence with the author reports on eleven covariates, of which the treatment group appears to have better access to water (24% of individuals compared to 3%) and private education (14% compared to 9%). The data collection process was unrelated to the deworming base trial and so unlikely to have been influenced by it, but data-analysis was not blinded. The risk of attrition bias is high with only 22 of the 50 parishes recruited by Alderman included in the sample (10 from the intervention group and 12 from the control group), and no assessment of the effects of migration. There was no pre-planned protocol.

**Credibility assessment:** Croke states that children who lived in intervention parishes during the base trial period had “test scores 0.2 to 0.4 standard deviations higher than those in control parishes”. using statistical significance at P < 0.05, this effect was not present in the raw data and only apparent after adjustment for age, gender and survey year. No formal power calculations were conducted, and the analysis is substantially underpowered to detect these effects, containing fewer than a third of the sample size calculated by Ozier. The authors found no evidence of an effect on school enrolment, but do not report this in the abstract.

**Discussion**

In summary, of the three included long-term follow-up studies, the Baird series reports possible effects on secondary school attendance and job sector choices, 9 to 11 years after a head start of 2.4 years of additional school-based deworming; the Ozier series reports possible externalities on cognitive development in children living in areas with school-based deworming during the first year of life; and Croke 2014 reports possible effects on English and maths test scores 10 to 11 years after less than two additional doses of deworming tablets during early childhood. All present these as clear evidence of benefit of deworming programmes.

Long-term studies of the effects of public health interventions are complex and difficult to do. We therefore acknowledge the hard work of the study authors and research teams. However, from our epidemiological standpoint we find substantial reason to doubt the validity and plausibility of these findings, given the information provided and the process of analysis that has been documented. As such, we believe they should be regarded as hypothesis generating, rather than reliable evidence of effects to support large scale deworming programmes in low and middle-income countries.

First, we note that these studies do not provide the evidence of cumulative effects from multiple rounds of deworming, that some have called for, and others have attributed to them [8][26]. In all three studies, most participants in both the intervention and control groups would have been ‘dewormed’ multiple times during their pre-school or school years, and the largest ‘intervention’ under evaluation was an additional 2.4 years of deworming medication in the Baird series. The majority of children in these studies would therefore be worm-free, or have reduced worm counts during much of their childhood, and consequently the consistent finding of no effect on height or weight across all three studies is unsurprising. More subtle nutritional pathways for the observed effects, such as via micronutrient status, also seem unlikely to act over such short durations.

Second, we are concerned about the selective reporting of favourable results in the abstracts, especially after multiple significance testing and post-hoc sub-group analyses. While all three papers herald from an economic discipline, we assess them against current epidemiological standards, and make no apology for that. The policy under evaluation is a public health programme, and the potential for bias exists irrespective of discipline. We do however acknowledge that some of the problems exist, at least in part, due to current norms within economics, and the reporting requirements of economic journals (such as strict word limits for abstracts). Whilst some economists may argue that the accuracy of conclusions is improved over time through the refinement and addition of new analyses, we are concerned that the process risks cumulative selective reporting, and our analysis provides some indicators that this may be the case.

Of note, none of these three studies worked to a pre-planned analytic protocol, and although these have been standard requirements within epidemiology for some time, they have only recently been recognized as important within economics [27]. This was also noted in the replication analyses of the primary trials [23][24]. Nevertheless, this does not obviate that the approach taken in the past may well lead to misleading results and conclusions, and statistical correction of multiple testing is insufficient to correct selective reporting.

The abstract to the Baird series exclusively presents the positive results, and leaves readers unaware of the multiple findings of no effect, and the conflicting findings within the analysis. For example:

* No evidence of effect on markers of nutrition (weight or height);
* No evidence of effect on multiple tests of cognition (the same tests as reported by the Ozier series);
* An increase in the need to repeat a school grade in the intervention group (an alternative explanation for the observed increase in years in school, and consistent with the finding of no overall increase in the number of grades achieved);
* No evidence of effect on secondary school attendance prior to sub-grouping by gender (only the whole sample is adequately powered to detect an effect), suggesting a potentially spurious subgroup finding;
* Little evidence of effect on secondary school attendance in females after adjustment for multiple significance testing (P = 0.084 after adjustment);
* No evidence of effect on monthly earnings.

We do know that post-hoc analyses increase the risk of type 1 errors (finding an effect when there is no effect present) [28][29]. Item 18 of the 2010 CONSORT statement for the reporting of randomized trials specifies that post-hoc analyses should be clearly labelled as such and considered as exploratory. In addition, the explanatory notes states that “Post hoc subgroup comparisons (analyses done after looking at the data) are especially likely not to be confirmed by further studies”[30].

At face value, there is a consistency of findings across the two remaining studies by Ozier and Croke. Both studies have substantial methodological and plausibility limitations which should temper their interpretation, but the observed effect after such a small deworming exposure probably deserves further consideration and should be amenable to testing through a well-designed randomized controlled trial.

More generally, there appears to be a tendency for advocates of deworming to ‘build a case’ for deworming, by drawing together evidence which supports their prior beliefs, and ignoring or dismissing the evidence that does not [31][32][9]. This ‘confirmation bias’ is common, but runs counter to current standards in transparent, evidence-informed decision making [33], and has led to the claims of these studies being cited *verbatim*, without appropriate appraisal [34].

Government ministries responsible for resource allocation, philanthropists supporting these programmes, and the public who are subject to them, require transparency about what effects could reasonably be expected. If a community in a given setting has a high prevalence of untreated worm infections, then mass-deworming programmes may well be an effective way to reach and treat a large number of children. If however, the problem is poor school attendance, or low educational attainment, then these are problems which probably require different solutions [35].

**Conclusion**

These three studies all have substantial problems in their methods and analysis, which question the credibility of the use of these studies to justify the effectiveness of deworming programmes. They help in generating hypothesis. Decisions about whether or not to implement mass treatment programmes, calculations around programme cost, and advocacy to the public, should be based on reliable estimates of effects, informed by robust evidence.

**Funding**

This work was supported through the Effective Health Care Research Programme Consortium, funded by UK aid from the UK Government for the benefit of developing countries (Grant: 5242). The views expressed in this review do not necessarily reflect UK government policy.

**Competing interests**

PG is Director of Evidence Building and Synthesis Research Consortium that receives money to increase the number of evidence-informed decisions by intermediary organizations, including WHO and national decision makers that benefit the poor in middle and low income countries. DS and SJ are employed as part of this Consortium. PG is the co-ordinator of a WHO Collaborating Centre for Evidence Synthesis for Infectious and Tropical Diseases (<http://apps.who.int/whocc/default.aspx>; UNK234) and one of the Centre’s aims is to help WHO in its role as an infomediary in communicating reliable summaries of research evidence to policy makers. PG is an author of the Cochrane review evaluating the effects of community-based deworming on health, nutrition, and school participation. PG receives support from COUNTDOWN, a grant to LSTM from DfID to promote control of neglected tropical diseases in developing countries, including soil transmitted helminths.

The funding source had no role in identifying the research topic, or in the design, data collection and analysis, decision to publish, or preparation of the manuscript.

**References**

1. de Silva NR, Brooker S, Hotez PJ, Montresor A, Engels D, Savioli L. Soil-transmitted helminth infections: updating the global picture. Trends in Parasitology 2003;19(12): 547–51
2. Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D et al. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet. 2006;367(9521): 1521-32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16679166> [Accessed 8 March 2016].
3. Baba AA, Ahmad SM, Sheikh KA. Intestinal ascariasis: the commonest cause of bowel obstruction in children at a tertiary care center in Kashmir. Pediatric Surgery International. 2009;25(12): 1099-102. Available from: doi: 10.1007/s00383-009-2486-8 [Accessed: 7 March 2016].
4. WHO. Prevention and control of schistosomiasis and soil-transmitted helminthiasis. Report of a WHO Expert Committee, Technical Report Series, No. 912. Geneva, World Health Organization; 2002. Available from: <http://apps.who.int/iris/bitstream/10665/42588/1/WHO_TRS_912.pdf> [Accessed: 7 March 2016].
5. Charles Mwandawiro. A new perspective on the War on Worms. Available from: <http://www.impatientoptimists.org/Posts/2015/09/A-New-Perspective-on-the-War-on-Worms#.Vukw52SLTZu> [Accessed: 15 March 2016].
6. Taylor-Robinson D, Maayan N, Soares-Weiser K, Donegan S, Garner P. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin, and school performance. Cochrane Database of Systematic Reviews. 2015(7). Available from: doi: 10.1002/14651858.CD000371.pub6 [Accessed: 8 October 2015].
7. Walson JL. Don't Shoot the Messenger. PLoS Neglected Tropical Diseases. 2015;9(10): e0004166. Available from: doi:10.1371/journal.pntd.0004166 [Accessed: 6 November 2015].
8. Montresor A, Addiss D, Albonico M, Ali SM, Ault SK, Gabrielli AF, et al. Methodological Bias Can Lead the Cochrane Collaboration to Irrelevance in Public Health Decision-Making. PLoS Neglected Tropical Diseases. 2015;9(10): e0004165. Available from: doi: 10.1371/journal.pntd.0004165 [Accessed: 31 January 2016].
9. Ahuja A, Baird S, Hicks JH, Kremer M, Miguel E, Powers S. When Should Governments Subsidize Health? The Case of Mass Deworming. The World Bank Economic Review. 2015; 1-16. Available from: doi:10.1093/wber/lhv008 [Accessed: 14 July 2015].
10. Evidence Action. Deworm the World Initiative. Available from: <http://www.evidenceaction.org/dewormtheworld> [Accessed: 3 March 2016].
11. Abdul Latif Jameel Poverty Action Lab. Deworming: A Best Buy for Development, J-PAL Policy Bulletin. 2012. Available from: https://www.povertyactionlab.org/sites/default/files/publications/2012.3.22-Deworming.pdf [Accessed: 3 March 2016].
12. WHO. Deworming campaign improves child health, school attendance in Rwanda. Available from: <http://www.who.int/features/2015/rwanda-deworming-campaign/en/> [Accessed: 3 March 2016].
13. Hicks JH, Kremer M, Miguel E. The Case for Mass Treatment of Intestinal Helminths in Endemic Areas. PLoS Neglected Tropical Diseases. 2015;9(10): e0004214. Available from: doi:10.1371/journal.pntd.0004214 [Accessed: 6 November 2015].
14. Give well. Deworm the world initiative, led by Evidence Action. Available from: <http://www.givewell.org/international/top-charities/deworm-world-initiative> [Accessed: 14 March 2016].
15. Higgins JP, Altman DG, Sterne JA. Assessing risk of bias in included studies. In Higgins JP, Green S. (eds.) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration; 2011. Available from: <http://handbook.cochrane.org/> [Accessed: 7 March 2016].
16. Noble WS. How does multiple testing correction work? Nature Biotechnology 2009; 27: 1135-1137.
17. Baird S, Hicks JH, Kremer M, Miguel E. Worms at work: Long-run impacts of a child health investment. 2016. Available from: <http://www.nber.org/papers/w21428> [Accessed: 12 January 2016].
18. Croke K. The long run effects of early childhood deworming on literacy and numeracy: Evidence from Uganda. 2014. Available from: <http://scholar.harvard.edu/files/kcroke/files/ug_lr_deworming_071714.pdf> [Accessed: 18 August 2016].
19. Ozier O. Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming. 2016. Available from: <http://economics.ozier.com/owen/papers/ozier_early_deworming_20160727.pdf> [Accessed: 18 August 2016].
20. Miguel E, Kremer M. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. Econometrica. 2004;72(1): 159–217. Available from: <http://www.jstor.org/stable/3598853> [Accessed: 31 August 2015].
21. Alderman H, Konde-Lule J, Sebuliba I,Bundy D, Hall A. Effect on weight gain of routinely giving albendazole to preschool children during child health days in Uganda: cluster randomized controlled trial. British Medical Journal. 2006. Available from: doi: 10.1136/bmj.38877.393530.7C [Accessed: 26 January 2016].
22. Bleakley H. Disease and Development: Evidence from Hookworm Eradication in the American South. Quarterly Journal of Economics. 2007;122(1):73–117. Available from: doi: 10.1162/qjec.121.1.73 [Accessed: 30 October 2015].
23. Aitken AM, Davey C, Hargreaves JR, Hayes RJ. Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication. International Journal of Epidemiology. 2015;1-9. doi: 10.1093/ije/dyv127
24. Davey C, Aitken AM, Hayes RJ, Hargreaves JR. Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a statistical replication of a cluster quasi-randomized stepped wedge trial. International Journal of Epidemiology. 2015;1-12. doi: 10.1093/ije/dyv128
25. Hicks JH, Kremer M, Miguel E. Commentary: Deworming externalities and schooling impacts in Kenya: a comment on Aitken et al. (2015) and Davey et al. (2015). International Journal of Epidemiology. 2015; 1-4. doi: 10.1093/ije/dyv129
26. Medley GF, Hollingsworth TD. MDA helminth control: more questions than answers. The lancet global health. 2015;3(10):e583-4. Available from: <http://wrap.warwick.ac.uk/id/eprint/69715> [Accessed: 10 March 2016].
27. Miguel E, Camerer C, Casey K, Cohen J, Esterling KM, Gerber A, et al. Promoting Transparency in Social Science Research. Science. 2014;343: 30-1.
28. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. British Medical Journal. 1995;310(6973): 170. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2548561/ [Accessed: 7 March 2016].
29. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. Health Technology Assessment. 2001;5(33). Available from: http://www.journalslibrary.nihr.ac.uk/\_\_data/assets/pdf\_file/0007/64978/FullReport-hta5330.pdf [Accessed: 1 March 2016].
30. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. British Medical Journal. 2010;340: c869. Available from: doi: 10.1136/bmj.c869 [Accessed: 7 March 2016].
31. WHO. The evidence is in: deworming helps meet the Millennium Development Goals. Available from: http://apps.who.int/iris/bitstream/10665/68876/1/WHO\_CDS\_CPE\_PVC\_2005.12.pdf [Accessed: 6 November 2015].
32. Albonico M, Allen H, Chitsulo L, Engels D, Gabrielli A-F, Savioli L. Controlling Soil-Transmitted Helminthiasis in Pre-School-Age Children through Preventive Chemotherapy. PLoS Neglected Tropical Diseases. 2008;2(3): e126. Available from: doi:10.1371/ journal.pntd.0000126 [Accessed: 7 March 2016].
33. Oxman AD, Lavis JN, Lewin S, Fretheim A. SUPPORT Tools for evidence-informed health Policymaking (STP) I: What is evidence-informed policymaking? Health Research Policy and Systems. 2009;7(Suppl 1):S1. Available from doi: 10.1186/1478-4505-7-S1-S1 [Accessed: 8 March 2016].
34. Lord CG, Ross L, Lepper MR. Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence. Journal of Personality and Social Psychology. 1979; 37(11):1098-109.
35. Krishnaratne S, White H, Carpenter E. Quality education for all children? What works in education in developing countries, Working Paper 20. New Delhi: International Initiative for Impact Evaluation (3ie); 2013.