RESEARCH ARTICLE

Toward Measuring *Schistosoma* Response to Praziquantel Treatment with Appropriate Descriptors of Egg Excretion

Piero L. Olliaro1,2,*, Michel Vaillant3, Aïssatou Diawara4, Jean T. Coulibaly5,6,7,8, Amadou Garba8,9,10, Jennifer Keiser7,8, Charles H. King11,12, Stefanie Knopp8,9,13, Aly Landouré14, Eliézer K. N’Goran5,6, Giovanna Raso8,9, Alexandra U. Scherrer15, José Carlos Sousa-Figueiredo16, Katarina Stete8,9,17, Xiao-Nong Zhou18,19, Jürg Utzinger8,9

1 UNICEF/UNDP/World Bank/WHO Special Programme on Research and Training in Tropical Diseases (TDR), World Health Organization, Geneva, Switzerland, 2 Centre for Tropical Medicine, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom, 3 Centre of Competence for Methodology and Statistics (CCMS), Luxembourg Institute of Health (LIH), Strassen, Luxembourg, 4 Department of Biology, Division of Science and Mathematics, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, 5 Unité de Formation et de Recherche Biosciences, Université Félix Houphouët-Boigny, Abidjan, Côte d’Ivoire, 6 Centre Suisse de Recherches Scientifiques en Côte d’Ivoire, Abidjan, Côte d’Ivoire, 7 Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland, 8 University of Basel, Basel, Switzerland, 9 Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland, 10 Réseau International Schistosomoses, Environnement, Aménagement et lutte (RISEAL-Niger), Niamey, Niger, 11 Center for Global Health and Diseases, Case Western Reserve University, Cleveland, Ohio, United States of America, 12 Schistosomiasis Consortium for Operational Research and Evaluation, University of Georgia, Athens, Georgia, United States of America, 13 Wolfson Wellcome Biomedical Laboratories, Department of Life Sciences, Natural History Museum, London, United Kingdom, 14 Institut National de Recherche en Santé Publique, Bamako, Mali, 15 Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland, 16 Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, 17 Center for Infectious Diseases and Travel Medicine, Department of Medicine, University Hospital Freiburg, Freiburg im Breisgau, Germany, 18 National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai, People’s Republic of China, 19 Key Laboratory of Parasite and Vector Biology of the Chinese Ministry of Health, WHO Collaborating Center for Malaria, Schistosomiasis and Filariasis, Shanghai, People’s Republic of China

* olia@g.who.int

Abstract

**Background**

The control of schistosomiasis emphasizes preventive chemotherapy with praziquantel, which aims at decreasing infection intensity and thus morbidity in individuals, as well as transmission in communities. Standardizing methods to assess treatment efficacy is important to compare trial outcomes across settings, and to monitor program effectiveness consistently. We compared customary methods and looked at possible complementary approaches in order to derive suggestions for standardizing outcome measures.

**Methodology/Principal Findings**

We analyzed data from 24 studies conducted at African, Asian, and Latin American sites, enrolling overall 4,740 individuals infected with *Schistosoma mansoni*, *S. haematobium*, or
S. japonicum, and treated with praziquantel at doses of 40–80 mg/kg. We found that group-based arithmetic and geometric means can be used interchangeably to express egg reduction rates (ERR) only if treatment efficacy is high (>95%). For lower levels of efficacy, ERR estimates are higher with geometric than arithmetic means. Using the distribution of individual responses in egg excretion, 6.3%, 1.7% and 4.3% of the subjects treated for S. haematobium, S. japonicum and S. mansoni infection, respectively, had no reduction in their egg counts (ERR = 0). The 5th, 10th, and 25th centiles of the subjects treated for S. haematobium had individual ERRs of 0%, 49.3%, and 96.5%; the corresponding values for S. japonicum were 75%, 99%, and 99%; and for S. mansoni 18.2%, 65.3%, and 99.8%. Using a single rather than quadruplicate Kato-Katz thick smear excluded 19% of S. mansoni-infected individuals. Whilst the effect on estimating ERR was negligible by individual studies, ERR estimates by arithmetic means were 8% lower with a single measurement.

Conclusions/Significance

Arithmetic mean calculations of Schistosoma ERR are more sensitive and therefore more appropriate to monitor drug performance than geometric means. However, neither are satisfactory to identify poor responders. Group-based response estimated by arithmetic mean and the distribution of individual ERRs are correlated, but the latter appears to be more apt to detect the presence and to quantitate the magnitude of suboptimal responses to praziquantel.

Author Summary

To identify whether a person is infected with parasitic worms, stool or urine samples are examined for worm eggs. The drug praziquantel is used against the parasitic disease schistosomiasis. However, there is no definitive agreement as to how the efficacy of praziquantel is best expressed. We put together a database from various studies of the efficacy of praziquantel against schistosomiasis. Efficacy was measured using customary methods: cure rate (CR: percentage of people with eggs in their stool/urine before treatment who became egg-negative after treatment); and egg reduction rate (ERR; percentage reduction in the number of eggs in the stool/urine after treatment, where the mean number of eggs from all people treated is calculated using either geometric or arithmetic means). We found that arithmetic and geometric means can be used interchangeably only if treatment efficacy is very high; arithmetic means are more sensitive to capture drops in efficacy expressed by ERR. A valid complement for drug efficacy monitoring is to study the distribution of individual responses in egg excretion that allows identifying in a single measure both those who had an adequate response to treatment and those who respond less well; e.g., the 5% of the patients with the lowest ERRs.

Introduction

Schistosomiasis is a parasitic disease caused by blood flukes of the genus Schistosoma. The three main species infecting humans are S. haematobium (causing urogenital schistosomiasis), S. japonicum, and S. mansoni (the latter two responsible for intestinal schistosomiasis) [1]. The backbone of the global strategy for controlling the morbidity caused by schistosomiasis is the
periodic administration of single-dose oral praziquantel (usually given at 40 mg/kg body weight). This strategy is termed “preventive chemotherapy”, whereby praziquantel is administered without prior diagnosis [2,3] to entire communities or target groups, most importantly school-aged children [4–6], depending on the level of endemicity [7,8].

How efficacy of antischistosomal drugs (and anthelmintic treatments in general) should be measured has been, and still is, a matter of debate in the research and disease control communities. One limitations of the current treatment outcome measure—i.e., parasite egg excretion—is that it is a proxy for drug effects on adult worms, which could also be confounded by various factors, including facultative temporary cessation of excretion by the adult worm [3]. Other, more direct but not yet widely used methods of worm vitality are the detection of specific antigens, like the circulating cathodic antigen (CCA) and the circulating anodic antigen (CAA) [9,10]. Clinical trials have traditionally used cure rates (CRs) as the main drug efficacy endpoint, and expressed results as the proportion of infected individuals who convert to a negative stool or urine sample post-treatment [11–13]. However, the World Health Organization (WHO) recently issued guidelines that recommend the egg reduction rate (ERR) as the primary outcome measure, especially when assessing programmatic treatment effectiveness [3]. This entails a quantitative diagnostic test based on the microscopic detection and enumeration of parasite eggs in small amounts of stool (usually 41.7 mg) and the estimation of the number of eggs per 1 g (EPG) of feces (S. japonicum or S. mansoni) or per 10 ml of urine (S. haematomium). The ERR measures the overall effect of treatment on the entire group of infected subjects treated (ignoring individual variability) and is expressed as the ratio between the mean of the pre- and post-treatment egg counts [3]. ERR is considered more suitable than CR to assess the impact of preventive chemotherapy on morbidity (which is commensurate to infection intensity) in the context of continuous risk of reinfection and in view of the low sensitivity of current diagnostic methods [12,14].

Which type of means (e.g., geometric mean (GM) or arithmetic mean (AM)) should be used to express treatment outcomes against helminthiases at the community level is an additional subject of debate [15,16]. Thus far, studies of treatment efficacy have predominantly reported results using GM egg counts [15] but recently, the use of AM egg counts has been advocated [17,18]. The issue is that egg counts are not normally distributed, even after logarithmic transformation, which would call for using GM [19]. However, GM hide extreme values (e.g., (i) a small proportion of individuals disproportionally contributing to total egg excretion, and (ii) individuals who do not respond to treatment), which are important when assessing the effects of interventions, and which are better captured by using AM.

A further complication when dealing with different studies is the diversity in methodologies, in particular: (i) which diagnostic method is used (e.g., single or multiple urine filtration for detection of S. haematobium eggs) [20] or the Kato-Katz technique for detection of S. japonicum or S. mansoni eggs in fecal samples, whose sensitivity depends upon the baseline infection intensity, the number of thick smears from a single sample, and the number of stool specimens examined [21–24]; and (ii) how many weeks post-treatment effects are measured, which also depends on Schistosoma species [25,26].

The objectives of this paper were to compare customary methods to assess the efficacy of praziquantel for treating schistosomiasis and to explain differences; to identify possible alternative approaches to express treatment effects on egg excretion; and to verify whether the size of treatment effects for intestinal schistosomiasis change when measured with a single or quadruplicate Kato-Katz thick smears. The overall aim of these analyses was to derive suggestions for standardizing outcome measures in future drug efficacy studies. For this purpose, we combined and analyzed available data from various studies where praziquantel was used to treat infections with different Schistosoma species.
**Methods**

**Ethics Statement**

All studies selected for the current secondary analyses had been approved by the relevant institutional review boards and ethics committees, and were conducted according to international ethics standards (for details, see individual publications [25–36]). Data received from the individual studies were completely anonymized.

**Datasets Analyzed**

We built a common database from 24 studies including 4,740 individuals assigned to three different treatment groups who had received either 40 mg/kg (18 studies, 3,713 individuals), 60 mg/kg (five studies, 690 individuals), or 80 mg/kg (one study, 337 individuals) praziquantel against *S. haematobium* [25,27–29,34,36], *S. japonicum* [31,32], and *S. mansoni* [21,23,26–28,30,31,33,35].

The main study characteristics, including *Schistosoma* species, praziquantel dose, age of participants, time-point of treatment follow-up, and diagnostic approach, are summarized in Table 1. Studies enrolled a total of 4,740 individuals, of whom 2,633 (55.5%) were infected with *S. haematobium*, 1,804 (38.1%) with *S. mansoni*, and the remaining 303 (6.4%) with *S. japonicum*. Studies generally enrolled children and adolescents except one study in the People’s Republic of China [32] and another one in Côte d’Ivoire [33]. The praziquantel dose was 40 mg/kg in 17 studies, 60 mg/kg in five studies, and 80 mg/kg in one study. Follow-up was within 3 weeks in 13 studies, four weeks in two studies, within 2 months in six studies, and longer in the remaining three studies.

For the detection of *S. haematobium* infection, two diagnostic approaches were employed: (i) a single urine filtration slide in five studies (one of them was carried out on the same sample but at two different time points) [25,26,28,29,34]; and (ii) duplicate urine filtration slides in one study [36]. For the diagnosis of *S. japonicum*, duplicate Kato-Katz thick smears from each of two stool specimens were subjected to microscopic examination in all studies. For the detection of *S. mansoni*, the most common diagnostic approach was duplicate Kato-Katz thick smears from each of two stool specimens. In one study, a single Kato-Katz thick smear was performed on five samples at baseline and triplicate Kato-Katz thick smears for four samples at follow-up [35,37].

**Efficacy Outcomes and Calculations**

Treatment response was assessed both at the overall group and the individual level. The AM and GM EPG values were calculated at pre- and post-treatment for *S. mansoni* and *S. japonicum* by multiplying the individual fecal egg counts (FECs) obtained by a single Kato-Katz thick smear (41.7 mg) by a factor of 24. For *S. haematobium*, egg counts are presented as eggs per 10 ml of urine. Drug efficacy was expressed as ERR and CR.

ERR (arithmetic (ERR\text{AM}) or geometric (ERR\text{GM})) was calculated as the ratio of the difference between the (arithmetic or geometric) means of the pre- and post-treatment EPG or eggs per 10 ml urine to the pre-treatment (arithmetic or geometric) mean EPG or eggs per 10 ml urine: $ERR = \frac{[(mean\ egg\ count_{pre-treatment} - mean\ egg\ count_{post-treatment}) / mean\ egg\ count_{pre-treatment}]}{100}$
GM egg counts were calculated as follows: 
\[
\text{exp} \left( \frac{\sum_{i=1}^{n_{jk}} \ln (x_{ijk} + c)}{n_{jk}} \right) - c,
\]
where \(x_{ijk}\) is the observed egg count for individual host \(i\), Schistosoma species \(j\), and study \(k\); \(n_{jk}\) is the number of hosts who provided a (pre- and post-treatment) sample for determination of infection intensity for each Schistosoma species and study, and \(c\) is a constant added to each count to allow inclusion of zero counts (negative test) [38]. Confidence intervals (CIs) for the ERR (calculated with AM and GM) were determined by using a bootstrap resampling method (with replacement) over 1,000 replicates and expressed as a univariate calculation of the 2.5th and 97.5th percentiles.
Individual ERR was calculated as the ratio of the difference between the pre- and post-treatment EPG or eggs per 10 ml urine to the pre-treatment EPG or eggs per 10 ml urine multiplied by 100.

CRs and 95% binomial CIs were the percentage of infected individuals negative for Schistosoma (in their urine or stool) at post-treatment follow-up. The distribution of individual responses in egg excretion was categorized as (i) negative (corresponding to CR), (ii) reduction, (iii) no change or increase, and further expressed in centiles to quantitate the fraction of poor responders.

Comparing Methods of Assessing Drug Efficacy

We compared (i) the ERRAM versus ERRGM and (ii) the CR versus ERR. The results were presented graphically in modified L’Abbé plots with 95% CIs for both comparisons, and additionally in Bland and Altman plots for ERR using AM and GM. The coefficient of determination ($R^2$) was also calculated.

A linear mixed model of the difference of the ERRs calculated as GM and AM ($\Delta\text{ERR}_{g,a}$) was developed to estimate which parameter could better predict the difference between the two ERR calculations, with the average of the two mean ERRs (GM and AM) set as covariate. Such a model was extrapolated from the Bland and Altman regression by further including predictive factors. In order to evaluate the effect of different factors on $\Delta\text{ERR}_{g,a}$, the ERR was calculated on the different strata defined by the combination of the categories of the following parameters: age, sex, treatment dose, and Schistosoma species. The same age categories were defined across all studies. The linear mixed model was estimated including these parameters as independent factors. 95% CIs of the difference between the two ERR calculations were calculated by using Tango’s score confidence interval which was shown in the literature to outperform other calculation methods in the case of correlated proportions [39,40].

Modeling was carried out through a shrinkage method of variable selection. Variables were first selected using the ElasticNet procedure, which is mixing a least absolute shrinkage and selection operator (LASSO) procedure and ridge regression [41]. Subsequently, a variance-covariance matrix structure was selected among unstructured, variance components, autoregressive, compound symmetry and Toeplitz structures, which minimized the Akaike information criterion corrected (AICC) for finite sample size. Post-hoc tests on each parameter were carried out with a Tukey adjustment. Pairwise differences in least square means (LSM) were calculated for fixed values of the average of the AM and GM ERRs (for the range 70–99%), thereby evaluating the influence of exogenous parameters on the bias between the two ERRs. This bias was evaluated in the Bland and Altman method by regressing the difference of the two methods by their mean.

All tests were two-tailed; a p-value of 5% was deemed significant. Only studies with treatment follow-up examination done within 90 days were included in the models (to minimize the effect of reinfection after praziquantel administration). Calculations and analyses were performed by using SAS system version 9.3 (SAS Institute, Cary, NC, United States of America).

Comparing Efficacy Assessments Based on Single versus Quadruplicate Kato-Katz Thick Smears for S. mansoni

For S. mansoni, we compared single (using the first thick smear on the first stool specimen) versus quadruplicate (using four thick smears of the same stool specimen) Kato-Katz thick smears for expressing CR and ERR (with AM and GM). In this sub-analysis, we only included individuals infected with S. mansoni whose first slide on the first fecal sample was positive.
Within this population, we calculated and compared the overall AM and GM pre- and post-treatment FECs and the respective ERRs and CRs based on single and quadruplicate Kato-Katz thick smears with 95% CIs calculated by boot-strapping. The results of the individual studies were presented graphically in Bland and Altman plots and in modified L’Abbé plots with 95% CIs, and the coefficient of determination ($R^2$) was calculated.

**Results**

**Study Characteristics**

The distributions of the raw egg counts at baseline by *Schistosoma* species for each study (including AMs and GMs) are presented in Fig 1. Efficacy outcomes were analyzed on a total of 4,375 individuals with pre- and post-treatment egg counts. Among them, 2,365 (54.1%) were infected with *S. haematobium*, 1,708 (39.0%) with *S. mansoni*, and the remaining 300 (6.9%) with *S. japonicum*. Details of baseline FECs and drug efficacy outcomes, including group-based means and individual responses by study and *Schistosoma* species, are presented in Table 2.

**Comparing ERRs Obtained with AM versus GM**

$\text{ERR}_{AM}$ ranged from 17.0% to 99.8% and $\text{ERR}_{GM}$ from 50.7% to 99.8% for *S. haematobium*; 81.5–100% ($\text{ERR}_{AM}$) and 95.9–100% ($\text{ERR}_{GM}$) for *S. mansoni*; 83.5–99.9% ($\text{ERR}_{AM}$) and 86.9–99.9% ($\text{ERR}_{GM}$) for *S. japonicum*. The 95% CIs estimated by boot-strapping tended to be wider with AM compared to GM. Among the 24 studies included in our analyses, six had an ERR <90% by AM (two if restricted to studies with 3 weeks’ follow-up), but only two by GM.

The modified L’Abbé plot (Fig 2A) indicates that ERRs tend to be higher when calculated using GM. The $R^2$ of the linear regression showed a strong linear correlation for *S. japonicum* ($R^2 = 1.00$, three studies) and *S. haematobium* ($R^2 = 0.88$, eight studies), but a weaker linear correlation for *S. mansoni* ($R^2 = 0.46$, 13 studies).

**Modeling Group-Based ERRs**

The mixed linear model found a significant relationship between the difference between $\text{ERR}_{AM}$ and $\text{ERR}_{GM}$ and their mean value. However, introducing the variables identified as predictors by the ElasticNet procedure, coupled with model averaging based on 1000 replicates (sampling with replacement)–mean baseline epg, year of study and species–rendered this relation non-significant (meaning that none of these covariates could explain the differences between $\text{ERR}_{AM}$ and $\text{ERR}_{GM}$.)

LSM pairwise comparisons and ERR (individually for AM and GM) models showed a significantly better consistency between ERRs calculated with AM and GM for *S. haematobium* than for *S. mansoni*. Study participants’ age was found to have an effect only for $\text{ERR}_{GM}$ (higher for school-aged children and adults than for preschool-aged children; S1 Table).

**Distribution of Individual Treatment Responses in Egg Excretion**

Group means and individual responses are presented in Table 2. Individual responses are also displayed graphically in Fig 3 (panel A as bar diagrams; panel B as centile distributions). Overall, 6.3% (ranging in individual studies from 0 to 39%), 1.7% (0–5%); and 4.3% (0–11%) of the subjects treated for *S. haematobium*, *S. japonicum*, and *S. mansoni* infection, respectively, had no reduction in their egg counts (ERR = 0). The $5^{th}$, $10^{th}$, and $25^{th}$ centiles of the subjects treated for *S. haematobium* had individual ERRs of 0% (ranging in individual studies from 0 to 99.7%), 49.3% (0–100%), and 96.5% (0–100%); the corresponding values for *S. japonicum* were
75.0% (0–97.4%), 99.0% (25.6–98.8%), and 99.0% (96.8–99%), and for S. mansoni 18.2% (0–97.4%), 65.3% (0–99.4%), and 99.8% (0–99.8%).

The centile distribution of iERRs in studies with ERR$_{AM}$ $<$ 90% was shifted to the right and clearly distinct from those with ERR$_{AM}$ $\geq$ 90% (Fig 4).

When considering only studies assessing outcomes within a maximum of 28 days (13 studies evaluated drug efficacy at 21 days, two at 28 days), only two (both treating S. mansoni with 40 mg/kg praziquantel with 21-day follow-up [27,30]) had an ERR$_{AM}$ $<$ 90%; in terms of individual responses, the 30th and 36th centile, respectively had ERRs $<$ 90%, and 10.8% and 7.1% of patients, respectively had no change in their egg counts (ERR = 0). Three additional studies, all with ERR AM $>$ 90%, had individual ERRs $<$ 90% in the 13th centile: one study on S.
Table 2. Summary of study outcomes, group-based mean responses, and individual responses by study and species.

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose</th>
<th>N</th>
<th>aMD0</th>
<th>cMD0</th>
<th>ERR</th>
<th>CR</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>ERR</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>ERR</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>ERR</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>ERR</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Côte d'Ivoire, 2011</td>
<td>40</td>
<td>18</td>
<td>13.44</td>
<td>4.34</td>
<td>99.38</td>
<td>99.00</td>
<td>100.00</td>
<td>98.80</td>
<td>96.81</td>
<td>100.00</td>
<td>94.4%</td>
<td>83.9%</td>
<td>100%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niger, 2007</td>
<td>40</td>
<td>377</td>
<td>33.09</td>
<td>11.73</td>
<td>96.50</td>
<td>95.09</td>
<td>97.60</td>
<td>96.33</td>
<td>95.58</td>
<td>97.00</td>
<td>69.2%</td>
<td>64.6%</td>
<td>73.9%</td>
<td>3.63%</td>
<td>4.52%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niger, 2009</td>
<td>40</td>
<td>401</td>
<td>95.99</td>
<td>25.52</td>
<td>95.73</td>
<td>92.23</td>
<td>98.14</td>
<td>96.72</td>
<td>95.90</td>
<td>97.30</td>
<td>51.4%</td>
<td>46.5%</td>
<td>56.3%</td>
<td>3.24%</td>
<td>4.23%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d'Ivoire, 2000</td>
<td>80</td>
<td>331</td>
<td>76.19</td>
<td>24.75</td>
<td>99.82</td>
<td>99.68</td>
<td>99.93</td>
<td>99.76</td>
<td>99.61</td>
<td>99.87</td>
<td>91.5%</td>
<td>88.5%</td>
<td>94.5%</td>
<td>0.30%</td>
<td>0.22%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mali, 2006</td>
<td>40</td>
<td>829</td>
<td>218.86</td>
<td>67.13</td>
<td>92.48</td>
<td>90.29</td>
<td>94.30</td>
<td>97.41</td>
<td>96.90</td>
<td>97.82</td>
<td>59.1%</td>
<td>55.8%</td>
<td>62.5%</td>
<td>5.43%</td>
<td>7.78%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mali, 2009</td>
<td>40</td>
<td>162</td>
<td>67.01</td>
<td>16.57</td>
<td>18.94</td>
<td>-19.93</td>
<td>45.07</td>
<td>58.08</td>
<td>39.31</td>
<td>71.80</td>
<td>67.0%</td>
<td>62.5%</td>
<td>72.5%</td>
<td>38.89%</td>
<td>56.79%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d'Ivoire, 2010</td>
<td>40</td>
<td>87</td>
<td>85.53</td>
<td>27.29</td>
<td>97.83</td>
<td>95.88</td>
<td>99.31</td>
<td>98.58</td>
<td>97.58</td>
<td>99.32</td>
<td>80.5%</td>
<td>72.1%</td>
<td>88.8%</td>
<td>13.72%</td>
<td>4.60%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kenya, 1990</td>
<td>40</td>
<td>162</td>
<td>238.94</td>
<td>69.42</td>
<td>55.20</td>
<td>41.15</td>
<td>67.96</td>
<td>90.67</td>
<td>85.95</td>
<td>94.10</td>
<td>65.1%</td>
<td>58.1%</td>
<td>68.1%</td>
<td>14.20%</td>
<td>31.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philippines, 2007</td>
<td>40</td>
<td>101</td>
<td>339.57</td>
<td>33.02</td>
<td>96.64</td>
<td>94.03</td>
<td>98.07</td>
<td>96.16</td>
<td>95.32</td>
<td>96.54</td>
<td>61.3%</td>
<td>53.1%</td>
<td>63.3%</td>
<td>6.25%</td>
<td>49.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China, 2007</td>
<td>40</td>
<td>100</td>
<td>4.48</td>
<td>1.42</td>
<td>83.55</td>
<td>48.80</td>
<td>96.43</td>
<td>86.95</td>
<td>78.28</td>
<td>93.62</td>
<td>78.0%</td>
<td>69.9%</td>
<td>86.1%</td>
<td>1.15%</td>
<td>31.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philippines, 2011</td>
<td>40</td>
<td>94</td>
<td>378.41</td>
<td>238.95</td>
<td>99.66</td>
<td>99.42</td>
<td>99.87</td>
<td>99.73</td>
<td>99.55</td>
<td>99.86</td>
<td>90.0%</td>
<td>86.6%</td>
<td>93.4%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mauritania, 2007</td>
<td>40</td>
<td>92</td>
<td>378.41</td>
<td>238.95</td>
<td>99.66</td>
<td>99.42</td>
<td>99.87</td>
<td>99.73</td>
<td>99.55</td>
<td>99.86</td>
<td>90.0%</td>
<td>86.6%</td>
<td>93.4%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanzania, 2007</td>
<td>40</td>
<td>92</td>
<td>378.41</td>
<td>238.95</td>
<td>99.66</td>
<td>99.42</td>
<td>99.87</td>
<td>99.73</td>
<td>99.55</td>
<td>99.86</td>
<td>90.0%</td>
<td>86.6%</td>
<td>93.4%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d'Ivoire, 2004</td>
<td>40</td>
<td>177</td>
<td>168.86</td>
<td>56.62</td>
<td>83.70</td>
<td>72.85</td>
<td>91.54</td>
<td>96.89</td>
<td>94.15</td>
<td>98.07</td>
<td>62.1%</td>
<td>55.0%</td>
<td>69.3%</td>
<td>7.34%</td>
<td>84.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d'Ivoire, 2007</td>
<td>40</td>
<td>49</td>
<td>111.58</td>
<td>31.74</td>
<td>87.06</td>
<td>49.89</td>
<td>99.17</td>
<td>96.12</td>
<td>92.59</td>
<td>99.54</td>
<td>74.6%</td>
<td>67.6%</td>
<td>80.3%</td>
<td>4.05%</td>
<td>40.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d'Ivoire, 2008</td>
<td>40</td>
<td>85</td>
<td>91.79</td>
<td>27.20</td>
<td>90.06</td>
<td>86.97</td>
<td>96.35</td>
<td>98.33</td>
<td>96.91</td>
<td>98.99</td>
<td>72.9%</td>
<td>67.5%</td>
<td>76.4%</td>
<td>2.35%</td>
<td>43.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Côte d'Ivoire, 2009</td>
<td>40</td>
<td>255</td>
<td>255.60</td>
<td>95.57</td>
<td>93.62</td>
<td>88.72</td>
<td>96.67</td>
<td>98.58</td>
<td>96.10</td>
<td>98.99</td>
<td>72.9%</td>
<td>67.5%</td>
<td>76.4%</td>
<td>2.35%</td>
<td>43.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N is the number assessed (egg counts at pre- and post-treatment) out of the total number enrolled. Group means by species are calculated by pooling all studies on that species. Group confidence interval (CI) are calculated by boot-strapping. Egg reduction rate (ERR) = 0 is the proportion of subjects who had no reduction in egg counts; the proportions in the 5th, 10th, and 25th centile columns represent the individual ERR achieved by that centile.
*haematobium* treated with 40 mg/kg and 21-day follow-up [27], and two studies on *S. mansoni* treated with 40 or 60 mg/kg with 28-day follow-up [35,37]. In these studies, 3.2%, 3.5% and 2.4%, respectively of individuals had no decrease in egg counts. Across these studies, ERRAM and iERR were highly correlated (R^2 0.95).

Comparing CR versus ERR

For all three *Schistosoma* species, the CR was systematically lower than ERR, regardless of whether AM or GM was employed, except in a single study [29]. ERR >90% corresponded to CRs ranging from 51.4% to 99%. Only when ERRs were very high (range: 97.7–100%) there was a good agreement between both indicators. The CR ranged from 82.1% to 100% (Fig 5).

Multiple versus Single Kato-Katz Thick Smear Examination for *S. mansoni*

Among the 1,435 individuals enrolled in the studies who were found positive for *S. mansoni* eggs in their stool based on quadruplicate Kato-Katz thick smears, 1,167 (81.3%) were diagnosed positive on the first Kato-Katz thick smear. In this subset, we found that the results in the individual studies were highly correlated (Fig 6A–6B; R^2 = 0.95 for AM and R^2 = 0.86 for GM). The same number of studies (n = 2) had AM ERR of less than 90% with either approach. On aggregate, the FECs at baseline were lower with quadruplicate than with a single Kato-Katz thick smear when using either AM (1,046 EPG versus 2,617 EPG) or GM (342 EPG versus 575 EPG); the ERR estimates were comparable only by GM (99.5% versus 99.8%) whereas by
AM, the estimate was lower with quadruplicate than single examinations (86.9% versus 94.9%) (Table 3). In order to verify if using a single Kato-Katz thick smear selected for a different sample, we compared the baseline FECs and ERRs assessed with quadruplicate Kato-Katz thick smears on the overall sample of these studies (n = 1,435) to that of this subgroup with the first Kato-Katz thick smear positive (n = 1,167): while the overall average baseline FECs was approximately threefold lower for the latter, there was no difference in ERR between the two groups for either AM and GM.

In contrast to ERRs, CRs derived from single versus quadruplicate Kato-Katz thick smears were poorly correlated (R² = 0.50) (Fig 7). With few exceptions, the CR assessed from single Kato-Katz thick smear (CR = 92.0%) was consistently higher than that from quadruplicate Kato-Katz thick smears (CR = 82.6%) in the individual studies and the difference between the two approaches increased with decreasing efficacy (Fig 7A and 7B). The overall CR was 88.0% applying a single and 79.0% using quadruplicate Kato-Katz thick smears (Table 4).

Fig 3. Distribution of individual ERR. A. Bar diagrams, B. Centile distribution.
Preventive chemotherapy with praziquantel is the current backbone of the global strategy to control the morbidity caused by schistosomiasis in high-endemicity areas [2,4,6,8]. Monitoring of praziquantel efficacy should accompany schistosomiasis control programs in order to identify promptly suboptimal responders; to that effect, the WHO has issued standard procedures for control programs based on one single measurement around 3 weeks post-treatment [3]. At the same time, more work is needed to improve the current evidence-base for decision-making: in order to provide reliable information, it is important to agree on a robust statistical approach to assess drug efficacy, especially in clinical trials, and to use standardized, quality-controlled diagnostic methods that are comparable from one setting to another [18,42]. To date, both CR and ERR (the latter based on either AM or GM), are used for assessing drug efficacy. The stool and urine sampling and diagnostic approaches vary across studies, and these issues have important ramifications for drug efficacy estimates.

Fig 4. Centile distribution of individual ERRs by species for all studies (A) and for studies with assessment within 28 days divided by ERR_{AM} ≥90% and <90% (B).

doi:10.1371/journal.pntd.0003821.g004

Discussion

Preventive chemotherapy with praziquantel is the current backbone of the global strategy to control the morbidity caused by schistosomiasis in high-endemicity areas [2,4,6,8]. Monitoring of praziquantel efficacy should accompany schistosomiasis control programs in order to identify promptly suboptimal responders; to that effect, the WHO has issued standard procedures for control programs based on one single measurement around 3 weeks post-treatment [3]. At the same time, more work is needed to improve the current evidence-base for decision-making: in order to provide reliable information, it is important to agree on a robust statistical approach to assess drug efficacy, especially in clinical trials, and to use standardized, quality-controlled diagnostic methods that are comparable from one setting to another [18,42]. To date, both CR and ERR (the latter based on either AM or GM), are used for assessing drug efficacy. The stool and urine sampling and diagnostic approaches vary across studies, and these issues have important ramifications for drug efficacy estimates.
The controversy over the use of AMs or GMs to measure anthelmintic treatment efficacy as assessed by ERR in parasitic nematodes of cattle, and more recently in those of humans, has already been expounded in several studies, with contrasting results [43] favoring either AMs [17], or GMs (see e.g. [15]) or pointing to inadequacies of both [44]. The overall aim of this study was to compare and contrast different approaches to express treatment effects on *Schistosoma* infections in order to derive indications for standardizing future studies of drug efficacy.

The first specific objective of this paper was to compare customary measures of drug efficacy. We first considered whether efficacy assessment based on ERR changes when using AM or GM egg counts for the three predominant human *Schistosoma* species. While both means were in the same range for all species and showed a moderate level of correlation, the discrepancy between AM and GM became wider with decreasing drug efficacy. As previously reported
by some [17,43,45] but not all authors [17], GM estimates tended to be higher than AM. These findings suggest that the two means can be used interchangeably if drug efficacy is high (ERR > 95%), but the difference between the two means is expected to increase as efficacy decreases. It is also worth noting that two out of the 13 studies with assessment of efficacy within 3 weeks would not meet the WHO threshold for acceptable efficacy of 90% [3] for ERR when calculated by AM, while only 8% would not meet the WHO threshold, if calculated by GM (2/24). Therefore, between the two, using AM, as suggested by WHO, appears to be more sensitive an approach to identify problems with the response to praziquantel.

We employed models to help further qualifying these findings using explanatory variables. We found that ERRs are more consistent between the AM and GM egg count values of *S. haematobium* and *S. japonicum* than of *S. mansoni* infection, and for school-aged children and adults than for preschool-aged children (but only when using GM). On the contrary, these findings are not accounted for by the baseline and post-treatment distributions of the raw egg counts (S2 Table), intensities of infection (S3 Table), or proportion of individuals with extreme values (S2 Table). Together, these findings are important to allow a meaningful comparison of newer studies using AM to older studies which would have used GM.

We then compared drug efficacy estimates by ERR (using AM or GM) versus CR. While generally used in the past, the CR is known to have some major limitations, and is no longer recommended by WHO for assessing the programmatic efficacy of drugs used in mass drug administration [12,43,45]. As expected, efficacy estimates by ERR and CR were hardly comparable, as they assess two different outcome measures (intensity versus presence of infection). At the onset of schistosomiasis control, the primary goal is to reduce morbidity, which is reflected by infection intensity, and hence, the ERR rather than CR might be the efficacy measure of choice. This choice is further justified by the relatively low sensitivity of widely used diagnostic methods, particularly the Kato-Katz technique for intestinal schistosomiasis [22,24,42] and the
The fact that current anthelmintic drugs only show low to moderate efficacies in terms of CR [13,46,47]. However, the situation is complex, and neither ERR nor CR alone provide a satisfactory description of the situation. The key questions about outcome measures are between effects on presence versus intensity of infection; and between measures of central tendency (for a group of individuals) versus individual subject responses. Our analyses indicate that the distribution of individual responses in egg excretion may be a better way of expressing results, as it comprises in one single measure drug effects on both presence and intensity of infection, and allows further detailing the distribution in centiles—which helps identifying and quantitating the presence of poor responders. ERRAM and iERR are correlated, but the latter appears to be more apt to detect the presence and to quantitate the magnitude of suboptimal responses to praziquantel. More than 10% of individuals had ERR < 90% in five studies, only two of which were identified by applying the ERRAM 90% threshold. In these two studies, with ERRAM ~82%, <30% of individuals did not achieve a 90% reduction in their egg counts, and 11–17% had no reduction at all.

These observations raise important questions as to which approach is best suited to assess drug efficacy for which purpose: while ERR is, currently, the preferred measure at the program

### Table 3. Data outcomes at baseline in individual positives on single or quadruplicate Kato-Katz thick smears.

<table>
<thead>
<tr>
<th>Study</th>
<th>Species</th>
<th>Dose (mg/kg)</th>
<th>% subjects with ERR = 0%</th>
<th>% subjects with iERR &lt;90%</th>
<th>iERR 5th centile</th>
<th>iERR 10th centile</th>
<th>iRR 25th centile</th>
<th>ERRAM(1) Lower limit 95% CI(2)</th>
<th>Upper limit 95% CI(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Côte d'Ivoire, 2001</td>
<td><em>S.</em> haematobium</td>
<td>40</td>
<td>0.0%</td>
<td>0.0%</td>
<td>98.8%</td>
<td>98.8%</td>
<td>99.4%</td>
<td>99.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Niger, 2009</td>
<td><em>S.</em> haematobium</td>
<td>40</td>
<td>3.2%</td>
<td>13.7%</td>
<td>42.5%</td>
<td>84.6%</td>
<td>96.4%</td>
<td>95.7%</td>
<td>92.2%</td>
</tr>
<tr>
<td>Côte d'Ivoire, 2010</td>
<td><em>S.</em> haematobium</td>
<td>40</td>
<td>1.1%</td>
<td>4.6%</td>
<td>82.4%</td>
<td>95.0%</td>
<td>99.5%</td>
<td>97.8%</td>
<td>95.9%</td>
</tr>
<tr>
<td>Philippines, 2007</td>
<td><em>S.</em> japonicum</td>
<td>40</td>
<td>0.0%</td>
<td>2.0%</td>
<td>97.4%</td>
<td>98.8%</td>
<td>98.8%</td>
<td>99.6%</td>
<td>99.2%</td>
</tr>
<tr>
<td>Philippines, 2007</td>
<td><em>S.</em> japonicum</td>
<td>60</td>
<td>0.0%</td>
<td>0.0%</td>
<td>96.8%</td>
<td>96.8%</td>
<td>96.8%</td>
<td>100.0%</td>
<td>99.9%</td>
</tr>
<tr>
<td>Uganda, 2012</td>
<td><em>S.</em> mansoni</td>
<td>40</td>
<td>10.8%</td>
<td>30.2%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>81.1%</td>
<td>82.2%</td>
<td>72.7%</td>
</tr>
<tr>
<td>Niger, 2007</td>
<td><em>S.</em> mansoni</td>
<td>40</td>
<td>7.1%</td>
<td>36.6%</td>
<td>0.0%</td>
<td>18.2%</td>
<td>63.2%</td>
<td>81.5%</td>
<td>75.9%</td>
</tr>
<tr>
<td>Brazil, 2007</td>
<td><em>S.</em> mansoni</td>
<td>40</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Brazil, 2007</td>
<td><em>S.</em> mansoni</td>
<td>60</td>
<td>0.0%</td>
<td>0.0%</td>
<td>97.4%</td>
<td>97.4%</td>
<td>97.4%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Mauritania, 2007</td>
<td><em>S.</em> mansoni</td>
<td>40</td>
<td>0.0%</td>
<td>1.1%</td>
<td>96.0%</td>
<td>97.2%</td>
<td>97.2%</td>
<td>98.8%</td>
<td>96.8%</td>
</tr>
<tr>
<td>Mauritania, 2007</td>
<td><em>S.</em> mansoni</td>
<td>60</td>
<td>0.0%</td>
<td>3.2%</td>
<td>97.2%</td>
<td>99.4%</td>
<td>99.4%</td>
<td>94.0%</td>
<td>85.1%</td>
</tr>
<tr>
<td>Tanzania, 2007</td>
<td><em>S.</em> mansoni</td>
<td>40</td>
<td>1.7%</td>
<td>6.7%</td>
<td>73.2%</td>
<td>97.8%</td>
<td>99.7%</td>
<td>95.9%</td>
<td>91.7%</td>
</tr>
<tr>
<td>Tanzania, 2007</td>
<td><em>S.</em> mansoni</td>
<td>60</td>
<td>0.0%</td>
<td>2.4%</td>
<td>92.5%</td>
<td>98.1%</td>
<td>99.8%</td>
<td>98.7%</td>
<td>97.0%</td>
</tr>
<tr>
<td>Côte d'Ivoire, 1997*</td>
<td><em>S.</em> mansoni</td>
<td>40</td>
<td>3.5%</td>
<td>12.9%</td>
<td>40.5%</td>
<td>81.1%</td>
<td>97.9%</td>
<td>96.1%</td>
<td>93.2%</td>
</tr>
<tr>
<td>Côte d'Ivoire, 1998*</td>
<td><em>S.</em> mansoni</td>
<td>60</td>
<td>2.4%</td>
<td>12.9%</td>
<td>43.8%</td>
<td>82.8%</td>
<td>98.3%</td>
<td>93.6%</td>
<td>88.7%</td>
</tr>
</tbody>
</table>

* follow-up 28 days
1Egg reduction rate (ERR).
295% confidence interval (CI) calculated using a bootstrap resampling method.

doi:10.1371/journal.pntd.0003821.t003
level, identifying poor responders is important in view of tracking trends in responses and signaling potential problem areas and emergence of drug resistance. Importantly, discussions should be held with a variety of stakeholders, including drug regulatory authorities, especially if new drug applications are forthcoming.

Lastly, we compared drug efficacy estimates (ERR calculated using AM and GM, and CR) obtained by a single versus quadruplicate Kato-Katz thick smears for *S. mansoni*. Obviously, a

Table 4. Drug efficacy outcomes at baseline in individuals detected positive for *S. mansoni* by a single or quadruplicate Kato-Katz thick smear examinations at baseline.

<table>
<thead>
<tr>
<th>Sample screened</th>
<th>Kato-Katz examination</th>
<th>n</th>
<th>¹AMD0</th>
<th>²GMD0</th>
<th>³ERR AM (± 95% CI)</th>
<th>ERR GM (± 95% CI)</th>
<th>⁵CR (± 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall sample positive on 4 KK examinations</td>
<td>4 KK</td>
<td>1,435</td>
<td>878.7</td>
<td>223.7</td>
<td>86.9 (83.6–90.0)</td>
<td>99.4 (99.3–99.5)</td>
<td>79.6 (77.4–81.7)</td>
</tr>
<tr>
<td>Subgroup positive on 4 and 1 KK examinations</td>
<td>4 KK</td>
<td>1,167</td>
<td>1046.3</td>
<td>342.1</td>
<td>86.9 (83.3–90.6)</td>
<td>99.6 (99.4–99.6)</td>
<td>79.0 (76.7–81.3)</td>
</tr>
<tr>
<td></td>
<td>1 KK</td>
<td>1,167</td>
<td>2617.5</td>
<td>574.8</td>
<td>94.92 (93.18–96.39)</td>
<td>99.83 (99.78–99.86)</td>
<td>88.1 (86.2–89.9)</td>
</tr>
</tbody>
</table>

1Arithmetic mean (AM) at baseline
2Geometric mean (GM) at baseline
3Egg reduction rate (ERR)
495% confidence interval (CI) calculated using a bootstrap resampling method
5Cure rate (CR).
6Kato-Katz.

doi:10.1371/journal.pntd.0003821.t004

Fig 7. Comparison of cure rate (CR) calculated from a single versus quadruplicate Kato-Katz thick smears, both with estimated 95% confidence intervals (CI). A. L’Abbé plot. For each point, the 95% CIs are shown. The red dashed diagonal line of no difference between the ERR calculated with both metrics. B. Bland-Altman plot. The red dashed lines show the interval that determines the agreement of the two metrics, set at 5%.

doi:10.1371/journal.pntd.0003821.g007
single Kato-Katz thick smear has a lower sensitivity than multiple Kato-Katz thick smears (here ~81%) [48], and the effect and ramifications thereof for estimating drug efficacy against *Schistosoma* and other helminth infections have been discussed [42,49–51]. At the same time, collecting single stool samples and examining single Kato-Katz thick smears is operationally more feasible and less expensive than multiple sampling and multiple thick smear examinations under a microscope. We found that, when single Kato-Katz thick smears are used, the initial intensities of infection are ~2.5 times higher than with quadruplicate thick smears; efficacy estimation by ERRs based on AM and CR are ~8–9% higher, respectively, whereas ERRs based on GM are similar. Taken together, these results reflect the lower sensitivity of a single Kato-Katz thick smear, which misses low-intensity infections both on enrolment and post-treatment; the bias appears to be proportionally greater for the initial infection intensity than for the treatment outcomes. Overestimation of treatment effects may be an issue with efficacies nearing the 90% ERR threshold. In order to ascertain whether this introduces a selection bias which could affect the estimation of efficacy [52] (i.e., if the sample positive on a single Kato-Katz thick smear is different from that positive on multiple Kato-Katz thick smears), we compared both baseline FECs and efficacy estimates in these two samples using the same diagnostic technique; we found that excluding the ~19% of subjects who were negative on a single Kato-Katz selects for subjects with marginally higher initial FECs, but has ultimately no effect on efficacy estimates. Currently there is no diagnostic ‘gold’ standard method to assess *Schistosoma* response to treatment [53]. Standards may need to be tailored to the study, whether a field survey (limited by practical imperatives) or a clinical trial (which could afford more complex conditions and costly diagnosis), and whether in high or low intensity infection setting. Part of the problem, however, is the limited sensitivity of the current diagnostic methods, particularly for the detection of low intensity infections, which cannot be corrected until and unless more reliable tests become available [6,42,54].

Conclusions

Using group means is practical when assessing sample effects, but may not be suited to detect small changes, especially those that may occur in early phases of decreasing drug efficacy. We estimate that the distribution of individual responses in *Schistosoma* egg excretion, which accounts for individual variability of responses to praziquantel treatment, allows measuring effects on both presence and intensity of infection, and helps identifying and quantitating poor responders. More research and larger databases will be required in order to identify meaningful thresholds—e.g., centile by which 90% ERR is achieved; ERR achieved by the lowest 5% or 10% centile—and also analyze in greater detail reasons for poor response.

Both approaches could be used in parallel and complement each other. It is important to agree on standardized outcome measures that are tailored to specific purposes, such as epidemiological surveys, routine monitoring, clinical trials, morbidity control, or elimination settings. Hence, we invite other groups to contribute to this discussion and scientific inquiry.

Supporting Information

S1 Table. Results of general linear models of distinct analyses of ERR based on geometric means and ERR based on arithmetic means.

(XLSX)

S2 Table. Central values of the distribution of counts of eggs per gram and log transformed counts of eggs per gram measured at baseline for all studies.

(XLSX)
S3 Table. Frequencies and proportions of species by intensities of infection at baseline. (XLSX)

Acknowledgments
Thanks to Dr. Jing Xu from the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, for kindly providing *S. japonicum* data. We also wish to thank Florilène Bouisset for statistical advice.

Disclaimer. The opinions expressed in this paper are those of the authors and may not reflect those of their employing organizations. PLO is a staff member of the World Health Organization (WHO) and MV is a staff member of LIH; the authors alone are responsible for the views expressed in this publication and it does not necessarily represent the decisions, policy, or views of their organizations.

Author Contributions
Conceived and designed the experiments: PLO MV AD. Performed the experiments: PLO MV AD. Analyzed the data: PLO MV AD. Contributed reagents/materials/analysis tools: MV. Wrote the paper: PLO MV AD JTC AG JK CHK SK AL EKN GR AUS JCSF KS XNZ JU.

References


