Sensitivity and Specificity of Multiple Kato-Katz Thick Smears and a Circulating Cathodic Antigen Test for *Schistosoma mansoni* Diagnosis Pre- and Post-repeated-Praziquantel Treatment

Poppy H. L. Lamberton¹*, Narcis B. Kabatereine², David W. Oguttu², Alan Fenwick¹, Joanne P. Webster¹

¹Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, ²Schistosomiasis Control Initiative at Vector Control Division – Ministry of Health, Kampala, Uganda

**Abstract**

**Background:** Two Kato-Katz thick smears (Kato-Katzs) from a single stool are currently recommended for diagnosing *Schistosoma mansoni* infections to map areas for intervention. This ‘gold standard’ has low sensitivity at low infection intensities. The urine point-of-care circulating cathodic antigen test (POC-CCA) is potentially more sensitive but how accurately they detect *S. mansoni* after repeated praziquantel treatments, their suitability for measuring drug efficacy and their correlation with egg counts remain to be fully understood. We compared the accuracies of one to six Kato-Katzs and one POC-CCA for the diagnosis of *S. mansoni* in primary-school children who have received zero to ten praziquantel treatments. We determined the impact each diagnostic approach may have on monitoring and evaluation (M&E) and drug-efficacy findings.

**Method/Principle Findings:** In a high *S. mansoni* endemic area of Uganda, three days of consecutive stool samples were collected from primary school-aged children (six – 12 years) at five time-points in year one: baseline, one-week-post-, four-weeks-post-, six-months-post- and six-months-one-week-post-praziquantel and three time-points in years two and three: pre-, one-week-post- and four-weeks-post-praziquantel-treatment/retreatment (n = 1065). Two Kato-Katzs were performed on each stool. In parallel, one urine sample was collected and a single POC-CCA evaluated per child at each time-point in year one (n = 367). At baseline, diagnosis by two Kato-Katzs (sensitivity = 98.6%) or one POC-CCA (sensitivity = 91.7%, specificity = 75.0%) accurately predicted *S. mansoni* infections. However, one year later, a minimum of three Kato-Katzs, and two years later, five Kato-Katzs were required for accurate diagnosis (sensitivity >90%) and drug-efficacy evaluation. The POC-CCA was as sensitive as six Kato-Katzs four-weeks-post and six-months-post-treatment, if trace readings were classified as positive.

**Conclusions/Significance:** Six Kato-Katzs (two/stool from three stools) and/or one POC-CCA are required for M&E or drug-efficacy studies. Although unable to measure egg reduction rates, one POC-CCA appears to be more sensitive than six Kato-Katzs at four-weeks-post-praziquantel (drug efficacy) and six-months-post-praziquantel (M&E).

**Citation:** Lamberton PHL, Kabatereine NB, Oguttu DW, Fenwick A, Webster JP (2014) Sensitivity and Specificity of Multiple Kato-Katz Thick Smears and a Circulating Cathodic Antigen Test for *Schistosoma mansoni* Diagnosis Pre- and Post-repeated-Praziquantel Treatment. PLoS Negl Trop Dis 8(9): e3139. doi:10.1371/journal.pntd.00031319

**Editor:** Stefanie Knopp, Swiss Tropical and Public Health Institute, Switzerland

**Received** January 27, 2014; **Accepted** July 22, 2014; **Published** September 11, 2014

**Copyright:** © 2014 Lamberton et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Field work, analyses and/or authors were funded by the Bill and Melinda Gates Foundation though the Schistosomiasis Control Initiative (PHLL, NBK, DWO, AT, JPW), the Medical Research Council (PHLL-PNBD studentship) and the Royal Society (JPW-University Research Fellowship). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* Email: poppy.lamberton@imperial.ac.uk

**Introduction**

Schistosomiasis remains a major public health concern despite praziquantel reaching over 30 million people in endemic areas in 2013 [1]. Goals to eliminate schistosomiasis by 2020 have been articulated by the World Health Organization’s (WHO) ‘Roadmap for Neglected Tropical Disease (NTD) Implementation’ [2], and the London Declaration of the NTD Coalition [3]. Accurate diagnostic techniques, recently highlighted by Gomes and colleagues [4], are essential for monitoring and evaluation (M&E) of mass drug administration (MDA) programs at all stages [5–9], and particularly when considering elimination [2,10,11] and/or drug-resistance pharmacovigilance [12,13].

The WHO recommends two Kato-Katz thick smears (Kato-Katzs) from a single stool [14] for *Schistosoma mansoni* diagnosis to determine prevalence to map areas for control interventions [15]. Kato-Katzs have assumed 100% specificity, but large inter- and intra-specimen variation [16–18] and low sensitivity for the detection of low intensity infections have been reported [19–21]. In Brazil, where M&E programs use only one Kato-Katz, *S. mansoni* prevalence has been significantly underestimated in low intensity regions [22,23]. This may be associated with overesti-
mated cure rates (CRs) and praziquantel efficacy [24], potentially missing drug resistance. Conversely, using one Kato-Katz can overestimate infection intensities [22]. These differences in sensitivity for prevalence versus infection intensity highlight complex interactions with egg reduction rates (ERRs) as true intensities decrease. Inter-day variation of excreted egg numbers post-treatment remains poorly understood. Two Kato-Katz are commonly used for annual M&E of program impact and/or drug-efficacy studies, without being rigorously tested in these scenarios. Detailed analyses of the sensitivity and specificity of single and multiple Kato-Katz following praziquantel treatment is urgently required.

It is not possible to directly measure S. mansoni adult worm numbers, due to their location in the mesenteric system, with eggs counted from Kato-Katz used as a proxy for infection intensity. Immunodiagnostics for adult worm circulating cathodic antigens (CCA) or circulating anodic antigens (CAA) detect current infections and are potentially more sensitive for diagnosis of cases in low transmission areas [25–27]. Recent developments using innovative immunomagnetic separation with several target CCAs [28] and novel monoclonal antibody diagnostics for serum CCA [27] show high sensitivity even in low endemic areas. Serum CAA tests are more accurate than urine CAA tests [29], with advances made on a, not yet commercially available, point-of-care-CCA (POC-CAA) [30]. In contrast, urine CCA tests were more accurate than serum CCA tests [29]. Being easier to collect and more socially acceptable than stool or blood, urine POC-CCAs were developed for rapid, non-invasive diagnostics and proposed as alternatives to Kato-Katz for S. mansoni prevalence mapping [31–37]. One urine POC-CCA is as sensitive as two Kato-Katz for S. mansoni diagnosis for prevalence mapping [35] or two years post-treatment [36] and has now been used to map S. mansoni prevalence in low intensity areas in Uganda [37]. Major POC-CCA limitations are their ability to only semi-quantitatively measure infection intensity [32,34,38], inaccuracy for S. haematobium infection detection [39,40], and not measuring soil-transmitted helminth (STH) infections. Intensity of infection measures are vital for control program M&E and drug-efficacy evaluation. Therefore knowledge on the ability of POC-CCAs to detect S. mansoni intensity reductions is essential [34–36,41].

We compared the accuracy of the currently available POC-CCA and one to six Kato-Katzs (two smears per day from three consecutive stool samples) for S. mansoni diagnosis, in primary-school children in Mayuge District, Uganda, at baseline and after up to ten praziquantel treatments per child over three years (STARD Checklist S1 and Figure 1). We evaluated the epidemiological implications of diagnostic methods for control program M&E and praziquantel-efficacy studies. We inform on the number of Kato-Katzs required to accurately detect S. mansoni infections pre- and post-praziquantel treatment and whether POC-CCAs are suitable alternatives. Our detailed longitudinal design enabled novel investigations into individuals’ recent versus total praziquantel treatments, aiding biological understanding of differences between Kato-Katz and POC-CCA results post-praziquantel treatment. We predicted that the accuracy of M&E and drug-efficacy findings are limited by Kato-Katz sensitivity at low infection intensities post-treatment. We also predicted that a single POC-CCA would have a higher sensitivity than multiple Kato-Katzs, and be more informative for prevalence monitoring as control programs progress.

Materials and Methods

Ethics Statement

Approvals were granted by the Uganda National Council of Science and Technology (Memorandum of Understanding: sections 1.4, 1.5, 1.6) and the Imperial College Research Ethics Committee (EC NO: 03.36. R&D No: 03/SB/033E). Verbal assent was given by every child before inclusion into this study and at school committee meetings comprising of parents, teachers, and community leaders before the onset of the study. Written consent for the children to participate in the study was attained from each head teacher. Participation was voluntary and children could withdraw or be withdrawn from the study at any time without obligation. Children were treated with 40 mg/kg praziquantel and 400 mg albendazole (active against STH infections) as detailed below.

Study Area

Samples were collected, between 2004 to 2006, from primary-school children, in a high S. mansoni-endemic area, in Mayuge district, Uganda from three schools on the shores of Lake Victoria: Bugoto Lake View, Bwondha, and Musubi Church of God. Children at Musubi were, to the authors’ knowledge, praziquantel-naive. Children at Bugoto and Bwondha had received 40 mg/kg praziquantel one year previously in 2003 [42]. Inclusion criteria were to have lived in the area since birth and to attend the schools sampled.

Study Cohort and Treatment

In 2004, samples were collected at five time-points: baseline, one-week-post-, four-weeks-post-, six-months-post- and six-months-one-week-post-praziquantel treatment (Figure 1). In 2005 and 2006, samples were collected pre-, one-week-post-, and four-weeks-post-praziquantel re/treatment. On the third day of sampling, at baseline, six-months, one-year, and two-years all children were treated with 40 mg/kg praziquantel and 400 mg albendazole (active against STH infections). At one-week post-treatment, children with infections of >100 S. mansoni eggs per gram of stool (EPG) were retreated with 40 mg/kg praziquantel.
At all other time-points all children with positive diagnoses for *S. mansoni* or STHs were retreated with 40 mg/kg praziquantel and 400 mg albendazole respectively.

Cohort and sample collection are described elsewhere [43]. In brief, 110 children from Bugoto, 110 from Bwondha and 68 from Musubi were recruited in 2004 with an equal sex ratio, aged six to 12 years, without prior knowledge of infection status and/or symptoms of *S. mansoni* infection. In addition, at one- and two-years, 30 praziquantel-naïve six year old children were recruited at each school and followed up with the original cohorts at the time points described above. This enabled monitoring of the impact of MDA on untreated children entering the school system, assessing diagnostic accuracies for Kato-Katzs and POC-CCA, in praziquantel-naïve and praziquantel-exposed children, as control programs progress.

### Intensity of Infection and Prevalence Measures

Diagnostic accuracy increases with the number of Kato-Katzs, however, in Brazilian low intensity regions, the additional benefit of more than six Kato-Katzs from repeated stools was negligible [21], supporting our six Kato-Katz ‘gold standard’. Stool samples, marked with unique child IDs, were collected on three consecutive days, between 10:00 and 12:00 hours. Two 41.7 mg Kato-Katzs were prepared per stool and read onsite using a compound microscope with natural light source, by highly trained personnel from the Ugandan Vector Control Division, Ministry of Health. *S. mansoni*, hookworm, *Ascaris lumbricoides*, and *Trichuris trichiura* egg counts were recorded. Five percent of slides were reread after the study for *S. mansoni*, *A. lumbricoides*, and *T. trichiura* egg counts for quality control, but no significant differences were observed. One urine sample per child was collected between 10:00 and 12:00 hours on the first day. In the first year, at all five time-points, POC-CCAs (European Veterinary Laboratory, The Netherlands) were performed, according to the producer’s protocols, by the first author, blind of other test results. Microhematuria was tested for using Hemastix (Bayer, United Kingdom).

### Statistical Analysis

SPSS version 19 (SPSS, Inc., Chicago, IL, United States of America) was used for all statistical analyses. The double entered data were not normally distributed and could not be normalized.

---

**Figure 1. Recruitment and inclusion of primary-school children and their samples in the study.** Final numbers of children who provided three stools for a total of six Kato-Katz thick smears (6KK) and one urine sample for a point-of-care circulating antigen test (POC-CCA) at each time-point. Final numbers and percentage female in parentheses. Totaling 1065 samples for Kato-Katz thick smears accuracy analyses and 367 for POC-CCA analyses over the whole study. Samples were collected pre-praziquantel treatment, one-week-post-praziquantel treatment (1 Wk-Post-Treat) and four-weeks-post-praziquantel treatment (4 Wks-Post-Treat) at four time points: Baseline, six months later, one year later and two years later. doi:10.1371/journal.pntd.0003139.g001
Table 1. Baseline prevalence and intensities.

<table>
<thead>
<tr>
<th>Parasite</th>
<th>Diagnostic approach</th>
<th>No. of children included</th>
<th>Prevalence % (95% CI)</th>
<th>Mean infection intensity (SE)</th>
<th>Infection intensity counts (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Light</td>
</tr>
<tr>
<td><em>S. mansoni</em></td>
<td>Six Kato-Katzs</td>
<td>96</td>
<td>94.8 (88.3–98.3)</td>
<td>249.8 (30.2)</td>
<td>36 (39.6)</td>
</tr>
<tr>
<td></td>
<td>Six Kato-Katzs &amp; one POC-CCA</td>
<td>76</td>
<td>94.7 (87.1–98.5)</td>
<td>259.0 (35.8)</td>
<td>28 (38.9)</td>
</tr>
<tr>
<td>Hookworm</td>
<td>Six Kato-Katzs</td>
<td>96</td>
<td>51.0 (40.6–61.4)</td>
<td>104.3 (18.4)</td>
<td>49 (100)</td>
</tr>
<tr>
<td><em>T. trichiura</em></td>
<td>Six Kato-Katzs</td>
<td>96</td>
<td>9.4 (4.4–17.1)</td>
<td>6.1 (2.4)</td>
<td>9 (100)</td>
</tr>
</tbody>
</table>

Overall baseline prevalence (measured by six Kato-Katz thick smears (Kato-Katzs) and/or one rapid urine based point-of-care circulating cathodic antigen test (POC-CCA) if trace readings are counted as positive (POC-CCA+t) or negative (POC-CCA-t−) and intensities (arithmetic mean eggs per gram of stool) of helminth infections in three primary schools in Mayuge District, Uganda. CI = confidence interval, SE = standard error, na = not applicable. Infection intensity categories are as per World Health Organization guidelines [15].

doi:10.1371/journal.pntd.0003139.t001
by transformation, therefore non-parametric tests were used. Individuals without the full six Kato-Katzs were excluded from the study (Figure 1). Arithmetic mean infection intensities were categorized as by the WHO (S. mansoni: light = 1–99 EPG, moderate = 100–399 EPG and high ≥400 EPG; A. lumbricoides: light = 1–4999 EPG; Hookworm: light = 1–1,999 EPG; T. trichiura: light = 1–999 EPG) [15]. Exact confidence intervals (CIs) were calculated for prevalence measures and standard errors for EPGs.

Inclusion bias and potential confounders. There were no significant differences between the final dataset (six Kato-Katzs) and the excluded dataset (<six Kato-Katzs) in S. mansoni intensities (Mann Whitney: U = 509672.5, d.f. = 1420, p = 0.12), STH presence (χ² = 0.90, d.f. = 2, p = 0.34) or microhematuria (χ² = 1.51, d.f. = 4, p = 0.83). It was therefore assumed that the final dataset was not biased by missing data. There were no significant differences in accuracies of POC-CCAs, with trace counted as negative (POC-CCA-t) (Fisher’s exact: P_A = 0.91, P_B = 0.91) or positive (POC-CCA+t) (P_A = 0.93, P_B = 0.87) between those infected with STHs (n = 107) or not (n = 260). A greater proportion of the microhematuria positive samples (19/367, of which 17 were female) were POC-CCA negative than expected (χ² = 17.62, d.f. = 2, p<0.001), although microhematuria was not thought to have biased results, with 12 negative, two trace, four positive (+) and one double positive (++). All urine samples were screened for S. haematobium but no eggs were observed, confidently excluding S. haematobium co-infections.

Diagnostic accuracy of tests pre- and post-praziquantel-treatment. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of POC-CCA+t, POC-CCA+t−, and one to five Kato-Katzs were calculated with 95% exact CIs. Differences in sensitivity between our six Kato-Katzs ‘gold standard’ and the current recommended two Kato-Katzs or one POC-CCA were determined using the McNemar test. The agreement between six Kato-Katzs and two Kato-Katzs or POC-CCA were assessed using Kappa (κ) statistics: κ=0 no agreement; κ=0.01–0.20 poor; κ=0.21–0.4 fair; κ=0.41–0.6 moderate; κ=0.61–0.8 substantial; κ=0.81–1 almost perfect [44]. Accuracies were also calculated for POC-CCA and one to five Kato-Katzs, using the combined ‘gold standard’ of POC-CCA+t and six Kato-Katzs [41]. Finally, the accuracies of one to five

![Figure 2. Sensitivity and negative predictive values of one to five Kato-Katzs for S. mansoni diagnosis.](A) Sensitivity and (B) Negative Predictive Values of one to five Kato-Katzs thick smears (1KK to 5KK) for S. mansoni diagnosis over 11 time-points using six Kato-Katzs as the ‘gold standard’. All individuals were treated with 40 mg/kg praziquantel after the start of the study (pre) and then again at six-months, one-year and two-years. All individuals providing stool samples with >100 eggs per gram (EPG) at one-week-post-praziquantel treatment and any infected children at all other time-points were re-treated with 40 mg/kg praziquantel. 95% confidence intervals are excluded for clarity, but can be seen in Table 2. doi:10.1371/journal.pntd.0003139.g002
Table 2. The accuracy of one to five Kato-Katzs for detecting *S. mansoni* infection.

<table>
<thead>
<tr>
<th>Kato-Katz</th>
<th>Baseline</th>
<th>One Week</th>
<th>Four Weeks</th>
<th>Six Months</th>
<th>Six Months One Week</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sens</td>
<td>Spec</td>
<td>NPV</td>
<td>Sens</td>
<td>Spec</td>
</tr>
<tr>
<td>One</td>
<td>83.50%</td>
<td>100%</td>
<td>25%</td>
<td>84.70%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(74.3–90.5)</td>
<td>(47.8–100)</td>
<td>(8.7–49.1)</td>
<td>(74.3–92.1)</td>
<td>(54.1–100)</td>
</tr>
<tr>
<td>One</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three</td>
<td>100%</td>
<td>100%</td>
<td></td>
<td>95.80%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(96.0–100)</td>
<td>(47.8–100)</td>
<td></td>
<td>(88.3–99.1)</td>
<td>(29.9–92.5)</td>
</tr>
<tr>
<td>Four</td>
<td>100%</td>
<td>100%</td>
<td></td>
<td>98.60%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(96.0–100)</td>
<td>(47.8–100)</td>
<td></td>
<td>(92.5–100)</td>
<td>(42.1–99.6)</td>
</tr>
<tr>
<td>Five</td>
<td>100%</td>
<td>100%</td>
<td></td>
<td>98.60%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(96.0–100)</td>
<td>(47.8–100)</td>
<td></td>
<td>(92.5–100)</td>
<td>(42.1–99.6)</td>
</tr>
<tr>
<td>One Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>59.40%</td>
<td>100%</td>
<td>57.40%</td>
<td>60.30%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(46.4–71.5)</td>
<td>(90.0–100)</td>
<td>(44.1–70.0)</td>
<td>(37.9–63.6)</td>
<td>(93.0–100)</td>
</tr>
<tr>
<td>Two</td>
<td>70.30%</td>
<td>64.80%</td>
<td>63.50%</td>
<td>68.80%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(57.6–81.1)</td>
<td>(50.6–77.3)</td>
<td>(50.4–75.2)</td>
<td>(57.1–79.2)</td>
<td>(36.1–80.9)</td>
</tr>
<tr>
<td>Three</td>
<td>79.70%</td>
<td>72.90%</td>
<td>87.30%</td>
<td>86.40%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(67.8–88.7)</td>
<td>(58.2–84.7)</td>
<td>(76.5–94.4)</td>
<td>(75.0–94.0)</td>
<td>(56.3–94.3)</td>
</tr>
<tr>
<td>Four</td>
<td>90.60%</td>
<td>85.40%</td>
<td>90.50%</td>
<td>89.50%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(80.7–96.5)</td>
<td>(70.8–94.4)</td>
<td>(80.4–96.4)</td>
<td>(78.5–96.0)</td>
<td>(56.3–94.3)</td>
</tr>
<tr>
<td>Five</td>
<td>96.90%</td>
<td>94.60%</td>
<td>96.80%</td>
<td>96.20%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(89.2–99.6)</td>
<td>(81.8–99.3)</td>
<td>(89.0–99.6)</td>
<td>(87.0–99.5)</td>
<td>(75.1–99.9)</td>
</tr>
<tr>
<td>Two Years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>60.30%</td>
<td>100%</td>
<td>60.30%</td>
<td>69.10%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(51.7–73.9)</td>
<td>(92.0–100)</td>
<td>(48.1–71.5)</td>
<td>(55.2–80.9)</td>
<td>(91.6–100)</td>
</tr>
<tr>
<td>Two</td>
<td>68.80%</td>
<td>68.80%</td>
<td>78.20%</td>
<td>77.80%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(63.6–83.8)</td>
<td>(55.4–79.8)</td>
<td>(65.0–88.2)</td>
<td>(64.4–88.0)</td>
<td>(23.4–83.3)</td>
</tr>
<tr>
<td>Three</td>
<td>83.00%</td>
<td>83.00%</td>
<td>92.70%</td>
<td>91.10%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Kato-Katzs and POC-CCA were calculated over time since each individual had first ever been treated, to explore differences between community and individual treatment history on diagnostic accuracies.

**Effect of mean infection intensity on accuracy.** Data from each school, at each time-point, were used to compare sensitivities of one to five Kato-Katzs against school arithmetic mean EPG, using Spearman’s rank correlation. Five best fit linear lines (for one to five Kato-Katzs) were calculated. The mean distance of data-points from these lines were compared, using Mann-Whitney, between two data-subsets to test whether the distance was greater for time-points one- or four-weeks-post-praziquantel in comparison to ‘pre-praziquantel’ (pre, six-months, one-year and two-years). No significant differences were found (all \( p > 0.05 \)) and data-points were combined for analysis of the effect of infection intensity.

**Effect of diagnostic test on treatment efficacy and reinfection measures.** Cure rates were determined, for each diagnostic method, as the proportion of *S. mansoni*-positive individuals at baseline who were negative four-weeks-post-praziquantel. Odds ratios were calculated for comparisons of prevalence at each time-point. Comparisons of infection intensities, over three days of Kato-Katzs, were performed using the Friedman test for repeated measures.

**Ability of Kato-Katzs and POC-CCA to measure infection intensity pre- and post-praziquantel.** Band strengths of the POC-CCAs (negative, trace, +, ++ and ++++) were compared with Kato-Katz infection intensity categories and individual children’s arithmetic mean EPG using Spearman’s rank coefficient.

**Results**

There were 1065 samples with six Kato-Katzs and 367 samples with six Kato-Katzs and a POC-CCA result (Figure 1). Baseline *S. mansoni* prevalence (\( \chi^2 = 0.38 \), d.f. = 2, \( p = 0.83 \)) and EPG intensity (Kruskal-Wallis: \( H = 3.416 \), d.f. = 2, \( p = 0.18 \)) were not significantly different between schools and all statistics were performed on the combined data. Baseline prevalence in the main (six Kato-Katzs) dataset was 94.8% with an arithmetic mean infection intensity of 249.8 EPG, similar to the POC-CCA dataset prevalence (94.7%) and intensity (259.0 EPG) (Table 1). Hookworm prevalence was 51.0%, whilst *A. lumbricoides* and *T. trichiura* infections were low at 1.0% and 9.4%, respectively (Table 1).

**Diagnostic Accuracies Pre- and Post-Praziquantel-Treatment**

**Kato-Katz.** There were no significant differences between the sensitivity and NPV of one to five Kato-Katzs between the POC-CCA and Kato-Katzs only datasets (all \( p > 0.05 \)) (Figure 2, Tables 2 and 3). At baseline, 97% of *S. mansoni* infections were detected by two Kato-Katzs, reaching 100% with a second day of sampling (Figure 2). At one-week-post-praziquantel, sensitivity of two Kato-Katzs was approximately 90%. However, by four-weeks-post-praziquantel, nearly half of infected individuals were wrongly classified as uninfected with the sensitivity of two Kato-Katzs being only 51.9%. Four Kato-Katzs had a higher sensitivity of 77.8%, but five Kato-Katzs were required for sensitivities above 90% at four-weeks-post-praziquantel. The sensitivity of two Kato-Katzs at six-month-post-praziquantel was 74.0%, but five Kato-Katzs were required to reach sensitivities above 90% as praziquantel
Table 3. The accuracy of one to six Kato-Katzs and one point-of-care circulating cathodic antigen test for diagnosing *S. mansoni* infections.

<table>
<thead>
<tr>
<th>Test Accuracy for <em>S. mansoni</em> Diagnosis Pre- and Post-treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gold Standard</strong></td>
</tr>
<tr>
<td><strong>6 Kato-Katzs</strong></td>
</tr>
<tr>
<td><strong>Baseline (n = 76)</strong></td>
</tr>
<tr>
<td><strong>6 Kato-Katzs &amp; POC-CCA-t</strong></td>
</tr>
<tr>
<td><strong>PPV</strong></td>
</tr>
<tr>
<td><strong>Baseline (n = 76)</strong></td>
</tr>
<tr>
<td><strong>Prev = 94.7% (87.2–97.9)</strong></td>
</tr>
</tbody>
</table>
### Table 3. Cont.

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>One Week</td>
<td>96.1% (88.0–99.2)</td>
<td>91.1% (80.4–99.2)</td>
<td>98.6% (92.6–99.8)</td>
<td>98.6% (92.6–99.8)</td>
</tr>
<tr>
<td>Four Weeks</td>
<td>61.0% (50.0–72.8)</td>
<td>49.0% (35.1–63.1)</td>
<td>66.0% (51.2–80.0)</td>
<td>66.0% (51.2–80.0)</td>
</tr>
<tr>
<td>Six Months</td>
<td>98.9% (89.2–100)</td>
<td>99.8% (98.5–100)</td>
<td>99.8% (98.5–100)</td>
<td>99.8% (98.5–100)</td>
</tr>
<tr>
<td>Six Months One Week</td>
<td>99.9% (98.2–100)</td>
<td>99.9% (98.2–100)</td>
<td>99.9% (98.2–100)</td>
<td>99.9% (98.2–100)</td>
</tr>
</tbody>
</table>

**Test Accuracy for S. mansoni Diagnosis Pre- and Post-treatment**

At baseline, only a quarter of the individuals found to be negative by one Kato-Katz were negative for six Kato-Katzes, with three Kato-Katzes required for a PPV of 100% (Figure 2A, Table 2). In each year of the study, NPVs peaked four weeks after treatment, but were at their lowest immediately before the next round of treatment (Figure 2B, Table 2).

The sensitivity of one to four Kato-Katzes was positively correlated with the mean infection intensity for one Kato-Katz ($r = 0.539$, $p = 0.001$); two Kato-Katzes ($r = 0.427$, $p = 0.02$); three Kato-Katzes ($r = 0.424$, $p = 0.02$); four Kato-Katzes ($r = 0.384$, $p = 0.03$), but not five Kato-Katzes ($r = 0.125$, $p = 0.50$) (Figure 3). Best fit lines indicated that in a school with an arithmetic mean infection intensity of, for example 300 EPG, two Kato-Katzes had 90% sensitivity, whereas in a school with an arithmetic mean infection intensity of 100 EPG, at least four Kato-Katzes were required for 90% sensitivity (Figure 3).

At baseline, the agreement between two Kato-Katzes and six Kato-Katzes was almost perfect ($k = 0.82$), however at other times the agreements were only substantial or moderate. The sensitivity of two Kato-Katzes was not significantly different to the ‘gold standard’ ($p = 0.15$) at baseline, but was significantly lower at all other times ($p < 0.001$ to 0.03).

**POC-CCA.** POC-CCA showed a high sensitivity at baseline of 91.7%, but a low sensitivity at one- and four-weeks-post-praziquantel of 66.7% and 75.0% respectively (Table 3). Sensitivity returned to baseline levels by six-months-post-praziquantel at 92.0% (Table 3). POC-CCA specificity varied with time since praziquantel (Table 3). Only one of the 19 children classified as negative by six Kato-Katzes had a negative POC-CCA result six-months-post-praziquantel. All POC-CCA+t and -t PPVs were ≥80% except at four-weeks-post-praziquantel. Conversely, the NPVs were continuously low except at four-weeks-post-praziquantel. POC-CCA+t showed a moderate agreement with the ‘gold standard’ at baseline ($k = 0.42$), but at all other time-points, only poor or no agreement. POC-CCA-t showed a fair agreement at baseline and six-months-one-week ($k = 0.34$), but a moderate agreement at four-weeks ($k = 0.44$). POC-CCA+t results were significantly different from six Kato-Katzes at one-week, four-weeks and six-months (all $p < 0.01$). POC-CCA-t results were significantly different from six Kato-Katzes at all time-points except four-weeks-post-praziquantel (all $p > 0.02$).

**Effect of Diagnostic Test on Treatment Efficacy and Reinfection**

Observed cure (measured at four-weeks-post-praziquantel treatment) and reinfection (measured at six-months) rates depended on sampling method and effort (Figures 4 and 5). Two Kato-Katzes underestimated S. mansoni reinfection whilst overestimating CRs (Figure 4A) (two Kato-Katzes CR = 81.5%; six Kato-Katzes CR = 70.4%). Cure rates determined with POC-CCA+t were 47.8% and 26.1% for POC-CCA+t. One-week-post-reinfection (from data at both one-week and six-months-one-week), prevalence was significantly lower when measured by POC-CCAs than by six Kato-Katzes (Figure 4B) (OR 0.33 (95% CI: 0.19, 0.59)).

Pre-re/treatment and at four-weeks-post-praziquantel-re/treatment in years zero, one, and two, the number of days of Kato-Katzes did not significantly affect the infection intensities (Figure 5) (all $p > 0.05$). However, at one-week-post-re/treatment (one-week, six-months-one-week, one-year-one-week and two-years-one-
Six Kato-Katzs and POC-CCA as the ‘Gold Standard’

Pre-treatment, one POC-CCA+t, and two Kato-Katzs had sensitivities above 90%, however the POC-CCA+t and t NPVs were extremely low, similar to one Kato-Katz (Table 3). At one-week-post-praziquantel, one POC-CCA was less sensitive and had lower NPVs than one Kato-Katz, whilst two Kato-Katzs had an NPV of only 55.6%. At four-weeks-post-praziquantel POC-CCA+t had high sensitivity and NPV (>80%), whilst POC-CCA+t and one to six Kato-Katzs all had sensitivities and NPVs of <60%. Indeed, two Kato-Katzs only detected a quarter of the infections. At six-months-post-praziquantel POC-CCA+t was more sensitive than six Kato-Katzs, however all diagnostics showed NPV values below 15%.

Effect of Individual Child’s Praziquantel Treatment on Test Accuracies

Some Bugoto and Bwondha children had been treated once before the study and new praziquantel-naïve cohorts were recruited each year. We therefore also analyzed our data along timelines specific for each individual child’s praziquantel exposure (Tables S2 and S3). Key results were not affected by re-analyzing the data in this manner. Pre-treatment, six-months, one-year and one-year-six-months POC-CCA+t+s showed high sensitivities but low NPVs throughout (Table S2). One-week-post-recent-praziquantel, three Kato-Katzs were required for >90% sensitivity in general, and four-weeks-post-recent-praziquantel four or five Kato-Katzs were required for >90% sensitivities (Table S3). An increased Kato-Katz sampling effort was required year on year to achieve sensitivities of >90% (Table S3), which was not clearly seen in the original M&E timeline (Figure 2, Table 2). Praziquantel-naïve children and children one-year-post-praziquantel required three Kato-Katzs for accurate S. mansoni diagnosis, whilst four Kato-Katzs were required at two-years, and five Kato-Katzs at three-years.

Discussion

We evaluated one to six Kato-Katzs and one POC-CCA for S. mansoni diagnosis before and after multiple rounds of praziquantel treatment, and how test choice affects M&E and drug-efficacy interpretations. Our data support using one POC-CCA+t or two Kato-Katzs for pre-treatment mapping in high endemicity areas [35–37]. However, as MDA continues, five Kato-Katzs were required for diagnosis of children after three to ten praziquantel treatments. Indeed, such high sensitivity and specificity interpretations were not observed again throughout this study.

POC-CCAs are shown to be more sensitive but less specific than two Kato-Katzs [33,37,45,46]. Our data show that one POC-CCA+t at four-weeks-post-praziquantel for praziquantel-efficacy studies and six-months-post-praziquantel for M&E, was more sensitive than six Kato-Katzs at the same time periods. Our POC-CCA+t baseline sensitivity (91.7%), from a 94.8% S. mansoni prevalence population, was comparable with that previously measured as eggs per gram of stool (EPG) from six Kato-Katz thick smears). S. mansoni infection intensities were measured at three primary schools, at 11 time points each, ranging from pre-treatment to two-years-four-weeks-post-praziquantel treatment. Lines are best fit linear lines.

doi:10.1371/journal.pntd.0003139.g003

Figure 3. Effect of S. mansoni arithmetic mean infection intensity on sensitivity of one to five Kato-Katzs. The sensitivity of one to five Kato-Katz thick smears (1KK to 5KK) for diagnosing S. mansoni infections at a range of community arithmetic mean S. mansoni infection intensities (measured as eggs per gram of stool (EPG) from six Kato-Katz thick smears). S. mansoni infection intensities were measured at three primary schools, at 11 time points each, ranging from pre-treatment to two-years-four-weeks-post-praziquantel treatment. Lines are best fit linear lines.
published from Côte d’Ivoire (sensitivity = 86.9%, prevalence = 91.8%) [33]. In contrast, our 73% sensitivity at four-weeks-post-praziquantel (prevalence = 34.2%) was greater than in the low prevalence Côte d’Ivoire region (sensitivity = 56.3%, prevalence = 32.9%) [33]. This may be explained by that study’s rigorous nine Kato-Katzs ‘gold standard’, with our six Kato-Katzs possibly still missing infections. In addition, in Côte d’Ivoire, three POC-CCAs were performed, increasing sensitivity, in comparison with our single POC-CCA [33]. The lack of POC-CCA reproducibility data, even from single urine samples [35], are a key limitation of our study. Though utilizing matching components, our accuracies from European Veterinary Laboratory POC-CCAs, may vary from the Rapid Medical Diagnostics’ POC-CCAs used in Côte d’Ivoire, however differences were not observed at higher prevalence.

At four-weeks-post-praziquantel, prevalence levels as indicated by six Kato-Katzs and one POC-CCA was nearly double (61.8%) than for just six Kato-Katzs (34.2%). Cure rates using two Kato-Katzs were >80% versus 70% with six Kato-Katzs, and only ~25% with POC-CCA-t+. Similar results have been seen for S. haematobium [47]. Further discordance between Kato-Katzs and POC-CCA at six-months (specificity of 5.3%) may be explained by high numbers of infections missed by Kato-Katzs. It is unlikely that POC-CCA false positives are the full explanation due to high specificity at four-weeks-post-praziquantel treatment, with potentially more ‘true’ negatives and only 1% of POC-CCA giving false positives in non-endemic areas [35]. We believe that the low POC-CCA specificities are, in part, due to low sensitivities of Kato-Katzs.

When six Kato-Katzs and one POC-CCA were the combined ‘gold standard’, baseline and one-week accuracies were relatively unaffected. Four-weeks-post-praziquantel Kato-Katzs sensitivities were substantially lower, having profound implications on what is a suitable ‘gold standard’ when communities have received multiple praziquantel treatments. 

---

**Figure 4. Effect of diagnostic technique and sampling effort on S. mansoni prevalence measures.** The effect of (A) number of Kato-Katz thick smears (Kato-Katzs), from one to six (1KK to 6KK) on the recorded prevalence of S. mansoni infection with multiple praziquantel treatments. All individuals were treated after sampling at pre, six-months, one-year and two-years. All individuals excreting more than 100 EPG at one-week and any infected children at all other time-points were re-treated. 95% confidence intervals are excluded for clarity. Difference in sensitivities between the current standard two Kato-Katzs and our ‘gold standard’ of six Kato-Katzs determined by the McNemar test is significant at all time-points except baseline. The effect of (B) diagnosis method using six Kato-Katzs or one point-of-care circulating cathodic antigen test (POC-CCA), with a trace counting as either a positive (POC-CCA-t+) or negative (POC-CCA-t−) on S. mansoni infection prevalence with 95% confidence intervals. Difference in sensitivities between the POC-CCA and our ‘gold standard’ of six Kato-Katzs determined by the McNemar test: * significant for POC-CCA-t− or † significant POC-CCA-t+. doi:10.1371/journal.pntd.0003139.g004
but is not fully applicable in praziquantel-efficacy studies, providing weighted prevalence rather than individual infection and clearance data.

Studies from the same region in Uganda demonstrating reduced *S. mansoni* infection prevalence and intensity levels in response to MDA [42,48], used only two Kato-Katzs and may have overestimated annual reductions [37]. We strongly recommend, as treatment campaigns continue, increased sampling efforts and/or alternative tools to accurately record program success and CRs, to detect early drug-resistance indicators, as for STHs [49]. The need for a higher number of Kato-Katzs for accurate diagnosis as the number of previous praziquantel treatments increases is not unexpected considering the small amount of stool used in each Kato-Katz and the progressively lower egg counts. Our baseline data indicated that two Kato-Katzs had a sensitivity of 90%, whereas in contrast, praziquantel-naive individuals (Table S3) required four Kato-Katzs for accurate predictions. This apparent conflicting result may be explained by the high number of praziquantel-naive recruits each year, sampled after several school-based MDA rounds, lowering infection intensities through reduced force-of-infections [50], supported by Figure 3, where Kato-Katz sensitivities decrease with EPG.

*S. mansoni* infection intensities were not expected to vary with Kato-Katz sampling effort. However, one-week-post annual or biannual treatment, mean intensities decreased from day one to day three, likely due to continued daily reductions in egg excretion post treatment. This, and our discordant POC-CCA and Kato-Katz results post treatment, raise interesting questions regarding parasite antigen and egg clearance, such as residual-egg clearance (with individuals with intensities of >100 EPG at one-week, not retreated, but negative at four-weeks), praziquantel-induced fecundity compensation and/or increased egg expulsion (with higher EPGs at one-week-day-one, than at baseline, as also observed in *S. haematobium* [51]).

In contrast, at four-weeks-post-praziquantel, positive POC-CCA results in egg negative individuals may have occurred due to juveniles unaffected by treatment, newly acquired infections, and/or worms which survived treatment, but with reduced or ceased egg production (embryostasis). Drug-induced embryostasis, has been demonstrated in *Oxyuris volvulus* [52] and *Ascaris suum* [53]. Embryostasis could explain our lower sensitivities (73%) at four-weeks-post-praziquantel than those observed in a stable, low transmission Western Kenyan region (prevalence 38.8%, sensitivity = 96%) [34]. In this Kenyan region a large proportion of individuals may be truly negative with no egg or antigen excretions. Embryostasis could significantly affect drug-resistance selection, with worms repeatedly exposed to praziquantel, without dying or being detected by standard parasitological techniques. Being impossible to sample adult worms directly, studies on worm antigens, egg production and molecular studies incorporating sibship analyses informing adult breeding numbers [54,55] post-praziquantel treatment may elucidate this.

As intensities decrease, costs of accurate diagnoses by Kato-Katzs will rise due to greater sampling requirements. Diagnoses using urine, rather than stool, remain quicker, cheaper on labor costs, more convenient, socially acceptable and may improve compliance [35]. In low endemicity areas, pooled urine samples for POC-CCAs could reduce costs further. However, key POC-CCA limitations are their inability to detect STHs, and inaccuracy measuring infection intensities and treatment resolutions. Multiple smears from one stool (versus multiple stools) and FLOTAC [49] may be viable diagnostic alternatives. If one sampling day can accurately detect schistosomiasis and STH infection intensities and ERRs, it may be highly cost-effective, warranting further research. As the geographical distribution of STH infections are more homogeneous than schistosomiasis, WHO recommends surveys of smaller subsets of schools for mapping and M&E [15]. We therefore recommend widespread POC-CCA use, with Kato-Katzs performed in a subsection of schools. For drug-efficacy studies we recommend at least six Kato-Katzs or one POC-CCA, with further research on clearance dynamics of eggs and antigens post treatment needed.

Conclusions

At least four Kato-Katzs (two smears per stool from two stools) are required for M&E, in the early years of a MDA program in a highly endemic area, increasing to six Kato-Katzs (two smears per stool from three stools) by year three. One POC-CCA is a suitable alternative to current prevalence M&E protocols, but they provide no

![Figure 5. *S. mansoni* intensity measures from one to three days of duplicate Kato-Katz thick smears. Data shown over multiple rounds of praziquantel treatments with standard error bars. doi:10.1371/journal.pntd.0003139.g005](image-url)
information on STHs and limited intensity data post treatment, therefore we recommend their use for *S. mansoni* M&E with Kato-Katzs performed in a subset of schools. For drug-efficacy studies, at least six Kato-Katzs (two smears per stool from three stools) are required for accurate prevalence assessment four-weeks-post-praziquantel treatment. POC-CCAs may be a promising alternative with low specificity findings potentially due to low Kato-Katzs sensitivity, however further work is required to elucidate POC-CCA’s full potential for drug-efficacy studies. Further work on improved ‘gold standards’ is required to elucidate discordant POC-CCA and Kato-Katzs results. Data on multiple Kato-Katzs from a single stool post-treatment would ascertain if accuracies of multiple days of Kato-Katzs or POC-CCAs could be matched, minimizing logistical costs without overestimating M&E success and drug efficacy, whilst retaining vitally important intensity data.

**Supporting Information**

**Checklist S1**  **STARD checklist.**

(DOC)

**Table S1**  *S. mansoni* infection intensity categories by six Kato-Katzs and a single POC-CCA. Proportion of World Health Organization infection intensity categories, as measured by six Kato-Katz thick smears, which are correctly identified by a
single-point-of-care circulating cathodic antigen test (POC-CCA) band strength, ranging from negative to three. Tests were performed pre-treatment (Baseline), one-week-post- (1 WK), four-weeks-post- (4 WKs), six-months-post- (6 Mths) and six-months-one-week-post (6 Mths 1 WK) praziquantel treatment. Percentage of tests correctly identified in parentheses. (DOCX)

Table S2: Accuracy of one to five Kato-Katz's for diagnosing S. mansoni infections over time since each child was first treated with praziquantel. The accuracy of one to five Kato-Katz thick smears (1KK to 5KK) and a single point-of-care circulating cathodic antigen test (POC-CCA) (comparing if trace readings are counted as positive (POC-CCA+τ) or negative (POC-CCA−τ)) for detecting S. mansoni infection over the time since each child was first treated with praziquantel (10 time-points, in real time from the start of this study) with six Kato-Katz thick smears (6KK) as the ‘gold standard’. Sens = sensitivity, Spec = specificity, NPV = negative predictive value, PPV = positive predictive value. (DOCX)

Acknowledgments
We would like to thank Prof. J. Russell Stothard, Dr. Emily Robinson, Vector Control Division technicians and drivers, and the school teachers and children for assistance and samples in the field and Dr. Charlotte Gower for detailed comments on the manuscript.

Author Contributions
Conceived and designed the experiments: JPW PHLL NBK AF. Performed the experiments: PHLL DWO. Analyzed the data: PHLL Contributed reagents/materials/analysis tools: NBK JPW AF. Wrote the paper: PHLL JPW NBK DWO AF.

References
30. Van Dam GJ. Schistosomiasis diagnosis by circulating antigen detection: from lab-based ultimate UCP-lateral flow diagnostic to field-available point-of-care rapid test; 2012; Rio de Janeiro, Brazil.