Title

Reproducibility matters: Intra- and inter-sample variation of the point-of-care circulating cathodic antigen test (POC-CCA) in two *Schistosoma mansoni* endemic areas in Uganda

Authors and affiliations

Elías Kabbas-Piñango[1], Moses Arinaitwe[2], Govert J van Dam[3], Adriko Moses[2], Annet Namukuta[2], Andrina Barungi Nankasi [2], Nicholas Khayinja Mwima[2], Fred Besigye [2], Joaquin M Prada*[4], Poppy Helen Louise Lamberton*[1]

[1] School of Biodiversity, One Health & Veterinary Medicine, Wellcome Centre for Integrative Parasitology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
[2] Vector Borne and NTD Control Division, Bilharzia and Worm Control program Uganda, Ministry of Health, Kampala, Uganda
[3] Leiden University Medical Center, Department of Parasitology, Leiden, Netherlands
[4] Department of Comparative Biomedical Sciences, Faculty of Health & Medical Sciences, University of Surrey, United Kingdom
* Denotes joint senior author

ORCID IDs: EKP: 0000-0002-6290-9493; GvD: 0000-0003-3773-0819; AM: 0000-0001-9748-1207; JMP: 0000-0003-4699-5931; PHLL: 0000-0003-1048-6318.

Abstract

Over 240 million people are infected with schistosomiasis. Detecting *Schistosoma mansoni* eggs in stool using Kato-Katz thick smears (Kato-Katzs) is highly specific but lacks sensitivity. The urine-based point-of-care circulating cathodic antigen test (POC-CCA) has higher sensitivity, but issues include specificity, discrepancy between batches and interpretation of trace results. A semi-quantitative G-score and latent class analyses making no assumptions about trace readings, have helped address some of these issues. However, intra-sample and inter-sample variation remains unknown for POC-CCAs.

We collected three days of stool and urine from 349 and 621 participants, from high- and moderate-endemicity areas, respectively. We performed duplicate Kato-Katzs and one POC-CCA per sample. In the high endemicity community, we also performed three POC-CCA technical replicates on one urine sample/participant. Latent class analysis was performed to estimate the relative contribution of intra- (test technical reproducibility) and inter-sample (day-to-day) variation on sensitivity and specificity. Within sample variation for Kato-Katzs was higher than between samples, with the opposite for POC-CCAs. A POC-CCA G3 threshold most accurately assesses individual infections. However, to reach the WHO Target product profile of the required 95% specificity for prevalence and monitoring and evaluation, a threshold of G4 is needed, but at the cost of reducing sensitivity.

Key Words

Schistosoma mansoni, diagnostics, latent class analysis, POC-CCA, intra-sample, inter-sample

Introduction

Schistosomiasis is a debilitating parasitic neglected tropical disease (NTD), caused by trematodes of the *Schistosoma* genus. There are six main species infecting humans: *S. mansoni*, *S. japonicum*, *S. guineensis*, *S. mekongi* and *S. intercalatum*, causing hepatointestinal schistosomiasis; and *S. haematobium* and its hybrids, causing urogenital schistosomiasis. People become infected by direct contact with freshwater contaminated with cercariae released from infected intermediate snail hosts. The cercariae burrow into the skin, before migrating, pairing up, and sexually reproducing in capillaries surrounding the intestines or bladder, depending on species. Worm pairs produce up to 300 eggs per day [1], a proportion of which are excreted in the faeces (intestinal species) or urine (*S. haematobium)*; eggs can also get retained in the liver or the bladder, respectively, causing inflammatory immune responses and the formation of granulomas [1]. Suboptimal sanitation enables excreted eggs to reach fresh water, hatch and infect new snails, completing the life cycle.

Approximately 240 million people are infected worldwide, 90% of them in sub-Saharan Africa. However, the total number is likely underestimated due to the lack of sensitive diagnostics [2]. Because adult worms lie within the capillaries, they are not accessible for direct diagnosis. All current diagnostic techniques provide indirect estimates of adult worm numbers: eggs excreted in stool or urine, antigens regurgitated by adult feeding worms, DNA from different life-cycle stages, or host antibodies against the parasite. In 2021, the Global Schistosomiasis Alliance Diagnostic Workstream published a list of all commercially available diagnostics for schistosomiasis [3], but within endemic settings diagnoses focus on the WHO endorsed egg microscopy and/or antigen detection.

In *S. mansoni* endemic settings, the mainstay diagnostic is microscopy, detecting eggs in stool using Kato-Katz thick smears (Kato-Katzs). It is highly specific and detects active infections, but lacks sensitivity in low intensity infections and regions [4–6] and particularly post treatment [7–9]. Sensitivity can be improved by increasing the number of smears per stool and/or stools sampled, but this increases logistical, temporal and financial costs. Artificial intelligence algorithms for automated or semiautomated identification of eggs, enable faster reading, but the sensitivity does not yet outperform humans [10]. Mathematical models have informed easy-to-use tools to estimate true *S. mansoni* prevalence from observed Kato-Katz prevalence to improve interpretation of sub-optimal sensitivity [11]. However, this can only improve population level prevalence indicators, and not individual diagnoses.

The point-of-care-circulating cathodic antigen test (POC-CCA) (Rapid Medical Diagnostics, Pretoria, RSA, currently distributed by ICT International on behalf of Rapid Medical Diagnostics) is an urine-based lateral flow assay that requires no equipment, is easy to read by the naked eye, enables high throughput and less processing than microscopy, uses the more popular sample of urine rather than stool, and is recommended for the detection of *S. mansoni* infections [12–14], and endorsed by the World Health Organization (WHO) since 2017 [15]. The POC-CCA is more sensitive than Kato-Katz, especially for low intensity infections [16–19] but issues exist with batch variation [20], low specificity including in samples from non-endemic areas [21–23] and cross-reactivity with other helminths [22], inter reader variability especially with trace results, and interpretation of these [24,25], all of which can affect individual diagnoses and prevalence estimates [26]. Improved standardisation and quality control is required for the POC-CCA to be more reliable [27], which becomes more important as regions and countries move towards the WHO goal of elimination as a public health problem [28]. Although a WHO endorsed, and commonly used, *S. mansoni* diagnostic, no

guidelines exist for POC-CCA based cut offs. Clark et al [29] have translated egg to antigen-based indicators for WHO targets, but these may be affected by intra- and inter-sample variation, which are currently unknown, and which may also vary with infection intensity and endemicity levels [30,31].

Traditionally, POC-CCA results were reported as negative (-), trace, and a range of positive intensities (+, ++, +++), based on the readers' interpretation of visual band strength. The G-score system improves on this using ten pre-printed test lines of increasing intensities, ranging from G1 to G10, enabling a wider range of semiquantitative results and lower inter-reader variation [32]. Recent latent class analyses have further elucidated 'trace' results [29,33], with the probability of G3 being a true positive being much higher than G2, resulting in a recommendation that G3 and above be classified as positive [29], improving population level predictions, as well as individual level diagnoses. In the absence of a perfect schistosomiasis diagnostic test (gold standard), a composite reference standard (CRS) can be built using different imperfect test results. However, this can lead to overestimation of prevalence if any tests are not 100% specific [34,35], such as POC-CCA [21,22] [23]. Alternatives to CRS include latent class analysis statistical methods, which have improved over time [29,36–39] and now no longer make assumptions about trace readings [29,33]. Latent class analyses have informed our understanding of true prevalence [36–39], clearance and reinfection [33] and WHO elimination targets [29]. If all costs are considered (test supply, transport, labour, and others), triplicate Kato-Katz is more expensive than a single POC-CCA, but a single POC-CCA is more expensive than a single Kato-Katz [40], and three days of POC-CCAs will be more expensive than three days of Kato-Katzs. Due to the higher sensitivity of POC-CCA and its slightly higher cost compared to one Kato-Katz [40], only one POC-CCA tends to be performed per person. Whilst data exist on the inter and intra-sample variation of Kato-Katz [41,42] little is known about intra-sample and inter-sample variation of antigen levels, or test reproducibility, nor how these affect the sensitivity and specificity of the results in comparison to the newly published WHO diagnostic target product profile (TPP). To our knowledge, only one study has investigated the use of POC-CCAs on repeated urine samples [43], however they did not report on variations between days, only on the final correlation of a composite reference standard, with Kato-Katz and another CCA-based test. Due to the slightly higher cost per test of POC-CCAs in comparison to Kato-Katz [40], information on the minimum number of POC-CCAs required for each WHO TPP end use is needed. Furthermore, use of different G-score cut offs will affect both the sensitivity and specificity of the POC-CCAs, and may therefore affect the minimum number of tests required for a specific case use.

[34,35][21,22][23][29,36–39][29,33][36–39][33][29][41,42][43]As with any diagnostic, it is important to minimise costs without reducing sensitivity or specificity below required thresholds [40], and information on the minimum number of POC-CCAs required to achieve the minimum, or desired, criteria for each WHO TPP end use is needed. Furthermore, use of different G-score cut offs will affect both the sensitivity and specificity of the POC-CCAs, and may therefore affect the minimum number of tests required for a specific case use.

The aim of this study was to quantify the effect of intra- and inter-sample POC-CCA variation in *S. mansoni* high and low endemicity communities to ascertain the accuracy of a single POC-CCA in correctly detecting a person's infection status and a community's endemicity level. Specifically, we address four key objectives: (i) to quantify the intra-sample variation of POC-CCA using three tests on the same urine (test reproducibility) and duplicate Kato-Katzs; (ii) to quantify the inter-sample variation of POC-CCAs and Kato-Katzs, using three samples from the

same person, over three different days; (iii) and to determine the minimum number of POC-CCAs needed to accurately report prevalence in higher and lower endemicity settings in comparison to the WHO TPP (sensitivity and specificity of single and multiple POC-CCAs) and (iv) how G-score thresholds affect each of these.

Methods

Cohort recruitment

In December 2021, 660 people were recruited in the villages of Kalachai A, Kateki, Kogala and Oburi, in Tororo, an inland district in the Eastern Region of Uganda, classified as low endemicity for *S. mansoni*. In May 2022, 386 people were recruited in Bugoto in Mayuge District, also in the Eastern Region of Uganda, but located on the shore of Lake Victoria and classified as high endemicity for *S. mansoni*. The two cohorts had an similar distribution of males and females (46.2%-53.8% male-female in Tororo, 43.1%-56.9% male-female in Mayuge, respectively), and all ages were considered for the study, see Supplementary Figure S1. Out of the recruited participants, 621 participants in Tororo and 349 participants in Mayuge provided at least one sample, and thus could be included in the analysis. Individuals with incomplete records (i.e. those who submitted at least one sample, but not all the samples) were included in the analysis as the Bayesian statistical framework used (see below) allows to infer missing data.

Sample and data collection

For both Tororo and Mayuge cohorts, all participants were asked to provide one stool and one urine sample on each of three days. Each stool sample was analysed using Kato-Katzs [44] prepared using 41.7 mg templates, and stained with malachite green. Duplicate Kato-Katzs were performed per stool (two smears from the same portion of stool after sieving), resulting in egg counts from six Kato-Katzs per person. For both endemicity settings, inter-sample CCA variation in urine collected over different days was assessed. Each urine sample was analysed using one POC-CCA (Schisto POC-CCA®, ICT International, Cape Town, RSA) following the manufacturer's instructions. In brief, 100 µL of urine were put into the POC-CCA sample well using an automatic pipette, and the test left on a flat surface for 20 minutes. Semiquantitative results (G1 to G10) were assigned by a trained reader following the G-score system [32].

Additionally, in Mayuge, three technical replicates (three POC-CCAs run on the same urine sample) were performed, as described above, to assess the intra-sample variation on one urine per person. Regardless of the number of technical replicates performed, all urines were homogenised prior to taking the aliquot used to run the POC-CCA. For the inter-sample variation, only the first POC-CCA from the technical replicates on the same sample in Mayuge was considered and used to report the observed prevalence.

Throughout the study two POC-CCA batches were used (210811080 and 211110105). For quality control purposes, and for each of the two batches used, one POC-CCA was run for each of the four reference standards (S-Series) containing 0, 80, 800 and 8000 ng/mL of the trichloroacetic acid-soluble fraction of *S. mansoni* adult worm antigen (AWA-TCA), containing approximately 3% CCA [32] with results shown in the Supplementary Table S1.

Statistical analyses

Data were double entered using Microsoft Excel® (Microsoft 365 MSO, version 2209) and checked for discrepancies and analysed using R (version 4.2.2). A descriptive analysis was initially carried out, with the prevalence and mean infection intensities and standard errors

estimated using the raw Kato-Katz data. Prevalences were also estimated using the raw POC-CCA data, using a range of positivity thresholds of the average G-Score of G2, G2.5, G3 and G4 or above [29]. When multiple G-Scores were available for the same sample, the average (arithmetic mean) G-Score was calculated taking the G-Scores for their numeric value (e.g. G1 = 1, G2 = 2, etc.). 95% confidence intervals were calculated by bootstrapping, extracting the 2.5th and 97.5th quantiles of 1000 bootstrap repeats.

To quantify the intra- and inter-sample variation, we extended a Bayesian latent class analysis framework recently developed [29,33,38]. Briefly, a latent (hidden) variable captures the true infection status of an individual (status = 0 for uninfected individuals – or with undetectable levels of infection, status = 1 for infected individuals). The value of this binary variable (either 0 or 1) is inferred by the model, given the outcomes of the different diagnostics, which were observed. Due to the high specificity of Kato-Katz, and the imperfect sensitivity of POC-CCA tests, including Kato-Katz data in our models increased the accuracy of predicting individual infection status, therefore enabling improved measures of sensitivity and specificity of the individual POC-CCAs.

For Kato-Katzs, we assume specificity is 100%, meaning that an individual with status = 0, must have zero eggs in all of their six raw egg counts. For individuals that are infected (status = 1), we assumed a gamma distributed infection intensity at the population level. For each individual, we allow the predicted egg count (infection intensity) to vary between days following a normal distribution, while multiple Kato-Katzs processed on the same day are assumed to be over-dispersed from the mean number of "expected" eggs excreted that day.

For the POC-CCAs, we assumed that the true intensity of the underlying antigen band is related in a non-linear way to the infection intensity. We used the same framework from Clark *et al.* 2021 and 2022 [29,33] – a logistic function – due to its flexibility, and used the posterior draws for the parameters of this function. Intra-sample and inter-sample variation were modelled assuming gaussian (normal) noise.

All model details can be found in https://github.com/joaquinprada/Schisto-CCA-reproducibility. The model was run using "jags" [45] and "runjags" [46] packages in R version 4.2.2 [47], with two independent chains, a 'burn-in' period of 20,000 iterations, 10,000 samples and a thinning of 10. Posteriors from previous work [29,38] were used as priors for some parameters, as mentioned above, for the remaining parameters, uninformative priors were used. Convergence was assessed by visual examination of the trace plots and the Gelman-Rubin statistic. The model was run independently in both settings, using the posterior estimates from the model runs in Mayuge as priors in the runs for Tororo, to account for the fact that more data were collected in Mayuge, with intra-sample variation data also existing.

Using the model posteriors, we conducted a simulation exercise to estimate the sensitivity and specificity of the POC-CCA test when performing one, two or three samples over consecutive days and considering different thresholds (G-score of G2, G2.5, G3 and G4 respectively), which was used to generate the Receiver Operating Characteristic (ROC) curve. We also calculated the squared error in the estimation of prevalence across the number of days of sampling and thresholds. These simulations were carried out in both endemicity settings. Percentage agreements and discrepancies for *S. mansoni* positivity were also calculated for the POC-CCA raw inter-sample data.

Ethical Clearance

Ethical approvals were granted from the Vector Control Division Research Ethics Committee of the Ministry of Health of Uganda (VCDREC/062), the Uganda National Council of Science and Technology (UNCST-HS 2193) and the University of Glasgow Medical, Veterinary and Life Sciences Research Ethics Committee (200160068). Before any data or sample collection, informed consent was given, by signature or thumb print, by recruited adults and by the parent or legal guardian of all children <18 years old; and informed assent was given by all recruited children aged eight and older.

Results

Samples from 621 participants, aged 1 to 85, were collected in Tororo; and from 349 participants, aged 3 to 83 years old, in Mayuge. The final sample was well gender-balanced, with 53.8% females and 46.2% males recruited in Tororo, and 56.9% females and 43.1% males in Mayuge. Prevalence as estimated from three days of duplicate Kato-Katzs was 29.1% in Tororo and 56.6% in Mayuge. Arithmetic mean infection intensities and their corresponding standard errors were 40.4±6.0 eggs per gram of stool (epg) in Tororo and 145±19.5 epg in Mayuge. The Model reproduced the distribution of infection intensity obtained through Kato-Katz, see Supplementary Figure S2. The prevalence observed in Tororo meant that it was actually an area which would be classified as moderate endemicity by the WHO [48] rather than the low endemicity area we had aimed to survey, and therefore it is described as that from here on.

G-score results of each of the four reference standards (S-series) run on one POC-CCA test per batch, are shown in the Supplementary Table S1. Both batches showed a lower intensity of the test line than the expected range for the highest reference standard, whilst batch 211110105 also showed a lower G-score than the expected range for S1 and S2.

<u>Intra-sample variation for prevalence from the raw data in comparison to the model estimate</u>

In Mayuge, 279 participants had three POC-CCAs performed on a given urine sample (Table 1). From the raw data only, when using any positive POC-CCA as a positive test outcome, then increasing the number of tests per urine sample, results in increasing prevalence estimates with G2 and G3 thresholds both overestimating the prevalence if more than one test is performed. However, one POC-CCA accurately estimates the prevalence in comparison with our best estimate from the model of true prevalence (59.5% (54.4-64.4, 95% credible interval (CI)) if using G-score 3 as the cut off, with little gained from increasing the number of tests per sample if using an average G-score.

Additionally, for the intra-sample variation studied in Mayuge, the percent of positivity of each of the three POC-CCA tests performed in a single urine sample is shown in the Supplementary Table S2.

Table 1. Intra-sample variation of the point-of-care circulating cathodic antigen test (POC-CCA) data in Mayuge (high endemicity setting), expressed as the sample prevalences where three technical replicates were performed. The model estimates are made using up to five POC-CCAs and six Kato-Katz thick smears per participant (three POC-CCAs on one day and one per day on two separate days). BtCI = Bootstrap confidence interval; BCI = Bayesian credible interval.

POC-CCA threshold	Any positive POC-CCA, % (95% BtCI)			Average POC-CCA result (%)			Model estimate, %
	1 test	2 tests	3 tests	1 test	2 tests	3 tests	(95% BCI)
G2	79.1 (74.2-83.4)	82.3 (78.1-86.7)	83.4 (78.9-87.5)	79.1 (74.2-83.4)	77.9 (73.1-82.8)	75.3 (70.3-80.6)	
G2.5	-	-	-	-	64.0 (58.4-69.5)	62.5 (57.0-68.1)	
G3	59.0 (53.0-64.9)	64.8 (58.8-70.3)	66.3 (60.6-72.0)	59.0 (53.0-64.9)	59.6 (53.8-65.6)	59.9 (54.1-65.2)	59.5 (54.4-64.4)
G4	48.7 (43.0-54.8)	56.1 (50.2-61.7)	58.1 (52.0-63.8)	48.7 (43.0-54.8)	50.0 (43.7-55.9)	50.9 (45.2-56.6)	

Inter-sample variation for prevalence from the raw data in comparison to the model estimate

In Mayuge, the high endemicity setting, the observed prevalence based on all six Kato-Katzs was 56.6%, a minor underestimated prevalence in comparison to the model estimated true prevalence of 59.5% (Table 2). If any positive POC-CCA result was used from across three days, then G-score cut-offs from G2 to G3, resulted in an overestimated prevalence. When considering the POC-CCA average of all tests performed, prevalence was 74.3, 62.7, 58.4 and 46.8% for G2, G2.5, G3 and G4 thresholds, respectively, with the G-score cut off of 3, most closely correlating to the model estimate. If fewer days were used then the cut off of G3 was still the most similar to the model estimated true prevalence.

In Tororo (Table 2), a moderate endemicity setting, the lowest observed prevalence was given by one day of duplicate Kato-Katzs (19.5%), with three days of duplicates (29.1%) still underestimating the true model estimated prevalence of 36%. Using any positive POC-CCA over three days of single POC-CCAs, over-estimated the prevalence when using any of the G-score thresholds tested. Whereas setting the POC-CCA threshold as either G2 or G2.5 overestimated the prevalence (52.3% or 41.9% respectively) when averaging all three tests performed per participant, compared to the model estimated prevalence of 36.0% (32.1-40.1, 95% CI). G4 was not a sensitive enough threshold (29.6% using the average of all tests), whilst G3 (37.7%) gave the observed prevalence closest to the model estimated true prevalence, as also seen in Mayuge. If fewer days were used then the cut off of G3 (for 2 days) or G4 for one day was the most similar to the model estimated true prevalence.

Additionally, for the inter-sample variation, the agreement and discrepancy in positivity (i.e. the proportion of study participants producing the same or a different POC-CCA outcome across the three days, respectively) is shown in the Supplementary Table S3.

Table 2. Inter-sample variation of Kato-Katz thick smears (Kato-Katzs) and point-of-care circulating cathodic antigen tests (POC-CCAs). Observed *Schistosoma mansoni* prevalence based on positivity percent of Kato-Katzs and POC-CCAs. The model estimates are made using up to three POC-CCA tests per participant performed in Tororo

(moderate endemicity), and up to five per participant (three on one day and one per day on two separate days) in Mayuge (high endemicity). The raw data for inter day variation however only use up to three tests per person, with no intra-sample replicates, using the result from the first test performed each day, to enable direct comparison between the two endemicity areas. The column "any positive test" shows the observed prevalence, considering an individual as positive if any of the diagnostic tests performed on any of their samples was positive for the corresponding technique. The column "average test result" takes the arithmetic mean of the values and - for POC-CCAs - establishes different thresholds (t) to consider individuals as positive. Any value >0 for Kato-Katzs results in that participant being classified as positive for *S. mansoni*. BtCI = Bootstrap confidence interval; BCI = Bayesian credible interval.

	Observed prevalence, % (95% BtCI)						**Model estimate, % (95% BCI)**
	Any positive test			**Average test result**			
	1 day	2 days	3 days	1 day	2 days	3 days	
Mayuge							
Kato-Katz	39.4 (34.5-44.5)	52.6 (47.7-58.0)	56.6 (51.1-61.2)	39.4 (34.5-44.5)	52.6 (47.7-58.0)	56.6 (51.1-61.2)	59.5 (54.4-64.4)
POC-CCA (t=G2)	71.3 (66.7-76.1)	83.8 (79.9-87.7)	91.0 (87.9-93.7)	71.3 (66.7-76.1)	72.0 (67.2-76.4)	74.3 (69.5-78.5)	
POC-CCA (t=G2.5)	-	-	-	-	61.0 (55.7-66.1)	62.7 (57.5-67.2)	
POC-CCA (t=G3)	54.0 (49.1-58.9)	63.6 (58.3-68.1)	71.7 (67.0-76.1)	54.0 (49.1-58.9)	55.8 (50.6-61.5)	58.4 (53.2-63.5)	
POC-CCA (t=G4)	43.4 (38.2-48.6)	54.9 (49.7-60.1)	61.8 (56.9-66.7)	43.4 (38.2-48.6)	45.4 (40.5-50.9)	46.8 (41.4-52.3)	
Tororo							
Kato-Katz	19.5 (16.4-22.5)	26.0 (22.7-29.3)	29.1 (25.6-32.7)	19.5 (16.4-22.5)	26.0 (22.7-29.3)	29.1 (25.6-32.7)	36.0 (32.1-40.1)
POC-CCA (t=G2)	55.2 (51.0-59.1)	63.8 (60.1-67.5)	67.3 (63.6-71.0)	55.2 (51.0-59.1)	53.7 (49.8-57.8)	52.3 (48.1-56.4)	
POC-CCA (t=G2.5)	-	-	-	-	45.6 (41.5-49.8)	41.9 (38.0-46.2)	
POC-CCA (t=G3)	41.7 (37.8-45.7)	49.8 (46.2-53.5)	52.3 (48.1-56.4)	41.7 (37.8-45.7)	40.0 (36.1-43.6)	37.7 (34.3-41.1)	
POC-CCA (t=G4)	33.0 (29.5-36.7)	40.3 (36.4-44.1)	42.1 (38.5-46.2)	33.0 (29.5-36.7)	30.3 (26.7-33.8)	29.6 (25.9-33.3)	

Intra- and inter-sample infection intensity model variation estimates

In the high endemic setting, Mayuge, the variance of the estimated "true" intensity of infection across multiple days and therefore, samples (inter-sample variation), for individuals estimated to be infected, was around 87 epg (43-117, 95% CI), while it was 1433 epg (936-2212, 95% CI) within the same sample (intra-sample variation) (Figure 1 top-left). Conversely, for the G-score, the inter-sample variation was 4.68 (4.1-5.36, 95% CI), while the intra-sample variation was much smaller, 0.71 (0.63-0.80, 95% CI) (Figure 1 top-right). In the moderate endemic setting, Tororo, the variance in the estimated "true" intensity of infection remains relatively similar between samples (47 epg, 31-72 CI), but is again much higher within samples (intra-sample variance: 550 epg, 364-810 CI) (Figure 1 bottom-left). Regarding the POC-CCA, while the intra-sample variation in Tororo was fixed, as there were no multiple tests performed the same day, the inter-sample variation increased slightly (6.32, 5.7-7 CI) compared to Mayuge (Figure 1 bottom-right).

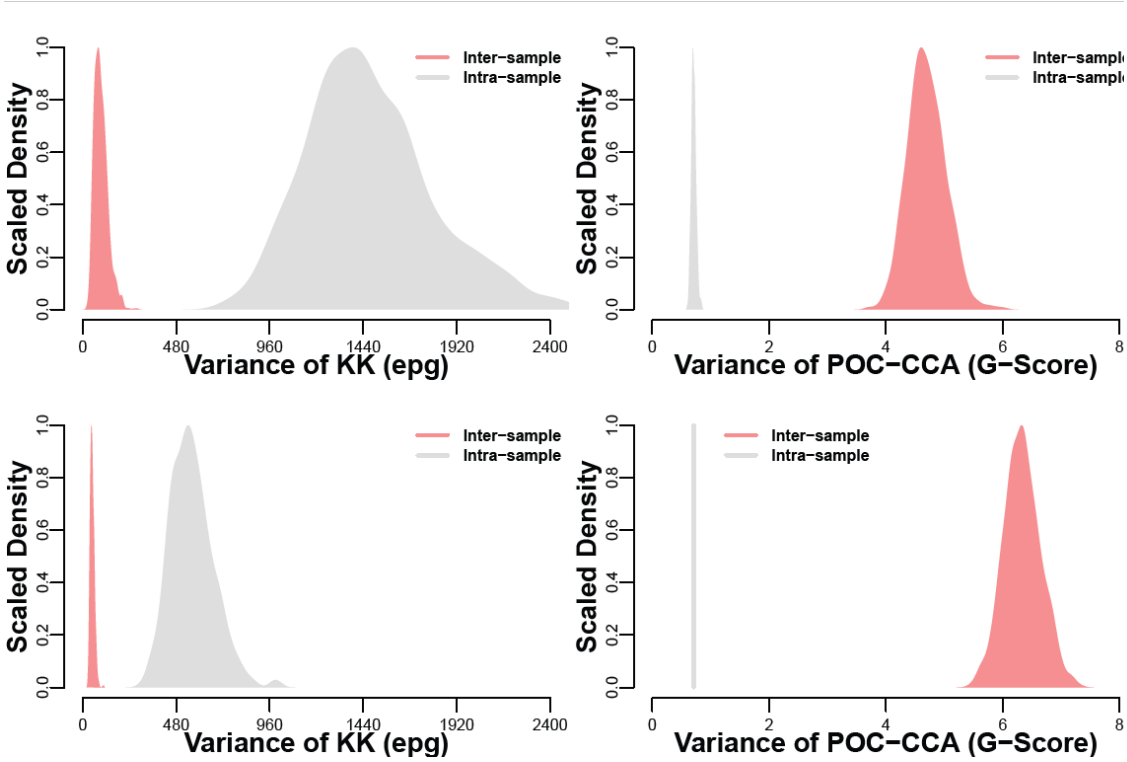

Figure 1. Inter-sample (red) and intra-sample (grey) variance for the Kato-Katz (left) and point-of-care circulating cathodic antigen test (POC-CCA) (right), for Mayuge (top) and Tororo (bottom). epg: eggs per gram in stool. G-Score: a semi-quantitative scale of antigen concentration from G1 to G10. Note: the intra-sample variation of POC-CCA in Tororo was fixed to the mean value from Mayuge, as no repeated tests were performed on the same day and sample in Tororo.

Number of POC-CCAs required to accurately report prevalence in low and high endemicity settings

The Receiver Operating Characteristic (ROC) curves for the POC-CCA indicate that higher sensitivity and specificity can be achieved when sampling over more than one day (Figure 2). Using a higher threshold significantly increases specificity at the cost of a marginal reduction in sensitivity up until G3, but sensitivity drops further in both settings when the threshold is higher at G4 (Figure 2). Using a G-score cut off of G3, and three days of sampling would result in near 100% sensitivity and slightly over 90% specificity in the higher endemicity area of Mayuge, and near 100% sensitivity and over 85% specificity in the moderate endemicity area of Tororo. Reducing the number of sampling days to only one, but still using the G3 cut off, lowers the sensitivity to 90% and specificity to 80% in Mayuge, but reduces it even further to approximately 85% and 75% respectively in Tororo (Figure 2).

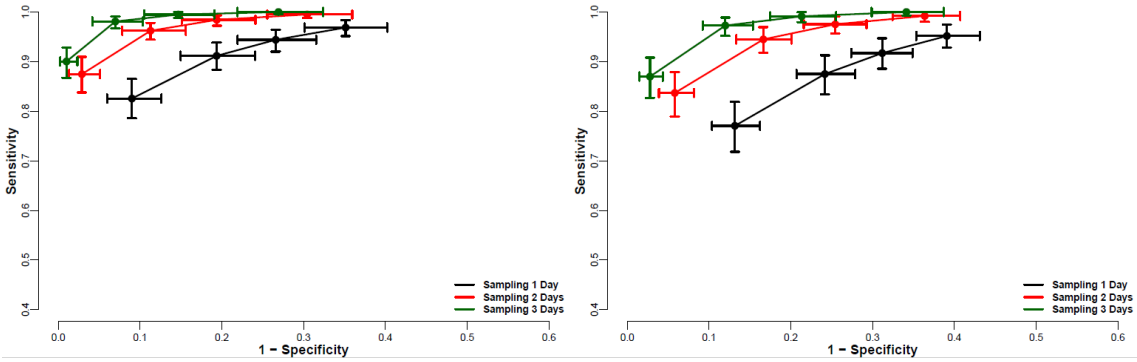

Figure 2. Receiver Operating Characteristic curve for the POC-CCA diagnostic depending on the number of sampling days (top – green – 3 days, middle – red – 2 days, and black – bottom – 1 day). The G-score threshold used moving from left (G4) to right (G2), for each colour (G4, G3, G2.5 and G2 respectively). Mayuge presented on the left, Tororo on the right.

The increased variance between days of the POC-CCA diagnostic in a moderate endemicity setting compared to a high endemicity setting has an impact in the error when estimating prevalence using this diagnostic alone. While increasing the threshold in the G-score reduces this error (except for G4 in the high endemic setting), increasing the number of days of sampling could also be key in a moderate endemicity settings (Figure 3).

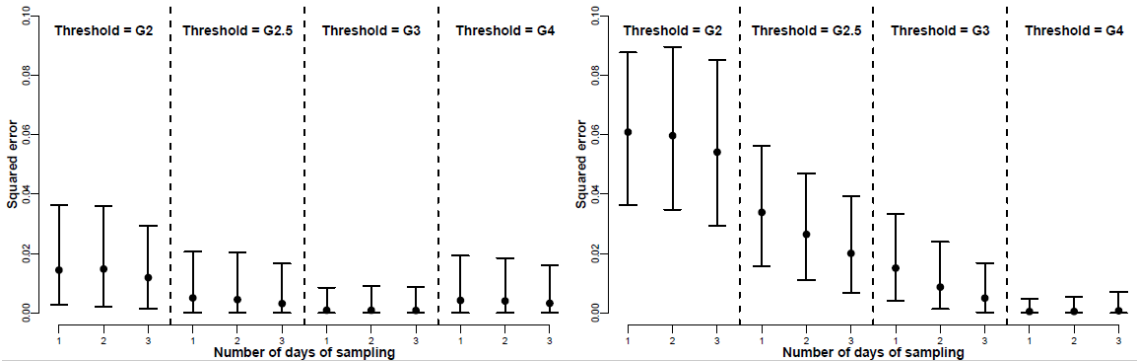

Figure 3. Squared error in the estimation of *S. mansoni* prevalence using point-of-care circulating cathodic antigen tests (POC-CCAs) depending on the number of days of sampling and the threshold used for the G-

score, for Mayuge – left and Tororo – right. The error is calculated by squaring the difference between the simulated prevalence and the estimated prevalence obtained from the diagnostic.

Discussion

Here, for the first time to our knowledge, we investigated the effect of intra- and inter-sample variation on POC-CCA results, identifying where the greatest degree of variation in results from POC-CCA comes from, to enable an informed recommendation on the number of POC-CCAs required and the G-score cut off to be used, to accurately predict community level infection as well as individual level diagnoses for WHO TPP requirements. Using an updated latent class analysis model we estimated the true infection status for each individual, as well as community prevalence based on different thresholds for the POC-CCA diagnostic. We compared raw field results against this to help provide recommendations for use directly in the field.

In brief we show that using G3 as a cut off, supporting previous work [33], one single POC-CCA test per person accurately reflects the population level prevalence in both a high and lower endemicity area, with two or more POC-CCAs not improving on this prevalence estimate. However, at an individual level this relates to only a 90% sensitivity and 80% specificity in the high endemicity area, and 85% sensitivity and 75% specificity in the moderate endemicity area. In contrast, performing POC-CCAs on urine samples collected over three multiple days, can raise sensitivity to near 100% and specificity to approximately 90% for both high and moderate endemicity areas. POC-CCAs on samples over multiple days can also improve infection intensity estimates and using the G3 as a cut off minimises error rates. In summary, G3 is the best cut off compared to the model estimates. However, our data indicate that this is not specific enough to meet the WHO TPP requirements for monitoring and evaluation.

In areas of ongoing *S. mansoni* monitoring and evaluation, the WHO TPP states that for a sample of 100 people surveyed, a minimum of 60% sensitivity and 95% specificity is required [49], therefore we would recommend two days of urine sampling with one POC-CCA per day, but with a cut off of G4, which sacrifices sensitivity for specificity in comparison to the G3 cut off (Figure 2). If only one day of sampling is to be performed, then the POC-CCA does not meet the required combined sensitivities and specificities for *S. mansoni* surveillance, interruption of transmission, nor monitoring and evaluation (Figure 2). Given that the lowest prevalence measured in this study is over 10% , we cannot provide recommendations on the use of the POC-CCA in interruption of transmission scenarios. However, our data, especially the drop in sensitivity when using G4 as the threshold, even when increasing the number of sampling days, suggests that the POC-CCA would highly likely also not be fit for use in interruption of transmission scenarios, as any of the two tests (initial or confirmatory) that are recommended by the WHO TPP [49].

Point-of-care diagnostics such as the POC-CCA pose an economic investment in the short term, but their improved sensitivity in comparison with microscopy, their ease of use and their acceptability may continue to make them key players in the fight against NTDs. However, further work, building on our findings into the cost-effectiveness of individual diagnoses and how this relates to sensitivity and specificity and downstream decisions requires further studies.

49Egg excretion appears to be fairly stable between days, with most of the variation in the observed counts due to within sample, rather than inter-sample variation (Figure 1). Therefore, taking multiple smears from the same stool will give a better estimate of true

infection intensity than processing single samples from repeated stools. Previous studies reported a higher variation in the presence or absence of *S. mansoni* eggs between samples from different days than within the same specimen [42], potentially due to more noise between samples where there is a difference in the true egg numbers within a given stool plus the variation in detecting those that are there. However, in line with our results, the variation in egg counts in infected people was found to be higher intra-sample than between days [42]. The lower intra-sample variation in Kato-Katz egg counts in the moderate endemicity area of Tororo, compared to the highly endemic area in Mayuge, is likely explained by the true lower intensity of infection. Increasing the number of stool smears and days when samples are taken for microscopy increases sensitivity. However, this approach, as previously demonstrated, does not compensate fully for the lack of sensitivity of Kato-Katzs, therefore using POC-CCAs can greatly improve on this, especially in lower endemicity areas. In comparison however, there appears to be greater inter-sample than intra-sample variation for the POC-CCAs. Given the more uniform nature of urine and the simplicity of homogenising samples, this is unsurprising, but it means that to improve reliability of POC-CCA interpretation, it is better to run tests on multiple days with multiple urine samples, with the added logistical and financial costs associated with this. One possible way to reduce costs could be to collect urine across multiple days, but to pool it by person prior to testing it with POC-CCAs, although this would not mitigate the logistical costs of multiple days of sampling, and could increase errors which may inflate prevalence measures if urine from truly uninfected people was cross-pooled with urine from infected people.

We have also shown that, for both high and moderate endemicity settings, the error of prevalence estimation decreases significantly when increasing the number of samples, and especially when using the G3 threshold as recommended, guided by our results here and previous work [33]. Setting G4 as the threshold would reduce the error in a low moderate endemic setting, with a corresponding increase in specificity, which might be worth considering when it is important to reduce false-positives (as for interruption of transmission scenarios as discussed above). In a high endemic setting though, setting the threshold at G4 would lead to a higher error in prevalence estimation than establishing it at G3. This contradicts however, the required sensitivity and specificity for the different WHO TPPs, and might indicate that a different diagnostic entirely may be better suited. For reference, G4 is the equivalent of the former light positive (a single +) [32], and this threshold, considering all traces (G2 and G3) as negative, underestimates the true prevalence for both high and low moderate endemicity *S. mansoni* areas. Considering all traces as positive would overestimate the prevalence, but it would still be more accurate than not considering them, with increased sensitivity closer to what a model estimate calculates than if traces are considered negative [36]. When using the G-score, considering G3 as positive is recommended [33] and strongly supported here both at a population level and individual diagnostic level.

Limitations

Whilst the G-score improves upon the older method of using trace and +, ++ and +++ scores, it is still only semi quantitative, and recorded by the naked eye, albeit in comparison to printed cassettes. Using electronic readers could eliminate inter-reader variability, but would add to the financial and logistical costs. Issues with POC-CCA batch-to-batch variation remain [20,27] and batch numbers are recommended to be reported in any associated publications. Implications of this can again be reduced by performing the G-score quality control check and reporting the data. Two batches of POC-CCA were used for this study (see Supplementary Table S1), and both batches showed a slightly lower intensity of the test band than expected.

Overall, they had similar performance to each other in our quality control check (i.e. same or one G-Score difference across the performance tests carried out), Table S1, and therefore the batch is unlikely to have affected any comparisons made here. In our study, we used two different batches for logistic reasons; with the second batch arriving when almost no test from the first batch were left, and therefore an additional inter-batch comparison would have not been possible. To fully address the effect of batches on inter and intra-sample variation, however, would have required a larger sample size and additional logistical challenges and although further test development and standardisation is recommended [27], this is outwith the remit of this paper. Furthermore, until manufacturers can guarantee standardisation, this issue will remain limiting generalisability of any studies using POC-CCAs from a limited number of batches. However, using the S-series quality control check, as we did here at least enables inter-batch comparisons both within and between papers.

Our sample was fairly balanced across genders and the age distribution was similar to the community distribution of age [50], and thus it was assumed to be representative of the population. However, it is possible that individuals recruited in this study are those with better access to treatment, which could lead to our sample underestimating the population-level disease prevalence. Conversely, people who know they are at risk may be more likely to contribute to the study. Whilst the recruitment aimed to randomly select people of different ages, it is possible that those who were recruited but did not provide all the samples may have biased the results, but this will have been minimised by the model's ability to infer missing data.

Recently, Mewamba *et al.* [51] analysed urines from 759 school-aged children in a *S. mansoni* endemic area in Cameroon and found that 55 samples that were traces with fresh urines, turned negative with the same POC-CCA batch, after being stored at -20°C for a year. If freezing affects antigens, this could affect prevalence estimates despite using the same diagnostic. Even though the manufacturer of POC-CCA states the stability of the antigens at +4°C for at least 7 weeks, and at -20°C for at least one year [12], assessing the effect of freeze-thawing on POC-CCAs would enable updated recommendations for its use in either scenario. In our study, all diagnostic tests were performed on freshly collected samples in the endemic setting, removing any issue associated with freezing. However, a portion of each urine sample was also frozen and future studies will benefit from comparing results of POC-CCA tests between fresh and frozen urine samples.

G4 is easily visible to most people reading a POC-CCA test, but G2 and sometimes even G3 are often not [25]. This inter-reader variation may also affect prevalence results, especially in low endemicity settings with low intensity of infection. In this study, only one person assigned G-scores to the POC-CCAs, t. However, inter-reader variation will be an interesting area for further study, especially in lower endemicity settings, m, where G-scores close to the threshold are more likely to be seen.

Finally, our study here was performed in only high and moderate *S. mansoni* settings, and sample sizes were only powered for investigating individual infections. Whilst soil-transmitted helminths and other commonly occurring co-infections are not thought to affect the specificity of POC-CCA tests, our results cannot be generalised to settings where *S. haematobium* may be co-endemic, and further work is also needed in low endemicity settings.

Summary

Combining different diagnostic techniques will usually improve accuracy, both at individual and population levels, especially if the combined diagnostics have high specificity. Kato-Katzs alone will require a minimum of three days with higher processing time and costs, and trained microscopists, and may still underestimate the true prevalence, especially in low endemicity settings. This technique has the advantage however, of being able to detect other helminths infections, so it should not be discarded in certain co-endemic areas. However, in areas where *S. mansoni* prevalence is expected to be lower, the use of the POC-CCA test will have significant benefits over microscopy, providing faster and more accurate results, improving precision mapping, and informing on MDA effectiveness and the potential to stop. We show for the first time that inter sample variation is far greater for POC-CCAs than intra-sample variation. At an individual level, the use of G3 as the threshold provides the best estimate of infection. However, at a population level G4 and a minimum of two to three POC-CCAs over different sampling days are needed to reach the required 95% specificity of the WHO TPP and predict the population level infection prevalence in either high or moderate endemicity areas. Multiple days are required to improve accuracy at individual level, especially for infection intensity measures. In areas of ongoing *S. mansoni* monitoring and evaluation, we recommend two days of POC-CCA, with a cut-off of G4. Whilst multiple tests are more costly, they are currently required to reach the WHO TPPs for schistosomiasis.

Acknowledgements

We would like to thank the communities in Kalachai A, Kateki, Kogala, Oburi, and Bugoto, for kindly participating in this study, as well their Village Health Teams and village elders for their assistance in mobilising the community. We also thank Alon Atuhaire and Ronald Lubowa, for their help in several logistical aspects of the field work; Dr Sergi Alonso, Dr Jessica Clark, Raheema Chunara, Thomas Arme and Rivka Lim for helping with the sample processing in the field, data entry and/or data cleaning, and Dr Thomas Crellen, for advising on data analysis. Finally, we would like to thank Dr Pytsje Hoekstra-Mevius for comments on the manuscript.

Funding statement

This work was supported by EKP's Medical Research Scotland PhD studentship awarded to PHLL, the primary supervisor (MRS PhD-1183-2017); the European Research Council (ERC Starting Grant awarded to PHLL, SCHISTO_PERSIST 680088); and the Engineering and Physical Sciences Research Council (EPSRC EP/T003618/1 awarded to JMP and PHLL).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

EKP's PhD focuses mainly on developing a POC test for the detection of the circulating anodic antigen (CAA). PHLL, JMP, MA, GvD have recently been awarded funding for a project to improve reproducibility of POC-CCA tests.

References

1. Gryseels B, de Vlas S. 1996 Worm Burdens in Schistosome Infections. *Parasitology Today* **12**. (doi:10.1016/0169-4758(96)80671-5)

2. Colley DG, Bustinduy AL, Secor WE, King CH. 2014 Human schistosomiasis. *The Lancet* **383**, 2253–2264. (doi:10.1016/S0140-6736(13)61949-2)

3. Global Schistosomiasis Alliance. 2021 GSA Communication Piece on Commercially Available Diagnostic Tests. See https://www.eliminateschisto.org/resources/communication-piece-commercially-available-diagnostic-tests.

4. Assaré RK, Tra MBI, Ouattara M, Hürlimann E, Coulibaly JT, N'Goran EK, Utzinger J. 2018 Sensitivity of the point-of-care circulating cathodic antigen urine cassette test for diagnosis of *Schistosoma mansoni* in low-endemicity settings in Côte d'Ivoire. *American Journal of Tropical Medicine and Hygiene* **99**, 1567–1572. (doi:10.4269/ajtmh.18-0550)

5. Sousa MS, Van Dam GJ, Pinheiro MCC, De Dood CJ, Peralta JM, Peralta RHS, De Francesco Daher E, Corstjens PLAM, Bezerra FSM. 2019 Performance of an Ultra-Sensitive Assay Targeting the Circulating Anodic Antigen (CAA) for Detection of *Schistosoma mansoni* Infection in a Low Endemic Area in Brazil. *Front Immunol* **10**, 1–16. (doi:10.3389/fimmu.2019.00682)

6. de Vlas S. 1992 Underestimation of *Schistosoma mansoni* Prevalences. *Parasitology Today*. **8**. (doi:10.1016/0169-4758(92)90144-q)

7. Coulibaly JT, N'Gbesso YK, Knopp S, N'Guessan NA, Silué KD, van Dam GJ, N'Goran EK, Utzinger J. 2013 Accuracy of Urine Circulating Cathodic Antigen Test for the Diagnosis of *Schistosoma mansoni* in Preschool-Aged Children before and after Treatment. *PLoS Negl Trop Dis* **7**. (doi:10.1371/journal.pntd.0002109)

8. Hoekstra PT *et al.* 2020 Efficacy of single versus four repeated doses of praziquantel against *Schistosoma mansoni* infection in school-aged children from Côte d'Ivoire based on Kato-Katz and POC-CCA: An open-label, randomised controlled trial (RePST). *PLoS Negl Trop Dis* **14**. (doi:10.1371/journal.pntd.0008189)

9. Mazigo HD, Kepha S, Kinung'hi SM. 2018 Sensitivity and specificity of point-of-care circulating Cathodic antigen test before and after praziquantel treatment in diagnosing *Schistosoma mansoni* infection in adult population co-infected with human immunodeficiency virus-1, North-Western Tanzania. *Archives of Public Health* **76**. (doi:10.1186/s13690-018-0274-4)

10. Meulah B *et al.* 2022 Performance Evaluation of the Schistoscope 5.0 for (Semi-)automated Digital Detection and Quantification of *Schistosoma haematobium* Eggs in Urine: A Field-based Study in Nigeria. *Am J Trop Med Hyg* (doi:10.4269/ajtmh.22-0276)

11. de Vlas S, Gryseels B, van Oortmarssen G, Polderman A, Habema J. 1993 A Pocket Chart to Estimate True *Schistosoma mansoni* Prevalences. *Parasitology Today* **9**, 305–307.

12. Rapid Medical Diagnostics. 2018 Schisto POC-CCA®: Rapid test for qualitative detection of Bilharzia (Schistosomiasis).

13. Stothard JR *et al.* 2009 An evaluation of urine-CCA strip test and fingerprick blood SEA-ELISA for detection of urinary schistosomiasis in schoolchildren in Zanzibar. *Acta Trop* **111**, 64–70. (doi:10.1016/J.ACTATROPICA.2009.02.009)

14. Cai P, Mu Y, Weerakoon KG, Olveda RM, Ross AG, McManus DP. 2021 Performance of the point-of-care circulating cathodic antigen test in the diagnosis of schistosomiasis japonica in a human cohort from Northern Samar, the Philippines. *Infect Dis Poverty* **10**. (doi:10.1186/s40249-021-00905-5)

15. World Health Organization. 2022 Enhancing implementation of schistosomiasis control and elimination programmes. See https://www.who.int/activities/enhancing-implementation-of-schistosomiasis-control-and-elimination-programmes.

16. Casacuberta M, Kinunghi S, Vennervald BJ, Olsen A. 2016 Evaluation and optimization of the Circulating Cathodic Antigen (POC-CCA) cassette test for detecting *Schistosoma mansoni* infection by using image analysis in school children in Mwanza Region, Tanzania. *Parasite Epidemiol Control* **1**, 105. (doi:10.1016/J.PAREPI.2016.04.002)

17. Tchuem Tchuenté LA *et al.* 2012 Evaluation of Circulating Cathodic Antigen (CCA) Urine-Tests for Diagnosis of *Schistosoma mansoni* Infection in Cameroon. *PLoS Negl Trop Dis* **6**. (doi:10.1371/journal.pntd.0001758)

18. Erko B, Medhin G, Teklehaymanot T, Degarege A, Legesse M. 2013 Evaluation of urine-circulating cathodic antigen (Urine-CCA) cassette test for the detection of *Schistosoma mansoni* infection in areas of moderate prevalence in Ethiopia. *Tropical Medicine and International Health* **18**, 1029–1035. (doi:10.1111/tmi.12117)

19. Danso-Appiah A, Minton J, Boamah D, Otchere J, Asmah RH, Rodgers M, Bosompem KM, Eusebi P, Vlas SJ De. 2016 Accuracy of point-of-care testing for circulatory cathodic antigen in the detection of schistosome infection: systematic review and meta-analysis. *Bull World Health Organ* **94**, 522. (doi:10.2471/BLT.15.158741)

20. Viana AG *et al.* 2019 Discrepancy between batches and impact on the sensitivity of point-of-care circulating cathodic antigen tests for *Schistosoma mansoni* infection. *Acta Trop* **197**. (doi:10.1016/j.actatropica.2019.105049)

21. Graeff-Teixeira C *et al.* 2021 Low specificity of point-of-care circulating cathodic antigen (POC-CCA) diagnostic test in a non-endemic area for schistosomiasis mansoni in Brazil. *Acta Trop* **217**. (doi:10.1016/j.actatropica.2021.105863)

22. Homsana A, Odermatt P, Southisavath P, Yajima A, Sayasone S. 2020 Cross-reaction of POC-CCA urine test for detection of *Schistosoma mekongi* in Lao

PDR: A cross-sectional study. *Infect Dis Poverty* **9**. (doi:10.1186/s40249-020-00733-z)

23. Casacuberta-Partal M *et al.* 2021 Specificity of the Point-of-Care Urine Strip Test for *Schistosoma* Circulating Cathodic Antigen (POC-CCA) Tested in Non-Endemic Pregnant Women and Young Children. *Am J Trop Med Hyg* **104**, 1412–1417. (doi:10.4269/AJTMH.20-1168)

24. Peralta JM, Cavalcanti MG. 2018 Is POC-CCA a truly reliable test for schistosomiasis diagnosis in low endemic areas? The trace results controversy. *PLoS Negl Trop Dis* **12**, e0006813. (doi:10.1371/journal.pntd.0006813)

25. Okoyo C, Simiyu E, Njenga SM, Mwandawiro C. 2018 Comparing the performance of circulating cathodic antigen and Kato-Katz techniques in evaluating *Schistosoma mansoni* infection in areas with low prevalence in selected counties of Kenya: A cross-sectional study. *BMC Public Health* **18**. (doi:10.1186/s12889-018-5414-9)

26. Hoekstra PT, Madinga J, Lutumba P, van Grootveld R, Brienen EAT, Corstjens PLAM, van Dam GJ, Polman K, van Lieshout L. 2022 Diagnosis of Schistosomiasis without a Microscope: Evaluating Circulating Antigen (CCA, CAA) and DNA Detection Methods on Banked Samples of a Community-Based Survey from DR Congo. *Trop Med Infect Dis* **7**. (doi:10.3390/tropicalmed7100315)

27. Colley DG, Ramzy RMR, Maganga J, Kinung'hi S, Odiere MR, Musuva RM, Campbell CH. 2023 The POC-CCA assay for detection of *Schistosoma mansoni* infection needs standardization in production and proper quality control to be reliable. *Acta Trop* **238**, 106795. (doi:10.1016/J.ACTATROPICA.2022.106795)

28. World Health Organization. 2021 Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030.

29. Clark J *et al.* 2022 Translating From Egg- to Antigen-Based Indicators for *Schistosoma mansoni* Elimination Targets: A Bayesian Latent Class Analysis Study. *Frontiers in Tropical Diseases* **3**. (doi:10.3389/fitd.2022.825721)

30. Leeflang MMG, Rutjes AWS, Reitsma JB, Hooft L, Bossuyt PMM. 2013 Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ. Canadian Medical Association Journal* **185**. (doi:10.1503/cmaj.121286)

31. Fuss A, Mazigo HD, Tappe D, Kasang C, Mueller A. 2018 Comparison of sensitivity and specificity of three diagnostic tests to detect *Schistosoma mansoni* infections in school children in Mwanza region, Tanzania. *PLoS One* **13**. (doi:10.1371/journal.pone.0202499)

32. Casacuberta-Partal M, Hoekstra PT, Kornelis D, van Lieshout L, van Dam GJ. 2019 An innovative and user-friendly scoring system for standardised quantitative interpretation of the urine-based point-of-care strip test (POC-CCA) for the diagnosis of intestinal schistosomiasis: a proof-of-concept study. *Acta Trop* **199**, 105150. (doi:10.1016/j.actatropica.2019.105150)

33. Clark J *et al.* 2022 Reconciling Egg- and Antigen-Based Estimates of *Schistosoma mansoni* Clearance and Reinfection: A Modeling Study. *Clinical Infectious Diseases* **74**, 1557–1563. (doi:10.1093/cid/ciab679)

34. Dendukuri N, Schiller I, De Groot J, Libman M, Moons K, Reitsma J, Van Smeden M. 2018 Concerns about composite reference standards in diagnostic research. *BMJ (Online)* **360**. (doi:10.1136/bmj.j5779)

35. Schiller I, van Smeden M, Hadgu A, Libman M, Reitsma JB, Dendukuri N. 2016 Bias due to composite reference standards in diagnostic accuracy studies. *Stat Med* **35**, 1454–1470. (doi:10.1002/sim.6803)

36. Clements MN *et al.* 2018 Latent class analysis to evaluate performance of point-of-care CCA for low-intensity *Schistosoma mansoni* infections in Burundi. *Parasit Vectors* **11**, 1–13. (doi:10.1186/s13071-018-2700-4)

37. Koukounari A, Jamil H, Erosheva E, Shiff C, Moustaki I. 2021 Latent Class Analysis: Insights about design and analysis of schistosomiasis diagnostic studies. , 1–23. (doi:10.1371/journal.pntd.0009042)

38. Prada JM, Touloupou P, Adriko M, Tukahebwa EM, Lamberton PHL, Hollingsworth TD. 2018 Understanding the relationship between egg- and antigen-based diagnostics of *Schistosoma mansoni* infection pre- and post-treatment in Uganda. *Parasit Vectors* **11**, 10–13. (doi:10.1186/s13071-017-2580-z)

39. Bärenbold O *et al.* 2018 Translating preventive chemotherapy prevalence thresholds for *Schistosoma mansoni* from the Kato-Katz technique into the point-of-care circulating cathodic antigen diagnostic test. *PLoS Negl Trop Dis* **12**. (doi:10.1371/journal.pntd.0006941)

40. Worrell CM, Bartoces M, Karanja DMS, Ochola EA, Matete DO, Mwinzi PNM, Montgomery SP, Secor WE. 2015 Cost analysis of tests for the detection of *Schistosoma mansoni* infection in children in western Kenya. *American Journal of Tropical Medicine and Hygiene* **92**, 1233–1239. (doi:10.4269/ajtmh.14-0644)

41. Booth M, Vounatsou P, N'goran EK, Tanner M, Utzinger J. 2003 The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Côte d'Ivoire. *Parasitology* **127**, 525–531. (doi:10.1017/S0031182003004128)

42. Utzinger J, Booth M, N'Goran EK, Müller I, Tanner M, Lengeler C. 2001 Relative contribution of day-to-day and intra-specimen variation in faecal egg counts of *Schistosoma mansoni* before and after treatment with praziquantel. *Parasitology* **122**, 537–544. (doi:10.1017/S0031182001007752)

43. Coulibaly JT *et al.* 2011 Accuracy of urine circulating cathodic antigen (CCA) test for *Schistosoma mansoni* diagnosis in different settings of Côte d'Ivoire. *PLoS Negl Trop Dis* **5**. (doi:10.1371/journal.pntd.0001384)

44. Katz N, Chaves A, Pellegrino J. 1972 A simple device for quantitative stool thick-smear technique in *Schistosoma mansoni*. *Rev Inst Med Trop Sao Paulo* **14**, 397–400.

45. Hornik K, Leisch F, Zeileis A, Plummer M. In press. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling.

46. Denwood MJ. 2016 runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *J Stat Softw* **71**. (doi:10.18637/jss.v071.i09)

47. R Core Team. 2022 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. See https://www.R-project.org/.

48. World Health Organization. 2022 WHO guideline on control and elimination of human schistosomiasis.

49. World Health Organization. 2021 Diagnostic target product profiles for monitoring, evaluation and surveillance of schistosomiasis control programmes.

50. Faust CL, Osakunor DNM, Downs JA, Kayuni S, Stothard JR, Lamberton PHL, Reinhard-Rupp J, Rollinson D. 2020 Schistosomiasis Control: Leave No Age Group Behind. *Trends Parasitol*. **36**, 582–591. (doi:10.1016/j.pt.2020.04.012)

51. Mewamba EM *et al.* 2021 Field assessment in Cameroon of a reader of POC-CCA lateral flow strips for the quantification of *Schistosoma mansoni* circulating cathodic antigen in urine. *PLoS Negl Trop Dis* **15**, e0009569. (doi:10.1371/JOURNAL.PNTD.0009569)