Title

Comparative evaluation of plasma biomarkers of *Schistosoma haematobium* infection in endemic populations from Burkina Faso

Short title

Plasma biomarkers of *Schistosoma haematobium* infection

Authors

Mireille Ouedraogo 1-2-3, Jana Christina Hey 4, Stan Hilt 5, Veronica Rodriguez Fernandez 7, Doris Winter 4, Ravo Razafindrakoto 6, Pytsje Hoekstra 5, Youssouf Kabore 3, Marco Fornili 7, Laura Baglietto 7, Issa Nebie 3, Govert J van Dam 5, Paul L Corstjens 8, Daniela Fusco 4-9, David Modiano 2, Fabrizio Bruschi 1, Valentina D Mangano 1

Affiliations

1. Dept. of Translational Research in Medicine and Surgery, University of Pisa, Italy
2. Dept. of Public Health and Infectious Diseases, University of Rome La Sapienza, Italy
3. Centre National de Recherche et Formation sur le Paludisme, Ouagadougou, Burkina Faso
4. Dept. Infectious Diseases Epidemiology, Bernard Nocht Institute for Tropical Medicine, Germany
5. Dept. of Parasitology, Leiden University Medical Center, the Netherlands
6. Centre d’infectiologie Charles Merieux, Antananarivo, Madagascar
7. Dept. of Clinical and Experimental Medicine, University of Pisa, Italy

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Infection with *Schistosoma haematobium* causes urogenital disease associated with organ dysfunction, bleeding, pain, and higher susceptibility to infections and cancer. Timely and accurate diagnosis is crucial for prompt and appropriate treatment as well as surveillance efforts, and the use of plasma biomarkers offers important advantages over parasitological examination of urine, including increased sensitivity and the possibility to use the same specimen for multiple investigations.

The present study aims to evaluate the diagnostic performance of different plasma biomarkers in endemic populations from Burkina Faso, West Africa. *Schistosoma* spp. Circulating Anodic Antigen (CAA), cell free *S. haematobium* DNA (cfDNA), class M and G antibodies against *S. haematobium* Soluble Worm Antigen Preparation (SWAP) and Soluble Egg Antigen (SEA) were measured in 406 plasma samples. Results of each biomarker test were compared to those of CAA, a Composite Reference Standard (CRS) and Latent Class Analysis (LCA).

An identical proportion of positive samples (29%) was observed as a result of CAA and cfDNA testing, with a substantial agreement (84%, Cohen $k=0.62$) between the results of the two tests, and a comparable agreement with the results of CRS and LCA. A higher positivity was
observed, as expected, as a result of specific antibody testing (47%-72%), with IgG showing a higher agreement than IgM with the three references. Also, higher IgG levels were observed in current vs past infection, and ROC analysis identified optimal cutoff values for improved testing accuracy.

This study provides compelling evidence that can inform the choice of the most appropriate diagnostic plasma biomarker for urogenital schistosomiasis in endemic areas, depending on the purpose, context, and available resources for testing. Either CAA or cfDNA testing can be used for the diagnosis of patients and for epidemiological investigations, even in absence of urine filtration microscopy, whereas anti-SWAP or anti-SEA IgG can be employed for surveillance and integrated monitoring of control interventions against poverty-associated diseases.

Keywords

Schistosomiasis, Schistosoma haematobium, diagnosis, plasma biomarkers, Neglected Tropical Diseases, Burkina Faso

Author summary

Urogenital schistosomiasis is a chronic debilitating disease affecting populations living in Africa and the Middle East and showing a strong association with poverty. Accurate detection of infection is important both for disease treatment and surveillance. Several tests
based on detection in plasma of parasite protein (CAA), parasite DNA or parasite-specific host antibodies (IgM and IgG against SWAP and SEA antigens) are available and this study aims at comparing them to evaluate their accuracy. The comparison showed that tests based on parasite CAA or DNA yield very similar results and therefore the test of choice for diagnosis or epidemiological investigations can be based on laboratory resources. Additionally, the comparison showed that IgG against SWAP and SEA outperform IgM, and that high accuracy can be achieved by identifying an optimal level to determine positivity (cut-off), making these antibody tests ideal for surveillance purposes.
Introduction

Human schistosomiasis is a poverty-associated parasitic diseases caused by trematode blood flukes of the genus *Schistosoma*. It is estimated that 250 million people are infected in 78 endemic countries, and that 90% of cases occur in Sub-Saharan Africa (SSA) [1-3].

*S. haematobium*, endemic in Africa and the Middle East, and since recent years in some areas of Europe [4-6], is the only species causing urogenital disease. Infection occurs by contact with fresh-water bodies infested with free-swimming cercariae larvae released by specific intermediate snail hosts of the genus *Bulinus*. Cercariae penetrate the skin of the human host, shed their fork and become schistosomula, which then migrate via venous circulation to the lungs, the heart and the liver where they mature to the adult stage. Adult worms exit the liver via the portal vein system and reach the venules of the bladder, where they reproduce and gravid females lay eggs. While most eggs remain trapped in venules and body tissues, some are eliminated with urine. When the human host discharges urine in fresh-water bodies, the eggs hatch and release miracidia, which penetrate the snail host and transform into sporocysts. After two generation of sporocysts larval reproduction, these transform into cercariae, able to initiate a new transmission cycle [7-9].

The disease is mostly accounted by immunopathogenic responses to the worms’ eggs accumulated around bladder and ureters resulting in granuloma formation and tissue fibrosis. Clinical manifestations include haematuria, obstructive uropathy, kidney damage and squamous cell carcinoma [7-9]. In females, genital schistosomiasis (FGS) causes
ulcerative lesions and fibrosis in vagina, cervix and uterus, resulting in bleeding, pain and higher risk of HIV infection [10-11]. Control of schistosomiasis mainly focuses on reducing disease burden through periodic, large-scale population treatment with praziquantel, but more comprehensive approaches include water, hygiene and sanitation strategies, health education and snail control [2-3].

Diagnosis of schistosomiasis is fundamental for appropriate treatment of symptomatic individuals as well as for monitoring the success of control efforts, and a variety of direct (detecting the parasite or parasite molecules) and indirect (detecting the human immune responses to the parasite) laboratory methods are available, each with different advantages and disadvantages. Microscopic examination of urine samples (urine filtration microscopy) for detection of S. haematobium eggs has optimal specificity based on distinct morphological features. However, sensitivity is hampered by the irregular timing of eggs release in urine and by their limited number, particularly in low intensity infections, and the examination of at least three specimens is required [12-13]. Other direct methods include the detection of intestinal antigens regurgitated by Schistosoma adult worms and therefore released in circulation, namely Circulating Cathodic Antigen (CCA) and Circulating Anodic Antigen (CAA), in either urine or serum/plasma samples by immunological assays such as ELISA and ICT, including the point-of-care CCA (POC-CCA) test and the upconverting particle lateral flow CAA (UCP-LF CAA) test [14]. Compared to CCA, CAA testing has shown a better sensitivity for the diagnosis of infection by different Schistosoma species, including S. haematobium, and in areas of low endemicity [15]. It has also been reported that CAA concentration shows little day to day fluctuations, positively correlates with the number of
adult worms, and decreases sharply after drug treatment, establishing it as a valuable test for both diagnosis and follow-up [16]. Nonetheless, it must be noted that infection with larval stages cannot be diagnosed by such antigen tests. Detection of cell-free DNA (cfDNA) of schistosomes in either stool/urine or plasma by molecular assays represents a promising method for early direct diagnosis, since it can detect all stages of infection including larval stages, and since optimal analytical sensitivity and specificity can be achieved by choice of multi-copy specific targets [17]. For *S. haematobium*, Real Time PCR protocols based on the detection of *Dra1*, a repeat sequence specific of this species and representing >15% of its genome, have been recently developed and tested in migrant and travellers as well as residents in endemic areas [18-22]. However, cfDNA may circulate for some time after successful treatment and its detection might not therefore always indicate current infection [17]. Regarding indirect methods, these are based on the detection of specific antibodies against parasite antigens in serum/plasma samples by immunological methods such as ELISA, ICT and more recently Luminex [23-24]. The main advantage of these methods is increased sensitivity compared to urine filtration microscopy, as demonstrated in non-endemic contexts [25-26]. However, as production of antibodies by plasma cells only starts about 1 week after exposure to a given antigen, a temporal delay is to be expected between infection and the detection of specific antibodies. More importantly, antibodies can still be detectable in plasma after drug treatment and parasite death, both because of antigen persistence and continued exposure to the immune system, and because of the long half-life of class G immunoglobulins (Ig). Therefore, seropositivity might reflect either history of infection or current infection or both. Antigens employed for *S. haematobium* serology have
long included soluble excretory/secretory products from the different stages, such as Cercarial Antigen Preparation (CAP), Soluble Worm Adult Preparation (SEA) and Soluble Egg Antigen (SEA) [21, 27], and, more recently, recombinant proteins of the tetraspanin family that were identified as promising candidates through immunomic analyses [28-29]. It is noteworthy that no commercial in vitro diagnostic tests for the specific diagnosis of *S. haematobium* exist to date.

As described above, several assays are available for laboratory diagnosis of *S. haematobium* using plasma samples. The use of plasma samples offers the important advantage of allowing the simultaneous detection of biomarkers of infection with different pathogens in the same specimen. This approach enables increased efficiency not only in the diagnosis of concomitant infections or the differential diagnosis of infections with overlapping clinical features in patients, but also in conducting epidemiological studies and surveillance activities of different infectious diseases such as malaria and NTDs, including the integrated monitoring of control interventions impact [30-32]. A further advantage is the possibility of retrospective analysis of archived plasma samples for further investigations.

Despite the benefits offered by plasma assays for *S. haematobium* diagnosis, to the best of our knowledge, a comparative evaluation of their diagnostic performance has not yet been conducted in endemic populations. To this aim, we took advantage of archived plasma samples collected during malariological surveys in an area of *S. haematobium* and malaria co-endemicity in Burkina Faso [33].
**Methods**

**Study design**

The present study aimed at measuring plasma biomarkers of *S. haematobium* infection in samples collected during a cross-sectional survey conducted among endemic populations from Burkina Faso, and at performing a comparative evaluation of their diagnostic performance.

The cross-sectional survey was conducted in August 2007 in two rural villages (Barkoumbilen and Barkoundouba) of shrubby savannah areas of Burkina Faso northeast of the capital town Ouagadougou, for malaria epidemiology purposes [34]. A 2 mL venous blood sample was collected from each participant in ethylenediaminetetraacetic acid (EDTA) tubes. Plasma was separated by centrifugation (3min at 2000 rpm), aliquoted and stored at -80°C until assays were performed. The study area is endemic for *S. haematobium* [1, 33].

Assays have been performed on 406 archived plasma samples to measure the following biomarkers: *Schistosoma* spp. Circulating Anodic Antigen (CAA), *S. haematobium* cell-free circulating DNA (cfDNA), *S. mansoni* cfDNA, IgM and IgG against *S. haematobium* Soluble Worm Adult Protein (SWAP) and Soluble Egg Antigen (SEA). The species-specific results of cfDNA testing were used to confirm that the study area is endemic for *S. haematobium* only.

The evaluation of the diagnostic performance of the different biomarker tests was conducted comparing the results to those of three different references (i.e. defining true infection status): i) Circulating Anodic Antigen; ii) Composite Reference Standard; iii) Latent Class Analysis.
**Ethics statement**

Study subjects or their guardians gave written informed consent for participation. The original malaria epidemiological study received approval from the ethical committee of the Ministry of Health of Burkina Faso (2007-048). The use of archived plasma samples for schistosomiasis research received further approval from the ethical committee of the Ministry of Health and Public Hygiene of Burkina Faso (2022-05-15).

**Circulating Anodic Antigen**

The Up-Converting Particle Lateral Flow (UCP-LF) assay was conducted in Leiden University Medical Center, the Netherlands, to detect Circulating Anodic Antigen (CAA) in plasma samples as previously described (SCAA20) [16, 35]. High Salt Lateral Flow (HSLF) buffer (100 mM HEPES pH 7.5, 270 mM NaCl, 0-5% v/v Tween-20, 1% w/v BSA, 50 µL) was pipetted into the wells of a 96-wells microplate. Each plasma sample (50 µL) was mixed with an equal volume of 4% trichloroacetic acid (TCA) and centrifuged for 5 minutes at 18,000 RCF to separate the supernatant from the pellet. The supernatant (20 µL) and a UCP solution (50 µL) containing reporter particles (400 nm $\text{Y}_2\text{O}_3\text{S}:\text{Yb}^{3+},\text{Er}^{3+}$) conjugated with mouse monoclonal anti-CAA antibody (25 µg/mg) were pipetted in the microplate wells. After an incubation of 1 hour at 37°C with shaking (900 rpm), LF strips with a test line comprising 200 ng mouse monoclonal anti-CAA antibody were added into the wells. A dilution series of a sample with known CAA concentration (pg/ml) was included in the microplate and used as a standard curve to quantify CAA concentrations and to validate the threshold of the assay (10 pg/ml). Strips were incubated overnight and scanned with an UPCON reader (980 nm excitation, 540
nm emission; Labrox Oy, Turku, Finland) for analysis. A threshold of 10 pg/ml (SCAA20, wet test format) was used to define CAA positivity [16, 35]. CAA levels were defined based on concentration ranges as follows: level 0=0-9 pg/ml range; level 1=10-99 pg/ml range; level 2=100-999 pg/ml range; level 3=1000-9999 pg/ml range; level 4=10000-100000 pg/ml range.

**Cell-free circulating DNA**

Cell free circulating DNA (cfDNA) was extracted from 100 µL plasma samples using the QIAamp Min Elute cfDNA Mini Kit according to the manufacturer’s protocol. Real Time PCR was performed at the Bernard Nocht Institute for Tropical Medicine, Germany, to amplify and detect *S. heamatobium* (*Dra1*) and *S. mansoni* (*Sm1-7*) target sequences (Sh-Forward Primer: 5’-GATCTCACCTATCAGACG AAAC-3’, Sh-Reverse Primer: 5’-TCACAACGATACGACCAAC-3’, Sh-Probe: 5’-Joe-TGTTGGAAGZGCCTGTTTCGCAA-BHQ1-3’; Sm-Forward Primer: 5’-CAACCGTTCTATGAA AATCGTTGT-3’, Sm-Reverse Primer : 5’-CCA CGCTCTC GCAAATAATCT-3’, Sm-Probe: 5’-Fam-TCCGAAACCACTGGACGGATTTTTATGAT-BHQ1-3’) as previously described [21-22]. Primers and probe were included for the detection of Phocid herpesvirus (PhHV) DNA as internal PCR control (PhHV—Forward Primer: 5’ GGG CGA ATC ACA GAT TGA ATC 3’, PhHV—Reverse Primer: 5’ GCG GTT CCA AAC GTA CCA A 3’, PhHV—Probe: 5’ Cy5.5-TTT TTA TGT GTC CGC CAC CA-BBQ 3’). Briefly, the reaction was performed in a total volume reaction of 25 µL containing 12.5 µl HotStarTaq Mastermix (Qiagen, Hilden, Germany), 5 mM MgCl2, 500 nM of each *Schistosoma* primer, 250 nM of each *Schistosoma* probe, 40 nM of the PhHV primer, 50 nM of the PhHV probe, 1.425 µg of the PhHV DNA template, 0.04 µg/µl
bovine serum albumin and 5 µL of DNA eluate. The amplification reaction (15 minutes at 95°C followed by 50 cycles of 15 seconds at 95°C and 60 seconds at 60°C with an initial touchdown from 65°C to 60°C in the first 11 cycles) was conducted using the Corbett Rotor-Gene 6000 (Qiagen) and results were analysed with the RotorGene 6000 Software v.7.87 (Qiagen). Samples with a clean sigmoid curve within the PCR cycles were considered positive.

Specific immunoglobulins

Specific IgM and IgG against *S. haematobium* Soluble Worm Adult Protein (SWAP) and Soluble Egg Antigen (SEA) were measured by ELISA using an in house protocol adapted from Mutapi and colleagues [27] as previously described [33]. Antigens were purchased from the Theodor Bilharz Institute, Giza, Egypt. A volume of 100 µL antigen (5µg/mL for SWAP and 10µg/mL for SEA) in coating buffer (0.5 M sodium carbonate pH 9.6) was added per well in 96-wells plates (Nunc Immunosorp), and plates were incubated overnight at 4°C. After 3 washings (PBS-0.03 % Tween 20, PBS-T), 200µL blocking buffer (5 % skimmed milk powder in PBS/T) was added per well and plates incubated 2h at room temperature (RT). After 3 washings, plasma samples were added in duplicate in blocking buffer with a total volume of 100µL (1:200 dilution for IgM and 1:100 dilution for IgG) and plates were incubated 2h at RT. After three washings 100 µL HRP-anti-human-IgG (Dako) at 1:1000 dilution in PBS-T was added and plates incubated for 2h at RT. After six washings 100 µL OPD (Sigma) substrate solution was added, and plates incubated for 20 min in the dark at RT. A volume of 25 µL of H₂SO₄ 1 M stop solution was added and plates were read with Synergy HT Biotek reader at the absorbance of 492 nm. A standard curve consisting of a 5-point dilution series of a pool
of positive control (patients with positive result of urine filtration microscopy) samples was included in each plate as well as a negative control consisting of a pool of negative control (naïve blood donors) samples and a blank control containing no plasma. The OD value of the blank control was subtracted from the OD values of test samples, and those were then fitted to account for experimental variation using the standard curves (fitted OD, fOD). The cut-off used for positivity was the mean absorbance plus 3 standard deviation of negative control samples (cutoff = mean OD neg + 3SD neg). Logarithm transformation has been applied to fOD values (logfOD) to normalise their distribution.

Statistical analysis

In reporting binary (i.e. positive/negative) test results, the proportion of positive results was computed with its 95% Confidence Interval (CI), shown using barplots and compared between population groups by logistic regression.

In reporting quantitative (i.e. antibody levels or antigen concentration) test results, the distribution was shown using boxplots and compared between population groups by linear regression. Correlation between quantitative test results was displayed using scatter plots and assessed by Spearman test.

The diagnostic performance of each test has been evaluated by comparison of binary results with three references: i) Circulating Anodic Antigen; ii) Composite Reference Standard; iii) Latent Class Analysis. For each comparison, sensitivity and specificity with their respective 95% CI were computed, as well as overall agreement and Cohen’s k statistics. Cohen’s k was interpreted as follows: values ≤ 0 as no agreement, 0.01–0.20 as none to slight, 0.21–0.40 as
fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement.

CAA was chosen as a reference since it is a validated marker of infection [15-16].

A Composite Reference Standard (CRS) was defined as follows: a sample was considered positive if one or both direct (i.e. CAA and/or cfDNA) test results were positive, while it was considered negative if both direct tests results were negative. The rationale of the algorithm is to define infection status based on direct markers tests, and it assumes that both CAA and cfDNA have 100% specificity [36].

Latent Class Analysis (LCA) modelling with two clusters was used to define a binary (positive/negative) LCA reference [26]. The model included all available manifest variables: CAA (positive/negative), cfDNA (positive/negative), anti-SWAP IgM level (log10-fOD), anti-SWAP IgG level (log10-fOD), anti-SEA IgG level (log10-fOD), anti-SEA IgM level (log10-fOD).

The LCA reference therefore defines infection status based on the results of all direct and indirect biomarker tests.

Receiver Operating Characteristic (ROC) analysis was used to evaluate the ability of antibody levels (logfOD) to discriminate between positive and negative reference results, by calculating Area Under Curve values and by computing specificity, sensitivity and overall agreement for each possible cut-off value. An optimal cut-off value was identified by filtering values associated with >75% agreement, sensitivity and specificity, ordering them from higher to lower agreement, sensitivity and specificity, and selecting the top value. A cutoff Antibody Index (AI) was calculated as the ratio between the value of the cutoff and the value of the negative control.
All analyses were performed in STATA v13.00, except for the LCA analysis that was conducted using R v. 4.3.1 and its package *flexmix*.

**Results**

Real Time PCR did detect *S. haematobium* cfDNA in the study sample, but not *S. mansoni* cfDNA (data not shown), as expected from the literature [1]. Therefore, it can be assumed that CAA in this area is a product of *S. haematobium* and therefore a biomarker of urogenital schistosomiasis. The proportion of positive results for each plasma biomarker of *S. haematobium* infection is shown in Table 1. CAA and cfDNA tests result in an identical proportion of positive subjects, equal to 29%. Specific antibody tests result in a higher proportion of positive subjects, ranging from 47% to 72%, with anti-SEA Ig showing higher positivity than anti-SWAP Ig.
Table 1. Proportion of positive subjects for plasma biomarkers of *S. haematobium*

The table shows, for each plasma biomarker of *S. haematobium* infection, the number of positive subjects (N+), the number of subjects in the study population (Ntot), and the proportion of positive subjects (%) with its lower (LCL) and upper (UCL) 95% Confidence Limits.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Positive subjects</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N+</td>
<td>Ntot</td>
<td>Proportion+ (%)</td>
<td>LCL</td>
<td>UCL</td>
</tr>
<tr>
<td>CAA</td>
<td>118</td>
<td>406</td>
<td>29.1</td>
<td>24.8</td>
<td>33.7</td>
</tr>
<tr>
<td>cfDNA</td>
<td>118</td>
<td>406</td>
<td>29.1</td>
<td>24.8</td>
<td>33.7</td>
</tr>
<tr>
<td>anti-SWAP IgM</td>
<td>192</td>
<td>406</td>
<td>47.3</td>
<td>42.5</td>
<td>52.2</td>
</tr>
<tr>
<td>anti-SWAP IgG</td>
<td>223</td>
<td>406</td>
<td>54.9</td>
<td>50.0</td>
<td>59.7</td>
</tr>
<tr>
<td>anti-SEA IgM</td>
<td>292</td>
<td>406</td>
<td>71.9</td>
<td>67.3</td>
<td>76.1</td>
</tr>
<tr>
<td>anti-SEA IgG</td>
<td>282</td>
<td>406</td>
<td>69.5</td>
<td>64.8</td>
<td>73.8</td>
</tr>
</tbody>
</table>

The breadth of antibody response to *S. haematobium* (*Sh*) antigens (i.e. the total number of specific Ig for which a positive result was observed) and the different patterns of positivity are shown in Table 2. Overall, 15% of subjects were seronegative for anti-*Sh* Ig. Among subjects who were seropositive for only one specific Ig (12%), the majority was positive for anti-SEA IgG. Among subjects who were seropositive for two specific Ig (18%), the two most and equally frequent combinations were anti-SWAP IgM/anti-SWAP IgG and anti-SWAP IgG/anti-SEA IgG. Among subjects who were seropositive for three specific Ig (26%), the majority was positive for all Ig except anti-SWAP IgM. Finally, about 29% of subjects were positive for all four specific Ig. In order to investigate the diagnostic potential of indirect
markers compared to direct ones, the positivity of CAA and cfDNA direct markers according
to the breadth of the anti-Sh Ig response was assessed by logistic regression analysis. It can
be observed that both CAA (OR=2.4, 95%CI=1.9-3.0, p-value<0.001) and cfDNA (OR=2.6,
95% CI=2.1-3.4, p-value<0.001) positivity increase significantly with the number of positive
specific Ig (Figure 1).

**Table 2. Patterns of antibody response against S. haematobium antigens.**

The table shows the breadth of the antibody response to *S. haematobium* antigens
(number of positive anti-Sh Ig) and the different positivity patterns of IgM and IgG
responses to SWAP and SEA antigens, with their relative proportion (%) in the total study
population.

<table>
<thead>
<tr>
<th>Breadth anti-Sh Ig+ (N)</th>
<th>Proportion (%)</th>
<th>anti-SWAP</th>
<th>anti-SEA</th>
<th>Proportion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IgM</td>
<td>IgG</td>
<td>IgM</td>
</tr>
<tr>
<td>0</td>
<td>14.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>26.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>29.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The formal evaluation of the diagnostic performance of the different biomarkers was conducted by comparing each test results to those of three different references (i.e. “true” infection status): i) Circulating Anodic Antigen (CAA); ii) Composite Reference Standard (CRS); iii) Latent Class Analysis (LCA). For each comparison, sensitivity and specificity were calculated as well as the overall agreement.

In comparison with CAA as the reference test (Table 3), cfDNA showed 73% sensitivity and 89% specificity, with a Cohen’s k of 0.64, indicating substantial agreement. Indeed, despite resulting in an identical proportion of positive samples, the two tests did not yield identical results (Supplementary Figure 1). The comparison of CAA and cfDNA results stratified by CAA level shows that the proportion of “false negative” cfDNA results decreases significantly with increasing CAA levels (OR=0.23, 95% CI=0.17-0.31, p-value<0.001; Figure 2). Among specific antibodies, Cohen’s k ranged from 0.09 to 0.37, indicating none to fair agreement, with the best performance being observed for anti-SWAP IgG, showing 88% sensitivity and 59% specificity.
Table 3. Diagnostic performance of plasma biomarkers compared to Circulating Anodic Antigen.

The table shows, for each plasma biomarker of *S. haematobium* infection, the sensitivity (Se) and specificity (Sp) with their lower (LCL) and upper (UCL) 95% Confidence Limits, the overall agreement and the Cohen’s K value, when compared to Circulating Anodic Antigen as the reference.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Circulating Anodic Antigen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Se</td>
</tr>
<tr>
<td>CAA</td>
<td>-</td>
</tr>
<tr>
<td>cfDNA</td>
<td>73%</td>
</tr>
<tr>
<td>anti-SWAP IgM</td>
<td>55%</td>
</tr>
<tr>
<td>anti-SWAP IgG</td>
<td>88%</td>
</tr>
<tr>
<td>anti-SEA IgM</td>
<td>94%</td>
</tr>
<tr>
<td>anti-SEA IgG</td>
<td>95%</td>
</tr>
</tbody>
</table>

A CRS was built using an algorithm that classified a sample as positive if it exhibits positive results of either or both direct marker tests (CAA+ OR cfDNA+), and as negative if they exhibit negative results of both direct marker tests (CAA- AND cfDNA-), irrespectively of the results of indirect marker tests. This CRS resulted in a 37% proportion of positive subjects (data not shown). In comparison with the CRS (Table 4), both CAA and cfDNA showed 79% sensitivity and 100% specificity (as expected), with an overall almost perfect agreement (92%, Cohen’s K=0.82). Among specific antibodies, Cohen’s k ranged from 0.11 to 0.43, indicating slight to moderate agreement, and once again the best performance was observed for anti-SWAP IgG, showing 85% sensitivity and 63% specificity.
Table 4. Diagnostic performance of plasma biomarkers compared to a Composite Reference Standard.

The table shows, for each plasma biomarker of *S. haematobium* infection, the sensitivity (Se) and specificity (Sp) with their lower (LCL) and upper (UCL) 95% Confidence Limits, the overall agreement and the Cohen’s K value, when compared to a Composite Reference Standard.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Composite Reference Standard</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Se</td>
<td>LCL</td>
<td>UCL</td>
<td>Sp</td>
<td>LCL</td>
<td>UCL</td>
<td>Agreement</td>
</tr>
<tr>
<td>CAA</td>
<td>79%</td>
<td>72%</td>
<td>85%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>92%</td>
</tr>
<tr>
<td>cfDNA</td>
<td>79%</td>
<td>72%</td>
<td>85%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>92%</td>
</tr>
<tr>
<td>anti-SWAP IgM</td>
<td>55%</td>
<td>47%</td>
<td>63%</td>
<td>57%</td>
<td>51%</td>
<td>63%</td>
<td>56%</td>
</tr>
<tr>
<td>anti-SWAP IgG</td>
<td>85%</td>
<td>79%</td>
<td>90%</td>
<td>63%</td>
<td>57%</td>
<td>68%</td>
<td>71%</td>
</tr>
<tr>
<td>anti-SEA IgM</td>
<td>94%</td>
<td>90%</td>
<td>98%</td>
<td>41%</td>
<td>35%</td>
<td>47%</td>
<td>61%</td>
</tr>
<tr>
<td>anti-SEA IgG</td>
<td>94%</td>
<td>90%</td>
<td>98%</td>
<td>45%</td>
<td>39%</td>
<td>51%</td>
<td>63%</td>
</tr>
</tbody>
</table>

Finally, result of each biomarker were compared with results of LCA modelling (Table 5), where a sample was classified as positive or negative based on all available biomarker data, and that resulted in a 35% proportion of positive subjects (data not shown). In comparison with the LCA, both CAA and cfDNA showed substantial agreement, with Cohen’s k equal to 0.62 and 0.75 respectively. Among specific antibodies, Cohen’s k ranged from 0.15 to 0.55, indicating slight to moderate agreement. Consistently with previous comparisons, the best performance was observed for anti-SWAP IgG, showing 97% sensitivity and 67% specificity.
Table 5. Diagnostic performance of plasma biomarkers compared to Latent Class Analysis classification.

The table shows, for each plasma biomarker of *S. haematobium* infection, the sensitivity (Se) and specificity (Sp) with their lower (LCL) and upper (UCL) 95% Confidence Limits, the overall agreement and the Cohen’s K value, when compared to Latent Class Analysis classification.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Se</th>
<th>LCL</th>
<th>UCL</th>
<th>Sp</th>
<th>LCL</th>
<th>UCL</th>
<th>Agreement</th>
<th>Cohen’s K</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAA</td>
<td>67%</td>
<td>60%</td>
<td>75%</td>
<td>92%</td>
<td>89%</td>
<td>95%</td>
<td>83%</td>
<td>0.62</td>
</tr>
<tr>
<td>cfDNA</td>
<td>76%</td>
<td>69%</td>
<td>83%</td>
<td>97%</td>
<td>94%</td>
<td>99%</td>
<td>89%</td>
<td>0.75</td>
</tr>
<tr>
<td>anti-SWAP IgM</td>
<td>58%</td>
<td>50%</td>
<td>66%</td>
<td>58%</td>
<td>52%</td>
<td>64%</td>
<td>58%</td>
<td>0.15</td>
</tr>
<tr>
<td>anti-SWAP IgG</td>
<td>95%</td>
<td>92%</td>
<td>99%</td>
<td>67%</td>
<td>61%</td>
<td>73%</td>
<td>77%</td>
<td>0.55</td>
</tr>
<tr>
<td>anti-SEA IgM</td>
<td>99%</td>
<td>97%</td>
<td>100%</td>
<td>43%</td>
<td>37%</td>
<td>49%</td>
<td>63%</td>
<td>0.34</td>
</tr>
<tr>
<td>anti-SEA IgG</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>47%</td>
<td>41%</td>
<td>53%</td>
<td>66%</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Receiver Operating Curve (ROC) analysis was conducted to further evaluate the diagnostic performance of *S. haematobium* specific antibodies. The Area Under Curve (AUC) statistics, a measure of the ability of a marker to correctly classify positive and negative cases as defined by a reference standard, was computed for each antibody test against the three references (CAA, CRS, LCA). The comparison of AUC allows to identify the marker showing the best diagnostic performance: the highest AUC value was observed for anti-SEA IgG, followed closely by anti-SWAP IgG, for each reference (Table 6; Supplementary Figures 2-4).
Table 6. Diagnostic performance of antibody tests resulting from ROC analysis.

The table shows, for each *S. haematobium* antibody test, the Area Under Curve (AUC) resulting from ROC analysis conducted to evaluate the diagnostic performance compared to Circulating Anodic Antigen (CAA), Composite Reference Standard (CRS), and Latent Class Analysis classifications.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Area Under Curve (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CAA</td>
</tr>
<tr>
<td>anti-SWAP IgM</td>
<td>0.567</td>
</tr>
<tr>
<td>anti-SWAP IgG</td>
<td>0.826</td>
</tr>
<tr>
<td>anti-SEA IgM</td>
<td>0.765</td>
</tr>
<tr>
<td>anti-SEA IgG</td>
<td>0.878</td>
</tr>
</tbody>
</table>

For anti-SWAP and anti-SEA IgG, ROC analysis was employed for further comparison with CAA to generate Sensitivity/Specificity plots (Figure 3), which show that selecting a higher cutoff value compared to the analytical one would result in better diagnostic performances. For instance, for anti-SWAP IgG, a cut-off of 0.196 log(fOD) corresponding to a 2.3 Antibody Index (AI) would result in 75% sensitivity, 76% specificity and 76% agreement. Similarly, for anti-SEA IgG, a cut-off of 0.242 log(fOD) corresponding to a 4.5 AI would result in 82% sensitivity, 80% specificity and 81% agreement (data not shown). Also, the distribution of antibody levels according to CAA positivity was investigated, as well as the correlation between antibody levels and CAA concentration ranges. The distribution of antibody levels was significantly higher in subjects that were positive to CAA than in negative subjects, for both anti-SWAP IgG (Exp β=15.3, 95%CI=9.8-23.8, p-value<0.001; Figure 4, left panel) and anti-SEA IgG (Exp β=13.3, 95%CI=9.4-18.8, p-value<0.001; Figure 4, right panel). Finally, a
scarce albeit significant positive correlation was observed between antibody levels and CAA concentration, both for anti-SWAP IgG (Spearman P-value<0.001, rho=0.425; Figure 5 left panel) and anti-SEA IgG (Spearman P-value<0.001, rho=0.437; Figure 5, right panel).

Discussion

In summary, the measurement of plasma biomarkers of *S. haematobium* infection among 406 subjects living in rural endemic villages of Burkina Faso, West Africa, has shown that both CAA and cfDNA testing result in a 29% proportion of positive samples. A 29% prevalence of urogenital schistosomiasis in this area confirms an estimate for Burkina Faso based on systematic review of survey data and geostatistical modelling (26.3%) [1]. However, it likely represents an underestimation, as suggested by the observation that CRS and LCA show a proportion of positive subjects of 37% and 35%, respectively. Indeed, despite an identical proportion of positive samples identified by CAA and cfDNA testing, the two tests did not yield identical results. Discordant CAA negative/cfDNA positive results could be attributed to differences in the stage specificity of the two markers, with cfDNA detecting all stages including larval ones while CAA detecting adults [37], and/or by recently cleared infections with cfDNA persisting in circulation longer than CAA [38]. On the other hand, discordant CAA positive/cfDNA negative results could be attributed to a suboptimal cfDNA sensitivity at low CAA concentrations because of an insufficient plasma volume used for nucleic acid extraction in this study (1000µl of plasma was used in a previous work [22] compared to the 100µl used in the present study), an hypothesis supported by the
observation that the proportion of “false negative” cfDNA results decreases significantly with increasing CAA levels. Despite some discordant results, there was a substantial agreement between CAA and cfDNA testing, along with a comparable agreement with the results of CRS and LCA. Taken together, the identical proportion of positive samples identified by the two tests, and the substantial agreement of the results, indicate that CAA and cfDNA are equivalent good plasma biomarkers of infection in areas of *S. haematobium* endemicity.

As expected, specific antibody testing resulted in a much higher proportion of positive samples (47%-72% range), as the presence of specific antibodies might reflect not only current infection but past exposure to the parasite. Interestingly, it was observed that the breadth of the antibody response to *S. haematobium* was positively associated with both CAA and cfDNA positivity, indicating that subjects harbouring antibodies against multiple *S. haematobium* antigens are more likely to be currently infected. The analysis of individual responses has shown that IgG against SWAP and SEA exhibit a higher agreement with the three references (CAA, CRS and LCA) compared to IgM. When compared to CAA, ROC analysis confirmed a better ability of IgG to distinguish positive and negative subjects, and showed that higher cutoff values result in greatly enhanced specificity and overall diagnostic performance of these indirect biomarkers. Indeed, IgG levels were significantly higher in CAA positive than CAA negative subjects, indicating that antibody levels are increased in subjects with current infection compared to subjects with past exposure, and highlighting the relevance of the choice of appropriate cutoff values for accurate detection of infection. However, only a moderate correlation between IgG antibody levels and CAA concentration...
was observed, indicating that IgG levels cannot be used as markers of *S. haematobium* infection intensity.

One limitation of the present study is the lack of results from urine filtration microscopy, as urine specimens were not collected during the original investigation survey. Also, the antibody response to CAP was not investigated, as this antigen could not be procured at the time of analysis. Since CAP is a cercarial antigen, data would have been beneficial for interpreting discordant results between cfDNA and CAA testing, by either supporting or refusing the hypothesis that very recent infections with larval stages would be detected by cfDNA but not CAA.

Despite the above limitations, the comparative evaluation of plasma biomarkers presented herein represents, to our knowledge, the first conducted among endemic populations. The data provide valuable insights that can inform the choice of testing to be employed in areas endemic for urogenital schistosomiasis, depending on the purpose, context, needs and available laboratory resources. For diagnostic “test-to-treat” purposes, when infection with *S. haematobium* is suspected, either CAA or cfDNA testing can be performed for confirmation, even in the absence of urine filtration microscopy, while anti-SWAP and anti-SEA IgG testing could be performed for supporting evidence. A similar strategy is currently recommended in non-endemic countries for the diagnosis of schistosomiasis [39]. It is noteworthy that only cfDNA is a species-specific marker, and therefore CAA positive results do not exclude an infection with other *Schistosoma* species. However, only CAA provides a quantitative measurement (i.e. CAA concentration) associated with worm burden that can be used for follow-up of treated patients [40]. Likewise, for epidemiological investigations,
either CAA or cfDNA testing can be performed to assess infection in areas of *S. haematobium* endemcity depending on laboratory resources, since they are equivalent good markers. As mentioned above, it should also be considered that only cfDNA testing would provide a confirmation of the species while only CAA testing results can be analyzed as a quantitative variable which might be an advantage for association studies. Finally, for surveillance purposes, once the appropriate cut-off has been established for a given epidemiological context to ensure optimal agreement with markers of current infection, anti-SWAP or anti-SEA IgG testing could be either employed, with the important advantages of allowing multiplexing with other antibody tests as well as automation.

The availability of different accurate diagnostic tests for *S. haematobium* infection based on the detection of plasma biomarkers represents a significant development in the control of urogenital schistosomiasis in endemic areas.

**Authors contributions**

MO, JH, SH, VRF, DW, RR and PH conducted the experiments. YK and IN coordinated the cross-sectional survey. MF and LB performed Latent Class Analysis. GJvD, PC, DF, DM, FB and VM designed experiments and contributed experimental resources. FB, DM and VDM designed the study. VDM performed data analyses and wrote the manuscript. All authors revised and approved the final version of the manuscript.
Funding

No specific funding was obtained to conduct the research herein presented. M Ouedraogo was supported by the PhD School in Microbiology, Infectious Diseases and Public Health, Sapienza University of Rome. VD Mangano was supported by the University of Pisa and the Minister of University and Research. D Fusco was supported by the German Centre for Infection.

Acknowledgments

The authors would like to thank participants for making the study possible; physicians, nurses, laboratory technicians, data management staff at Centre National de Recherche et Formation sur le Paludisme for their skilled work; Zeno Bisoffi at Centre for Tropical Diseases in Negrar, Italy, for contributing sera samples of patients infected with *S. haematobium*; Carlota Dobano, Gemma Moncunill and colleagues at Barcelona Institute for Global Health, Spain, for critical discussion of the results; the Italian Network for Neglected Tropical Diseases for promoting research, control and advocacy activities for the fight against NTDs.

Data availability

Data have been deposited in the Repository Figshare and are (DOI: 10.6084/m9.figshare.25311364).

REFERENCES


for diagnosis, monitoring and screening of Schistosoma infections utilizing lateral-flow based assays and upconverting phosphor labels. Parasitology. 2014


Schistosomiasis in non-endemic areas: Italian consensus recommendations for screening, diagnosis and management by the Italian Society of Tropical Medicine and Global Health (SIMET), endorsed by the Committee for the Study of Parasitology of the Italian Association of Clinical Microbiologists (CoSP-AMCLI), the Italian Society of Parasitology (SoIPa), the Italian Society of Gastroenterology and Digestive Endoscopy (SIGE), the Italian Society of Gynaecology and Obstetrics (SIGO), the Italian Society of Colposcopy and Cervico-Vaginal Pathology (SICPCV), the Italian Society of General Medicine and Primary Care (SIMG), the Italian Society of Infectious and Tropical Diseases (SIMIT), the Italian Society of Pediatrics (SIP), the
FIGURES

Figure 1. CAA and cfDNA positivity according to the breadth of the antibody response to *S. haematobium* antigens.

The figure shows the proportion of subjects showing results to CAA (blue) and cfDNA (green) testing according to the breadth of the antibody response to *S. haematobium* antigens. Bars represent proportion and whiskers represent the standard error of the proportion.
Figure 2. Proportion of cfDNA negative results according to CAA levels.

The figure shows the proportion of negative results according to CAA levels, i.e. concentration ranges. Bars represent proportion and whiskers represent the standard error of the proportion.
Figure 3. Cutoff levels of anti-SWAP and anti-SEA IgG and diagnostic performance compared to Circulating Anodic Antigen.

The figure shows the plot of sensitivity (blue) and specificity (red) compared to Circulating Anodic Antigen for each cutoff value of anti-SWAP IgG (left panel) and anti-SEA IgG (right panel). The grey vertical line indicates the analytical cutoff.
Figure 4. Distribution of anti-SWAP and anti-SEA IgG levels according to Circulating Anodic Antigen positivity.

The figure shows the boxplot distribution of anti-SWAP (left panel) and anti-SEA (right panel) IgG levels according to Circulating Anodic Antigen (CAA) positivity (negative subjects, Neg; positive subjects; Pos). The horizontal line represents the 50% percentile (median), the box lower and upper limits represent the 25% and 75% percentiles respectively, and lower and upper whiskers represent the 5% and 95% percentiles respectively, while the dots are outliers of the distribution.
Figure 5. Correlation of anti-SWAP and anti-SEA IgG levels with Circulating Anodic Antigen concentration.

The figure shows scatterplots illustrating the correlation of anti-SWAP IgG levels (log fOD; left panel) and anti-SEA IgG levels (log fOD; right panel) with CAA concentration (log pm/ml). Dots represent actual data points, the line represent the fitted data, and the grey area represent the 95% confidence interval of the fitted line.
Figure 6. Sensitivity and specificity of *S. haematobium* plasma biomarkers compared to Circulating Anodic Antigen, Composite Reference Standard and Latent Class Analysis.

The figure shows forest plots of sensitivity (blue) and specificity (red) of different *S. haematobium* plasma biomarkers compared to Circulating Anodic Antigen (CAA, left panel), Composite Reference Standard (CRS, mid panel) and Latent Class Analysis (LCA, right panel). Dots represent the estimate and whiskers represent the 95%CI of the estimate.