Check for updates

RESEARCH ARTICLE

Visual diagnosis of female genital schistosomiasis in Zambian women from hand-held colposcopy: agreement of expert image review [version 1; peer review: 3 approved, 3 approved with reservations]

Amy Sturt [ID][1,2], Henrietta Bristowe[3], Emily Webb [ID][4], Isaiah Hansingo [ID][5], Comfort Phiri[6], Maina Mudenda[5], Joyce Mapani[5], Tobias Mweene[6], Bruno Levecke [ID][7], Piet Cools[7,8], Govert van Dam[9], Paul Corstjens [ID][10], Helen Ayles[6,11], Richard Hayes[4], Suzanna Francis[4], Lisette van Lieshout[9], Bellington Vwalika[12], Eyrun Kjetland[13,14], Amaya Bustinduy [ID][11]

[1]Infectious Diseases, Department of Veterans Affairs Healthcare System, Palo Alto, CA, 94025, USA
[2]Geographic Medicine and Infectious Diseases, Stanford University School of Medicine, Palo Alto, USA
[3]Department of Microbiology and Infectious Diseases, Kings College Hopsital NHS Foundation Trust, London, UK
[4]MRC International Statistic and Epidemiology Group, London School of Hygiene & Tropical Medicine, London, UK
[5]Department of Obstetrics and Gynecology, Livingstone Central Hospital, Livingstone, Zambia
[6]Zambart, Lusaka, Zambia
[7]Department of Translational Physiology, Infectiology, and Public Health, Ghent University, Merelbeke, Belgium
[8]Department of Diagnostic Sciences, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium
[9]Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands
[10]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands
[11]Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, UK
[12]Department of Obstetrics and Gynecology, University of Lusaka, Lusaka, Zambia
[13]Department of Infectious Diseases, Oslo University Hospital, Oslo, Norway
[14]Discipline of Public Health Medicine, College of Health Sciences, University of KwaZulu-Natal, KwaZulu-Natal, South Africa

Open Peer Review

Approval Status ? ✓ ✓ ? ✓ ?

	1	2	3	4	5	6
version 1 10 Jan 2023	? view	✓ view	✓ view	? view	✓ view	? view

1. **Louise Arenholt**, Aalborg University, Aalborg, Denmark
   North Denmark Regional Hospital, Hjørring, Denmark

Abstract

Female genital schistosomiasis (FGS) can occur in *S. haematobium* infection and is caused by parasite egg deposition in the genital tract. Confirming a diagnosis of FGS is challenging due to the lack of a diagnostic reference standard. A 2010 expert-led consensus meeting proposed visual inspection of the cervicovaginal mucosa as an adequate reference standard for FGS diagnosis. The agreement of expert human reviewers for visual-FGS has not been previously described.

Methods:
In two Zambian communities, non-menstruating, non-pregnant,

sexually-active women aged 18-31 years participating in the HPTN 071 (PopART) Population-Cohort were enrolled in a cross-sectional study. Self-collected genital swabs and a urine specimen were collected at a home visit; trained midwives performed CVL and hand-held colposcopy at a clinic visit. *S. haematobium* eggs and circulating anodic antigen (CAA) were detected from urine. Two expert reviewers independently diagnosed visual-FGS as the presence of sandy patches, rubbery papules or abnormal blood vessels in digital cervicovaginal images obtained by hand-held colposcopy. PCR-FGS was defined as *Schistosoma* DNA detected by real-time PCR in any genital specimen (CVL or genital swab).

Results:
Of 527 women with cervicovaginal colposcopic images, 468/527 (88.8%) were deemed interpretable by Reviewer 1 and 417/527 (79.1%) by Reviewer 2. Visual-FGS was detected in 35.3% (165/468) of participants by expert review of colposcopic images by Reviewer 1 and in 63.6% (265/417) by Reviewer 2. Cohen's kappa statistic for agreement between the two expert reviewers was 0.16, corresponding to "slight" agreement. The reviewers made concordant diagnoses in 38.7% (204/527) participants (100 negative, 104 positive) and discordant diagnoses in 31.8% (168/527) participants.

Conclusions: The unexpectedly low level of correlation between expert reviewers highlights the imperfect nature of visual diagnosis for FGS based on cervicovaginal images obtained with a hand-held colposcope. This finding is a call to action for improved point-of-care diagnostics for female genital schistosomiasis

Keywords
Female genital schistosomiasis, Schistosoma haematobium, hand-held colposcopy

2. **Alexander B. Odaibo** (iD), University of Ibadan, Ibadan, Nigeria

3. **Daniela Fusco** (iD), Bernhard Nocht Insitute for Tropical Medicine, Hamburg, Germany

4. **Vanessa Christinet**, ASCRES, Lausanne, Switzerland

5. **Susannah Colt**, Rhode Island Hospital, Providence, USA

6. **Soledad Colombe** (iD), Institute of Tropical Medicine, Antwerp, Belgium

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Amy Sturt (amysturt@stanford.edu)

Author roles: Sturt A: Data Curation, Formal Analysis, Project Administration, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Bristowe H**: Formal Analysis, Writing – Original Draft Preparation, Writing – Review & Editing; **Webb E**: Data Curation, Formal Analysis, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Hansingo I**: Resources, Supervision, Writing – Review & Editing; **Phiri C**: Project Administration, Writing – Review & Editing; **Mudenda M**: Investigation, Writing – Review & Editing; **Mapani J**: Investigation, Writing – Review & Editing; **Mweene T**: Investigation, Writing – Review & Editing; **Levecke B**: Resources, Writing – Review & Editing; **Cools P**: Investigation, Resources, Writing – Review & Editing; **van Dam G**: Investigation, Writing – Review & Editing; **Corstjens P**: Investigation, Writing – Review & Editing; **Ayles H**: Resources, Writing – Review & Editing; **Hayes R**: Resources, Supervision, Writing – Review & Editing; **Francis S**: Supervision, Writing – Review & Editing; **van Lieshout L**: Investigation, Writing – Review & Editing; **Vwalika B**: Investigation, Writing – Review & Editing; **Kjetland E**: Investigation, Writing – Review & Editing; **Bustinduy A**: Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Sturt A, Bristowe H, Webb E *et al*. **Visual diagnosis of female genital schistosomiasis in Zambian women from hand-held colposcopy: agreement of expert image review [version 1; peer review: 3 approved, 3 approved with reservations]** Wellcome Open Research 2023, **8**:14 https://doi.org/10.12688/wellcomeopenres.18737.1

First published: 10 Jan 2023, **8**:14 https://doi.org/10.12688/wellcomeopenres.18737.1

Introduction

Female genital schistosomiasis (FGS), primarily caused by *S. haematobium* infection, is a neglected tropical disease associated with poverty, inadequate sanitation, and limited access to safe drinking water[1,2]. FGS occurs when schistosome eggs destined for excretion via the urinary bladder are deposited in the female genital tract. These tissue-deposited eggs can be associated with chronic inflammation and characteristic genital mucosal lesions[3]. Visual-FGS refers to the identification of these characteristic mucosal changes, such as sandy patches (grainy and homogeneous), rubbery papules, and abnormal blood vessels by visual inspection of the cervicovaginal mucosa[3]. The visual detection of FGS-associated lesions requires the insertion of a vaginal speculum, a good light source, and a lens providing at least 15x magnification. A standard colposcope has traditionally been used in research settings for visual-FGS diagnosis, but the bilharzia and HIV (BILHIV) study demonstrated recently that hand-held colposcopy could also be used to decentralize colposcopy services[4–6].

Confirming a diagnosis of FGS is challenging as there is not a widely accepted diagnostic reference standard for research, diagnosis, and screening[2]. A 2010 expert-led consensus meeting proposed visual inspection of the cervicovaginal mucosa as an adequate reference standard for FGS diagnosis[7]. However, the mucosal changes in visual-FGS are non-specific and have also been associated with sexually transmitted infections (STI), human papillomavirus (HPV) infection, and cervical precancer[3]. Diagnostic methods that are not adequately specific for FGS diagnosis may lead to over-treatment with praziquantel and may overlook the diagnosis and treatment of STIs and cervical cancer. Although there is little evidence of praziquantel resistance in humans[8], indiscriminate treatment may theoretically increase the risk of the development of praziquantel resistance[9]. Since cervicovaginal visualization is widely promoted[10] for FGS screening and diagnosis, we aimed to further evaluate the agreement of human expert reviewers for the diagnosis of visual-FGS.

Methods

Study setting and participants

The cross-sectional bilharzia and HIV (BILHIV) study[11] was nested in the HPTN 071 (PopART) cluster randomized trial[12]. As previously described, after the 36-month HPTN 071 (PopART) visit, community workers made home visits to women expressing interest in the BILHIV study[11]. Between January and August 2018, eligible women who were 18–31 years old, not pregnant, sexually active, and resident in one of two urban communities that participated in HPTN 071 (PopART) in Livingstone, Zambia were enrolled in the BILHIV study. The primary aim of the BILHIV study was to compare the performance of genital self-sampling (cervical and vaginal swabs) to clinic-based cervicovaginal lavage (CVL) for the detection of *Schistosoma* DNA by quantitative PCR (qPCR); results have been reported[11]. A specific pre-specified BILHIV study objective (the subject of the current manuscript) was to compare agreement of expert review of images obtained through hand-held colposcopy for the diagnosis of visual-FGS.

Home and clinic-based sample collection

As previously described, the home visit included written informed consent, a questionnaire, genital self-sampling (cervical and vaginal), and collection of a urine specimen[11]. Enrolled women who were not currently menstruating were then invited to attend Livingstone Central Hospital cervical cancer clinic, where midwives collected CVL. After speculum insertion, normal saline (10 mL) was flushed across the cervix and vaginal walls for one minute with a bulb syringe and CVL fluid was collected from the posterior fornices.

Hand-held colposcopy and image review

At the clinic, cervicovaginal images were captured with a portable colposcope (EVA System, MobileODT, Tel Aviv, Israel) and independently evaluated by two expert reviewers for any of the four recognized FGS cervicovaginal manifestations: grainy sandy patches, homogenous yellow sandy patches, rubbery papules, and abnormal blood vessels[13]. At their discretion, expert reviewers could exclude images that they felt could not be evaluated due to technical issues, image quality, or limited cervical visualization. If any of the four recognized FGS cervicovaginal manifestations was present, the participant was categorized as "visual-FGS". If none of the four cervicovaginal manifestations were present the participant was categorized as "visual-FGS not detected"[13]. The expert reviewers were both senior practicing physicians at the Professor level, who have training and expertise in standard colposcopy. Reviewer 1 (EFK) is an infectious diseases physician and Reviewer 2 (BV) is an obstetrician and gynecologist. Both reviewers have extensive practical and research-based expertise in evaluating and diagnosing FGS in endemic settings. Additionally, both reviewers contributed as authors of the 2015 WHO FGS Pocket Atlas[13]. Each reviewer was informed of the study setting and methods, but both were blinded to the study participants' FGS, *Schistosoma*, and STI status.

Women with at least one of the visual manifestations of FGS[3,13] or with any positive urine or genital *Schistosoma* diagnostic result were treated free-of-charge with 40 mg/kg praziquantel. Testing for STIs was not performed at the point-of-care and participants with suspected STIs were offered syndromic management, as per local guidelines[14]. In line with national and local clinic protocols adapted to real-world resource limitations, human papillomavirus (HPV) testing was not performed.

In parallel with BILHIV study procedures, participants could choose to engage in free cervical cancer screening using the visual inspection with acetic acid (VIA) technique. In the subset of women who engaged in cervical cancer screening, midwives applied 3-5% acetic acid to the cervix after CVL collection, as previously described[15]. An opaque white reaction was classified as positive and no change as negative[16].

Urine microscopy, and circulating anodic antigen

Up to 50mL of fresh urine was centrifuged and examined by microscopy for *S. haematobium* eggs. The participant was considered to have urinary schistosomiasis if a pellet contained at least one *S. haematobium* egg[11]. A lateral flow assay utilizing

up-converting reporter particles for the quantification of circulating anodic antigen (CAA) was performed on urine samples, as previously described[11,17]. Analyzing the equivalent of 417 µL urine (wet reagent, UCAAhT417), a test result indicating a CAA value >0.6 pg/mL was considered positive[18].

qPCR for detection of Schistosoma DNA
DNA extraction, amplification and detection of the Schistosoma-specific internal-transcribed-spacer-2 (ITS-2) target by real-time (qPCR) was performed at Leiden University Medical Center, as previously described, using 200 µL of CVL, cervical or vaginal swab fluid[11,19].

Other infections
Laboratory-based fourth-generation HIV-1 testing (Abbott Architect HIV Ag/Ab Combo Assay) was performed for HPTN 071 (PopART) Population Cohort participants at each study visit[12]. STIs were quantified among a subset of participants by qPCR using the S-DiaCTNG™ (for C. trachomatis and N. gonorrhea) and S-DiaMGTV™ kits (for M. genitalium and T. vaginalis) (Diagenode Diagnostics, Seraing, Belgium) on DNA from cervical swabs at Ghent University (Ghent, Belgium) according to the manufacturer's instructions.

Consent
The study was approved by the University of Zambia Biomedical Research Ethics Committee (011-08-17), the Zambia National Health Research Authority and the London School of Hygiene and Tropical Medicine Ethics Committee (14506). Permission to conduct the study was given by Livingstone District Health Office and the Livingstone Central Hospital superintendent.

Statistical Methods
The planned sample size of the BILHIV study was based on calculations related to the primary BILHIV study objective, as previously described[11]. Participant characteristics were summarized by median and interquartile range (IQR) for continuous variables, and by frequency and percentage for categorical variables. Participants missing data for a specific variable were excluded from analysis involving that variable. The primary analysis evaluated the agreement between the two expert reviewers using Cohen's kappa statistic. A secondary analysis evaluated the association between visual-FGS (exposure) and abdominal, genitourinary, and reproductive manifestations (outcomes). Crude associations were evaluated using chi-squared tests, and logistic regression was used to calculate crude and adjusted odds ratios (OR) for the association of visual-FGS with clinical manifestations; this was done separately for each expert reviewer's diagnosis of visual-FGS. In this study we employed various diagnostic tests to evaluate urinary Schistosoma infection (CAA and urine microscopy), and FGS (portable colposcopy, and Schistosoma DNA on CVL and genital swabs) as previously described[20–22]. Another secondary analysis evaluated each diagnostic method for its association with the presence of visual-FGS, separately for each expert reviewer. Due to small numbers, for evaluating the association of visual-FGS with

PCR-FGS, we used a composite definition of PCR-FGS or "any positive genital PCR", defined as any positive cervical or vaginal swab or CVL specimen. Chi-squared tests were used to assess crude associations, and logistic regression was used to calculate crude and age-adjusted odds ratio (OR) of the various Schistosoma and FGS diagnostics with the presence or absence of visual-FGS. For both secondary analyses, exact logistic regression was used for analyses where 5 or fewer participants in a particular exposure category had the outcome. Due to the exploratory nature of these analyses, we did not adjust for multiple comparisons. Data were analyzed using STATA 15.1 (Stata Corporation, College Station, TX).

Results
Baseline characteristics and demographics
The BILHIV study enrolled 603 eligible women, 527 (87.4%) of whom had cervicovaginal images captured by portable colposcopy. Of the 527 women with images, 468 (88.8%) were deemed interpretable by Reviewer 1 and 418 (79.3%) by Reviewer 2 (Figure 1). Each reviewer designated a proportion of images uninterpretable, leading to differences in denominators. The median age of the participants was 24 years (range 22 – 28) and 323 (61.3%) had attended some secondary school (Table 1). The majority of participants were married, had previously been pregnant, and had been sexually active within the last six months. There was no association between visual FGS, as identified by any expert review, and current or childhood water use.

Prevalence of visual-FGS and interrater agreement
Visual-FGS was detected in 35.3% (165/468) of participants by expert review of digital images from hand-held colposcopy by Reviewer 1 and in 63.6% (265/417) by Reviewer 2. The Cohen's kappa statistic for interrater agreement between the two expert reviewers was 0.16, corresponding to "slight" agreement (Table 2). The reviewers made concordant diagnoses in 38.7% (204/527) participants (100 concordant negative, 104 concordant positive) and discordant diagnoses in 31.8% (168/527) cases (reviewer 1 positive, reviewer 2 negative in 32; reviewer 2 positive and reviewer 1 negative in 136). Both reviewers agreed 14 images were unevaluable. A further 26.7% (141/527) images were discordant in evaluability by the expert reviewers (Reviewer 1, n=45; Reviewer 2, n=96).

Visual FGS and Schistosoma laboratory tests
Of the 527 participants, 6.1% (32/527) had urinary S. haematobium infection, as diagnosed by urine microscopy, and 14.9% (78/525) had a detectable urine CAA. There was no association between S. haematobium egg-positive urine microscopy or urine CAA and visual-FGS, as defined by Reviewer 1 or Reviewer 2's assessmen (Table 3).

PCR-FGS, defined as any positive Schistosoma qPCR from a genital sample, was diagnosed in 5.0% (30/603) of participants [3.4% (18/527) cervical swab, 2.7% (14/527) vaginal swab, and 2.7% (14/527) CVL]. There was no association between visual-FGS and PCR-FGS, 10 of the 165 women (6.0%)
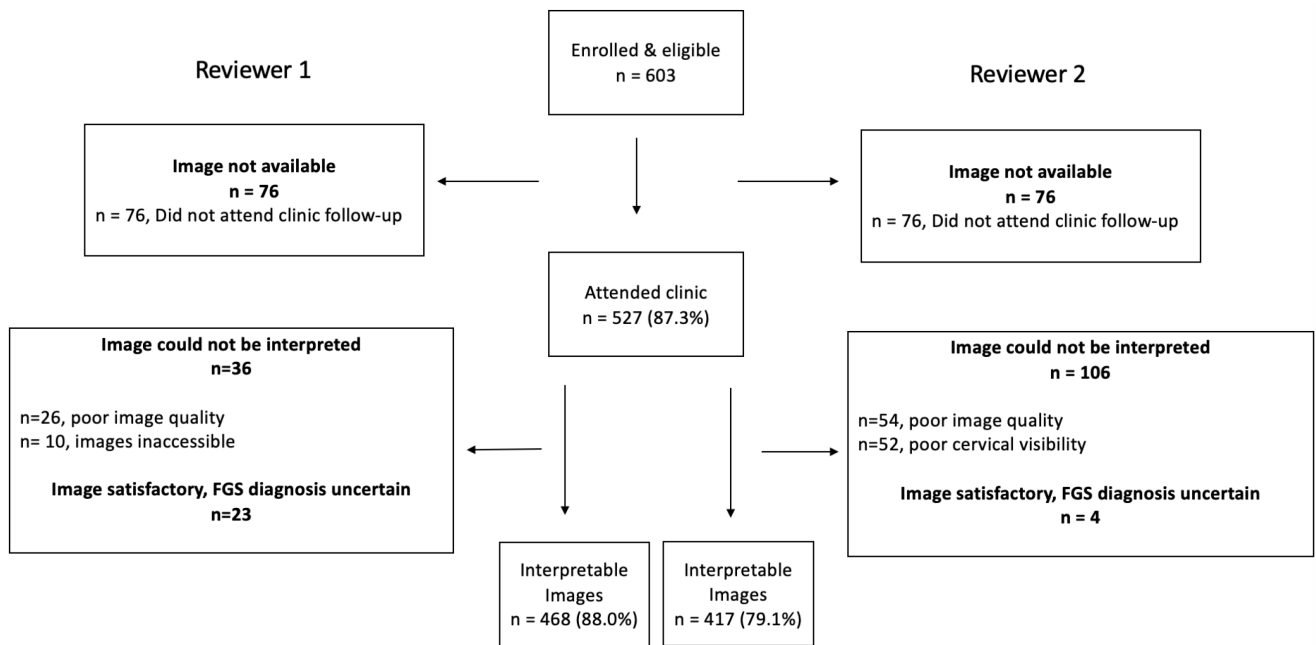
Figure 1. Flowsheet of cervicovaginal image review after hand-held colposcopy in BILHIV study participants.

Table 1. Demographics and reproductive health characteristics of the BILHIV study participants who underwent portable colposcopy (n=527).

Participant Characteristics		Study population (n=527)
Age in years – Median (IQR)		24 (22 – 28)
Marital status	Single	213 (40.4)
	Married or cohabitating	292 (55.4)
	Divorced or separated	22 (4.2)
Education (highest level)	None or any primary school	155 (29.4)
	Any secondary school	323 (61.3)
	Trade, degree or higher	49 (9.3)
District	Community A	290 (55.0)
	Community B	237 (45.0)
Household members	1-3	162 (30.7)
	4-5	210 (39.9)
	6+	155 (29.4)
Employment status	Unemployed	363 (68.9)
	Employed	164 (31.1)
Current water contact	None	447 (84.8)
	Any	80 (15.2)

Participant Characteristics		Study population (n=527)
Childhood water contact	None	151 (28.7)
	Any	376 (71.3)
Reproductive Health Characteristics		
Age at sexual debut (years)*	8-16	221 (42.0)
	17-19	228 (43.4)
	20-24	77 (14.6)
Lifetime sexual partners	1	145 (27.5)
	2	134 (25.4)
	3	108 (20.5)
	4+	140 (26.6)
Prior pregnancy*	No	74 (14.1)
	Yes	452 (85.9)
Currently sexually active**	No	65 (12.4)
	Yes	460 (87.6)
Condom use with last sex†	No	381 (73.7)
	Yes	136 (26.3)
HIV-1 infection††	Not detected	407 (77.8)
	Detected	116 (22.2)
Any STI ~	Not detected	138 (63.8)
	Detected	73 (34.6)
Any Hormonal Contraception~~	No	201 (38.1)
	Yes	326 (61.9)
VIA‡	Negative	213 (89.9)
	Positive	24 (10.1)

*Participants who responded with "no answer" (n=1) are not shown in the table

**Defined as any sexual activity in the last 6 months; Participants who responded with "no answer" (n=2) are not shown in the table

†Participants who responded with "no answer" (n=10) are not shown in the table

††Participants with missing data (n=4) are not shown in the table

~STI were evaluated in a sub-set of women from the BILHIV study, missing values (n=316) are not shown in the table)

~~Any hormonal contraception is defined as use of injectable agents, implants, or oral contraceptive pills

‡VIA results were not collected in the BILHIV study and were not available for all participants, participants with missing data (n=366) are not shown in the table

identified by Reviewer 1 as having visual-FGS had PCR-FGS and 17 of the 265 women (6.4%) identified by Reviewer 2 as having visual-FGS had PCR-FGS , compared to 4.9% and 3.9% among women identified by Reviewer 1 and 2 as not having visual-FGS respectively (Table 3).

Symptoms

The association between abdominal, genitourinary, and reproductive manifestations and visual-FGS is shown in Table 4. Neither vaginal discharge, vaginal bleeding after sex, the presence of external genital sores, dysuria nor abdominal pain was

Table 2. Agreement of expert reviewers for the presence or absence of visual-FGS.

Reviewer 2	Reviewer 1	FGS not detected	Visual FGS	Not evaluable	Total	Total evaluable	Cohen's Kappa (+ SE)*·†	Interpretation
Visual FGS not detected		100 (19.0)	32 (6.1)	20 (3.8)	152 (28.8)	152 (36.5)		
Visual FGS		136 (25.8)	104 (19.7)	25 (4.7)	265 (50.3)	265 (63.5)	0.16 (0.04)	None to slight
Not evaluable		67 (12.7)	29 (5.5)	14 (2.7)	110 (20.9)			
Total		303 (57.5)	165 (31.3)	59 (11.2)	527 (100.0)	417 (100)		
Total evaluable		303 (64.7)	165 (35.3)		468 (100)			

FGS – female genital schistosomiasis, SE – standard error

*total evaluable for Reviewer 1: visual FGS not detected 303/468 (64.7), visual FGS detected 165/468 (35.3); total evaluable for Reviewer 2: visual FGS not detected 152/417 (36.5), visual FGS detected 265/417 (63.5);

†Cohen's kappa is restricted to those participants where both Reviewer 1 and Reviewer 2 provided a diagnosis

Table 3. Associations of *Schistosoma* diagnostics with visual-FGS.

Participant Characteristics		Visual FGS not detected	Visual FGS detected	Crude OR	P-value*	Adjusted OR**	P-value†
***Schistosoma* diagnostics (Reviewer 1)**		n=303	n=165				
Eggs on urine microscopy	Not detected	287 (94.7)	151 (91.5)	reference	0.3	reference	0.2
	Detected	16 (5.3)	14 (8.5)	1.66 (0.73 – 3.74)		1.68 (0.74 – 3.81)	
Circulating anodic antigen~	Not detected	261 (86.1)	131 (80.4)	reference	0.1	reference	0.1
	Detected	42 (13.9)	32 (19.6)	1.52 (0.92 – 2.52)		1.49 (0.91 – 2.51)	
Any positive genital PCR	DNA not detected	288 (95.1)	155 (93.9)	reference	0.8	reference	0.7
	DNA detected	15 (4.9)	10 (6.0)	1.24 (0.54 – 2.82)		1.27 (0.55 – 2.94)	
***Schistosoma* Diagnostics (Reviewer 2)**		n=152	n=265				
Eggs on urine microscopy††	Not detected	147 (96.7)	243 (91.7)	reference	0.06	reference	0.07
	Detected	5 (3.3)	22 (8.3)	2.65 (0.95 – 9.17)		2.65 (0.95 – 9.19)	
Circulating anodic antigen~	Not detected	131 (86.7)	215 (81.4)	reference	0.2	reference	0.2
	Detected	20 (13.3)	49 (18.6)	1.49 (0.85 – 2.62)		1.49 (0.85 – 2.63)	
Any positive genital PCR	DNA not detected	146 (96.1)	248 (93.6)	reference	0.3	reference	0.3
	DNA detected	6 (3.9)	17 (6.4)	1.69 (0.64 – 4.33)		1.67 (0.64 – 4.36)	

*Chi squared p-value unless otherwise noted

**Adjusted for age

†Likelihood ratio test p-value

Missing values not included in the table: ~ (n=2)

††Odds Ratios and p-values obtained through exact logistic regression in both crude and adjusted analyses

Table 4. Associations of genitourinary and abdominal symptoms with visual-FGS.

Participant Characteristics		Visual FGS not detected	Visual FGS detected	P-value*	Crude OR	P-value	Adjusted OR**	P-value †
Signs & Symptoms (Reviewer 1)		n= 303	n=165					
Vaginal discharge	Not present	273 (90.1)	139 (84.2)	0.06	reference	0.06	reference	0.1
	Present	30 (9.9)	26 (15.8)		1.70 (0.97 – 2.99)		1.58 (0.90 – 2.80)	
Dyspareunia	Not present	239 (78.9)	139 (84.2)	0.2	reference	0.2	reference	0.1
	Present	64 (21.1)	26 (15.8)		0.70 (0.42 – 1.15)		0.68 (0.41 – 1.12)	
Vaginal bleeding after sex	Not present	289 (95.4)	153 (92.7)	0.2	reference	0.2	reference	0.3
	Present	14 (4.6)	12 (7.3)		1.62 (0.73 – 3.59)		1.48 (0.66 – 3.31)	
Vaginal sores	Not present	280 (92.4)	154 (93.3)	0.7	reference	0.7	reference	0.6
	Present	23 (7.6)	11 (6.7)		0.87 (0.41 – 1.83)		0.84 (0.40 – 1.77)	
Dysuria	Not present	259 (85.5)	140 (84.9)	0.9	reference	0.9	reference	0.8
	Present	44 (14.5)	25 (15.1)		1.05 (0.62 – 1.79)		0.92 (0.53 – 1.59)	
Hematuria	Not present	295 (97.4)	155 (93.9)	0.07	reference	0.07	reference	0.1
	Present	8 (2.6)	10 (6.1)		2.38 (0.92 – 6.20)		2.18 (0.84 – 5.67)	
Abdominal pain	Not present	217 (71.6)	123 (74.5)	0.5	reference	0.5	reference	0.4
	Present	86 (28.4)	42 (25.5)		0.86 (0.56 – 1.33)		0.82 (0.53 – 1.26)	
Delay in conception††	No	235 (94.4)	111 (86.7)	0.01	reference	0.01	reference	<0.01
	Yes	14 (5.6)	17 (13.3)		2.57 (1.22 – 5.40)		2.74 (1.29 – 5.83)	
Signs & Symptoms (Reviewer 2)		n=152	n=265					
Vaginal discharge	Not present	133 (87.5)	236 (89.1)	0.6	reference	0.6	reference	0.6
	Present	19 (12.5)	29 (10.9)		0.86 (0.46 – 1.60)		0.84 (0.45 – 1.57)	
Dyspareunia	Not present	131 (86.2)	208 (78.5)	0.05	reference	0.05	reference	0.05
	Present	21 (13.8)	57 (21.5)		1.71 (0.99 – 2.95)		1.71 (0.99 – 2.95)	
Vaginal bleeding after sex	Not present	144 (94.7)	251 (94.7)	1.0	reference	1.0	reference	1.0
	Present	8 (5.3)	14 (5.3)		1.00 (0.41 – 2.45)		0.99 (0.40 – 2.42)	
Vaginal sores	Not present	146 (96.1)	244 (92.1)	0.1	reference	0.1	reference	0.1
	Present	6 (3.9)	21 (7.9)		2.09 (0.83 – 5.31)		2.10 (0.83 – 5.31)	
Dysuria	Not present	132 (86.8)	231 (87.2)	0.9	reference	0.9	reference	0.9
	Present	20 (13.2)	34 (12.8)		0.97 (0.54 – 1.76)		0.95 (0.52 – 1.75)	
Hematuria†††	Not present	150 (98.7)	250 (94.3)	0.03	reference	0.05	reference	0.05
	Present	2 (1.3)	15 (5.7)		4.49 (1.02 – 40.99)		4.44 (1.00 – 40.63)	
Abdominal pain	Not present	108 (71.0)	192 (72.5)	0.8	reference	0.8	reference	0.7
	Present	44 (28.9)	73 (27.5)		0.93 (0.60 – 1.45)		0.92 (0.59 – 1.44)	
Delay in conception‡	No	112 (91.8)	193 (91.9)	1.0	reference	1.0	reference	1.0
	Yes	10 (8.2)	17 (8.1)		0.99 (0.44 – 2.22)		1.01 (0.45 – 2.92)	

*Chi squared p-value

**Adjusted for age and district of residence

†Likelihood ratio test p-value

††(Reviewer 1) Declined to answer (n=33) and not applicable (n=95) are not included in the table (54 visual FGS not detected; 37 FGS; 37 missing)

†††Odds Ratios and p-values obtained through exact logistic regression in both crude and adjusted analyses

‡(Reviewer 2) Declined to answer (n=33) and missing (n=95) are not included in the table; (30 visual FGS not detected; 55 FGS detected; 43 missing)

associated with the presence of visual FGS as diagnosed by either expert reviewer (Table 4). Self-reported delay in conception was associated with the presence of visual-FGS, as assessed by Reviewer 1, both in crude analysis and after adjusting for age and community of residence (aOR 2.74 [1.29 – 5.83], p<0.01). Visual-FGS, as defined by Reviewer 2's assessment of hand-held colposcopy images, was associated with hematuria (aOR 4.44 [1.00 – 40.63] p=0.05) and dyspareunia (aOR 1.71 [0.99 – 2.95], p=0.05), albeit with weak evidence of an association (Table 4).

Discussion

Diagnostics for neglected tropical diseases should be accurate, accessible, and affordable, with specimen collection that is easy[23]. Making a diagnosis of FGS is challenging as there is currently not a widely accessible, sensitive and non-invasive reference standard for either diagnosis or screening which confirms *Schistosoma* genital involvement at the point-of-care. In a 2010 expert-led consensus meeting, visual imaging of the vagina and cervix with photocolposcopic methods was proposed as an adequate reference standard for FGS visual diagnosis[24]. Imaging is currently the only widely available point-of-care diagnostic tool for FGS diagnosis outside of the research setting and the BILHIV study sought to use hand-held colposcopy to enable community-based FGS diagnosis[11]. Visual imaging can be useful in the assessment of *Schistosoma*-related morbidity, praziquantel treatment response, and defining the natural history of visual-FGS. Additionally, hand-held and traditional colposcopy have the logistical advantage that they can be integrated with existing cervical cancer screening programmes[25]. However, visual imaging has important limitations. Firstly, interpretation of visual imaging is subjective. Secondly, visual imaging lacks specificity as the characteristic sandy patches can also be associated with STI and the abnormal blood vessels can also be associated with cervical precancer[3]. This study shows "slight" agreement between senior, highly experienced expert reviewers, highlighting the imperfect nature of human expert review of images obtained with hand-held colposcopy for FGS.

Visual FGS-diagnosis is a widely accepted diagnostic tool for evaluating *Schistosoma*-associated genital morbidity. However, visual-FGS screening is often centralized in settings with access to traditional colposcopy and is invasive, requiring vaginal speculum insertion and trained medical professionals (physicians, nurses, or midwives) to visualize the cervix and vagina at high resolution[11]. Additionally, visual-FGS diagnosis requires a full inspection of the mucosal surfaces of the vagina and cervix. If metal specula are used, post-examination autoclaving and appropriate disinfection further constrains the settings in which this diagnostic strategy can be seamlessly implemented. Disposable specula have risks and benefits. While hygienic and convenient, disposable plastic specula may not be sturdy enough when rotated to inspect the anterior and posterior vaginal walls and may contribute to missed visual-FGS diagnoses[4]. A good light source is needed for optimal cervicovaginal visualization[4], as well as a device which can provide at least 15x magnification, ideally a colposcope,

hand-held colposcope, or digital camera[6]. Thus, colposcopy, whether hand-held or traditional, for visual-FGS diagnosis is not readily scalable for use as a population-based screening technique.

In this current work, without complete STI and HPV testing or cervicovaginal biopsy on each participant, it is challenging to assess the significance of the sandy patches and abnormal blood vessels identified by the clinical expert reviewers. Notably, researchers in Tanzania performed macroscopic cervicovaginal examinations comparing *S. haematobium* endemic and non-endemic areas, finding 75% of participants in endemic areas had cervical lesions (including sandy patches, edema, erosions and petechiae) compared with 36% of women in non-endemic areas[26]. The Tanzanian study illustrates the limited specificity of visual techniques, since one-third of the women had cervical lesions in communities where *S. haematobium* is not endemic.

Other diagnostic approaches such as PCR-based methods, have been implemented in research settings but are not yet field-deployable[11]. Antigen, antibody, and pathogen-based diagnostics (such as microscopy) are useful diagnostic adjuncts for *Schistosoma* infection, but do not confirm the involvement of genital tissue. Future diagnostic algorithms may be optimized by first performing a microbiologic *S. haematobium* diagnosis prior to performing screening for genital involvement[11]. Promising pathogen detection strategies that can be implemented at the point-of-care include isothermal DNA amplification methods. These field-deployable molecular assays should be further developed for use at the point-of-care to identify *Schistosoma* DNA in self-collected genital swabs[27].

Our study did not show a consistent association between expert diagnosis of visual-FGS and abdominal, genitourinary and reproductive symptoms. Reviewer 1's evaluation suggested an association between self-reported delay in conception and visual-FGS and Reviewer 2's evaluation suggested a weak association with hematuria and dyspareunia in participants with visual-FGS. A retrospective study from Tanzania evaluating histopathology reported tubal schistosomiasis in 4 patients reporting with infertility[28] and a cross-sectional study from Zimbabwe found strong evidence that the presence of *S. haematobium* in pap smear was associated with infertility in women aged 20 – 49 years, after adjusting for age and HIV status[29]. While alluring to consider the association of delayed conception identified by one reviewer with visual-FGS in isolation, the association would have been strengthened by consistency of the findings across reviewers. Additionally, in interpreting this result, it is important to consider the possibility of a type 1 error when large numbers of statistical tests are performed.

Previous work on visual-FGS has compared visual imaging to other diagnostic standards[30] or have used computerized algorithms[31,32], or a combination of human reviewers and a digital gridding technique to evaluate visual-FGS[5]. A recent Madagascan study utilized human reviewers together with a

digital image gridding technique to review images of women with known FGS-associated clinical lesions and found a Fleiss kappa of 0.55 ("moderate" agreement) for detecting rubbery papules. Reviewers in that study achieved a higher agreement than that described in our study, potentially by undergoing an initial consensus rating exercise to reach agreement on uniform rating of images. However, it is notable that in the Madagascan study, all images were thought to contain FGS lesions, removing the burden of uncertainty and highlighting that in settings where images are known to contain FGS lesions, agreement between reviewers was at best "moderate"[5]. Our approach in the BILHIV study illustrates a real-world scenario where expert reviewers may not necessarily have the opportunity for consensus agreement prior to consultation. This is the first study to assess the agreement of human expert reviewers for diagnosing visual-FGS with hand-held colposcopy, where both reviewers were blinded to the participants' FGS and *Schistosoma* diagnostic status. In this study, both expert reviewers are experienced clinical Professors who have expertise in diagnosing FGS in endemic settings and contributed as authors to the 2015 WHO FGS Pocket Atlas[13].

While our approach is unique, this work has some limitations. The prevalence of urinary schistosomiasis and also PCR-FGS were low, thus limiting precision in effect sizes and power to detect association when comparing PCR-FGS and urinary schistosomasis with visual-FGS. Additionally, the urban setting, relatively narrow age range of the participants and low urinary *S. haematobium* prevalence may limit generalizability. Future additional work in a setting with higher schistosomiasis prevalence would be needed to definitively exclude an association between symptoms, standard *Schistosoma,* and FGS diagnostics and visual-FGS. Additionally, future work could incorporate artificial intelligence, such as computer algorithms to detect the characteristic color change caused by involvement of the genital mucosa with FGS[32] or the use of digital gridding techniques[5]. Additionally, a initial consensus rating exercise could be incorporated into future work with human expert review for FGS-associated lesions. The presence or absence of the specific FGS lesion (sandy patch, rubbery papule, abnormal blood vessels) was not consistently documented along with the presence or absence of visual-FGS, limiting analysis by lesion type. Study participants self-reported their time-to-conception status, thus results may be subject to recall bias. STI testing was only performed on a subset of the study population and visual inspection with acetic acid data were not obtained within the BILHIV study, thus data on these variables are incomplete[15]. Without complete STI and HPV testing or cervicovaginal biopsy on each participant, it is challenging to assess the significance of the sandy patches and abnormal blood vessels identified by the clinical expert reviewers. Thus, we cannot exclude residual or unmeasured confounding.

In conclusion, with only "slight" agreement between experienced expert reviewers who identified visual-FGS from digital images obtained during point-of-care colposcopy, we suggest caution when visual imaging is used as a stand-alone FGS diagnostic. Our findings highlight the imperfect nature and challenges of visual diagnosis of FGS as a research and clinical endpoint and is a call to action for improved point-of-care diagnostics and diagnostic pathways for female genital schistosomiasis.

Data availability

Underlying data
LSHTM Data Compass. Data for: Expert review of images from hand-held colposcopy in Zambian women with genital schistosomiasis. DOI: https://doi.org/10.17037/DATA.00003182[33].

- Female genital schistosomiasis (FGS) can occur in the setting of urinary S. haematobium infection, a neglected tropical disease associated with poverty, inadequate sanitation, and limited access to safe drinking water. Confirming a diagnosis of FGS is challenging as there is not a widely accepted diagnostic reference standard for research, diagnosis, and screening. A 2010 expert-led consensus meeting proposed visual inspection of the cervicovaginal mucosa as an adequate reference standard for FGS diagnosis. However, the mucosal changes in visual-FGS are non-specific and have also been associated with bacterial STI, human papillomavirus (HPV) infection, and cervical pre-cancer. Since cervicovaginal visualization is widely promoted for FGS screening and diagnosis, we wished to further evaluate the inter-rater reliability, correlation, and agreement of human expert reviewers in visual-FGS.

- This data is under restricted access due to the assurance given to participants that responses would be kept completely confidential. This is particularly important due to the sensitivity of the data produced. The data set can be accessed by completing the Request Form, which requires that the intended use for the data is specified. Data available under the LSHTM Data Compass Data Sharing Agreement.

Author contributions
Amy S. Sturt – data curation, formal analysis, BILHIV project administration, visualization, original manuscript preparation, manuscript editing and revision

Henrietta Bristowe – formal analysis, original manuscript preparation, manuscript editing and revision

Emily L. Webb – data curation, formal analysis, supervision, original manuscript preparation, manuscript editing and revision

Isaiah Hansingo – resources, supervision, writing – review and editing

Comfort R. Phiri – BILHIV project administration, writing – review and editing

Maina Mudenda – investigation, writing – review and editing

References

1. WHO: **Schistosomiasis**. Geneva: World Health Organization; 2022; [accessed November 10, 2022].
   Reference Source

2. Bustinduy AL, Randriansolo B, Sturt AS, *et al.*: **An update on female and male genital schistosomiasis and a call to integrate efforts to escalate diagnosis, treatment and awareness in endemic and non-endemic settings: The time is now.** *Adv Parasitol.* 2022; **115**: 1–44.
   PubMed Abstract | **Publisher Full Text**

3. Kjetland EF, Ndhlovu PD, Mduluza T, *et al.*: **Simple clinical manifestations of genital Schistosoma haematobium infection in rural Zimbabwean women.** *Am J Trop Med Hyg.* 2005; **72**(3): 311–9.
   PubMed Abstract | **Publisher Full Text**

4. Norseth HM, Ndhlovu PD, Kleppa E, *et al.*: **The colposcopic atlas of schistosomiasis in the lower female genital tract based on studies in Malawi, Zimbabwe, Madagascar and South Africa.** *PLoS Negl Trop Dis.* 2014; **8**(11): e3229.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Arenholt LTS, Aaroe KK, Norderud K, *et al.*: **Cervical lesion proportion measure using a digital gridded imaging technique to assess cervical pathology in women with genital schistosomiasis.** *PLoS Negl Trop Dis.* 2022; **16**(7): e0009995.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Søfteland S, Sebitloane MH, Taylor M, *et al.*: **A systematic review of handheld tools in lieu of colposcopy for cervical neoplasia and female genital schistosomiasis.** *Int J Gynaecol Obstet.* 2021; **153**(2): 190–9.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Kjetland EF, Leutscher PD, Ndhlovu PD: **A review of female genital schistosomiasis.** *Trends Parasitol.* 2012; **28**(2): 58–65.
   PubMed Abstract | **Publisher Full Text**

8. Tushabe JV, Lubyayi L, Sserubanja J, *et al.*: **Does Intensive Treatment Select for Praziquantel Resistance in High-Transmission Settings? Parasitological Trends and Treatment Efficacy Within a Cluster-Randomized Trial.** *Open Forum Infect Dis.* 2020; **7**(4): ofaa091.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Summers S, Bhattacharyya T, Allan F, *et al.*: **A review of the genetic determinants of praziquantel resistance in *Schistosoma mansoni*: Is praziquantel and intestinal schistosomiasis a perfect match?** *Front Trop Dis.* 2022; **3**: 933097.
   Publisher Full Text

10. Wells N, Chappuis F, Beran D: **Spotlight on experiences of medicine unavailability: access to medicines challenges for NCDs and NTDs - the contrasting cases of insulin and praziquantel.** *Expert Rev Clin Pharmacol.* 2020; **13**(4): 341–53.
    PubMed Abstract | **Publisher Full Text**

11. Sturt AS, Webb EL, Phiri CR, *et al.*: **Genital self-sampling compared with cervicovaginal lavage for the diagnosis of female genital schistosomiasis in Zambian women: The BILHIV study.** *PLoS Negl Trop Dis.* 2020; **14**(7): e0008337.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Hayes RJ, Donnell D, Floyd S, *et al.*: **Effect of Universal Testing and Treatment on HIV Incidence - HPTN 071 (PopART).** *N Engl J Med.* 2019; **381**(3): 207–18.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. World Health Organization: **Female genital schistosomiasis: a pocket atlas for clinical health-care professionals.** World Health Organization, 2015; Accessed April 30, 2021.
    Reference Source

14. Zambian Ministry of Health: **Guidelines for the Etiological and Clinical Management of Sexually Transmitted Infections in Zambia.** 2017.

15. Rafferty H, Sturt AS, Phiri CR, *et al.*: **Association between cervical dysplasia and female genital schistosomiasis diagnosed by genital PCR in Zambian women.** *BMC Infect Dis.* 2021; **21**(1): 691.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Sankaranarayanan R, Wesley R, Thara S, *et al.*: **Test characteristics of visual inspection with 4% acetic acid (VIA) and Lugol's iodine (VILI) in cervical cancer screening in Kerala, India.** *Int J Cancer.* 2003; **106**(3): 404–8.
    PubMed Abstract | **Publisher Full Text**

17. Corstjens PL, De Dood CJ, Kornelis D, *et al.*: **Tools for diagnosis, monitoring and screening of *Schistosoma* infections utilizing lateral-flow based assays and upconverting phosphor labels.** *Parasitology.* 2014; **141**(14): 1841–55.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Corstjens P, de Dood CJ, Knopp S, *et al.*: **Circulating Anodic Antigen (CAA): A Highly Sensitive Diagnostic Biomarker to Detect Active *Schistosoma* Infections-Improvement and Use during SCORE.** *Am J Trop Med Hyg.* 2020; **103**(1_Suppl): 50–7.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Obeng BB, Aryeetey YA, de Dood CJ, *et al.*: **Application of a circulating-cathodic-antigen (CCA) strip test and real-time PCR, in comparison with microscopy, for the detection of *Schistosoma haematobium* in urine samples from Ghana.** *Ann Trop Med Parasitol.* 2008; **102**(7): 625–33.
    PubMed Abstract | **Publisher Full Text**

20. Sturt AS, Webb EL, Patterson C, *et al.*: **Cervicovaginal Immune Activation in Zambian Women With Female Genital Schistosomiasis.** *Front Immunol.* 2021; **12**(181): 620657.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Sturt AS, Webb EL, Himschoot L, *et al.*: **Association of Female Genital Schistosomiasis With the Cervicovaginal Microbiota and Sexually Transmitted Infections in Zambian Women.** *Open Forum Infect Dis.* 2021; **8**(9): ofab438.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Sturt AS, Webb EL, Phiri CR, *et al.*: **Female Genital Schistosomiasis and HIV-1 Incidence in Zambian Women: A Retrospective Cohort Study.** *Open Forum Infect Dis.* 2021; **8**(7): ofab349.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Land KJ, Boeras DI, Chen XS, *et al.*: **REASSURED diagnostics to inform disease control strategies, strengthen health systems and improve patient outcomes.** *Nat Microbiol.* 2019; **4**(1): 46–54.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

24. Kjetland EF, Norseth HM, Taylor M, *et al.*: **Classification of the lesions observed in female genital schistosomiasis.** *Int J Gynaecol Obstet.* 2014; **127**(3): 227–8.
PubMed Abstract | **Publisher Full Text**

25. Engels D, Hotez PJ, Ducker C, *et al.*: **Integration of prevention and control measures for female genital schistosomiasis, HIV and cervical cancer.** *Bull World Health Organ.* 2020; **98**(9): 615–24.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Poggensee G, Kiwelu I, Weger V, *et al.*: **Female genital schistosomiasis of the lower genital tract: prevalence and disease-associated morbidity in northern Tanzania.** *J Infect Dis.* 2000; **181**(3): 1210–3.
PubMed Abstract | **Publisher Full Text**

27. Archer J, Patwary FK, Sturt AS, *et al.*: **Validation of the isothermal *Schistosoma haematobium* Recombinase Polymerase Amplification (RPA) assay, coupled with simplified sample preparation, for diagnosing female genital schistosomiasis using cervicovaginal lavage and vaginal self-swab samples.** *PLoS Negl Trop Dis.* 2022; **16**(3): e0010276.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Swai B, Poggensee G, Mtweve S, *et al.*: **Female genital schistosomiasis as an evidence of a neglected cause for reproductive ill-health: a retrospective histopathological study from Tanzania.** *BMC Infect Dis.* 2006; **6**: 134.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

29. Kjetland EF, Kurewa EN, Mduluza T, *et al.*: **The first community-based report on the effect of genital *Schistosoma haematobium* infection on female fertility.** *Fertil Steril.* 2010; **94**(4): 1551–3.
PubMed Abstract | **Publisher Full Text**

30. Galappaththi-Arachchige HN, Holmen S, Koukounari A, *et al.*: **Evaluating diagnostic indicators of urogenital *Schistosoma haematobium* infection in young women: A cross sectional study in rural South Africa.** *PLoS One.* 2018; **13**(2): e0191459.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

31. Holmen SD, Kleppa E, Lillebø K, *et al.*: **The first step toward diagnosing female genital schistosomiasis by computer image analysis.** *Am J Trop Med Hyg.* 2015; **93**(1): 80–6.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

32. Holmen SD, Kjetland EF, Taylor M, *et al.*: **Colourimetric image analysis as a diagnostic tool in female genital schistosomiasis.** *Med Eng Phys.* 2015; **37**(3): 309–14.
PubMed Abstract | **Publisher Full Text**

33. Sturt A, Webb E, Bristowe H, *et al.*: **Data for: Expert review of images from hand-held colposcopy in Zambian women with genital schistosomiasis.** [Data Collection]. London School of Hygiene & Tropical Medicine, London, United Kingdom. 2022.
http://www.doi.org/10.17037/DATA.00003182

Open Peer Review

Current Peer Review Status: ❓ ✔ ✔ ❓ ✔ ❓

Version 1

Reviewer Report 21 February 2023

❓   **Soledad Colombe** 🆔

Department of Public Health, Institute of Tropical Medicine, Antwerp, Belgium

I thank the editor and authors for the opportunity to review this article, which brings to light relevant information for those working on female genital schistosomiasis. In this study, the authors tested the validity of the recommended visual diagnosis of FGS by assessing the concordance in visual diagnosis from two independent senior physicians. They found very low agreement between the independent reviewers, highlighting the complexity of using visual imaging for diagnosis of female genital schistosomiasis. The article was overall very sound and well written. I have made a few comments below on both format and scientific content, which I hope will be useful to the authors.

Comments on format:
- Please define CVL in the abstract before using the accronym.

- I would quickly define in the abstract what is meant by "expert reviewers", and would also define it sooner in the methods in the main text.

Comments on scientific content:
I am missing more in-depth description and discussion of the discordance in visual diagnosis between the two expert reviewers.
- a) I noticed that the discordance was already quite large for the number of images that could not be interpreted (36 vs 106). What does "images inaccessible" in Figure 1 mean? Could you explain this discordance?

- b) Regarding the discordance in visual diagnosis of FGS: I understand that the presence or absence of a specific FGS lesion was not systematically recorded but did you notice a systematicity in the discordance? Would this discordance have been solved in the majority of cases after discussions between the two expert reviewers?

- c) Does the conclusion of "diagnostic uncertain" (n=23) of reviewer 1 overlap with the images discarded by reviewer 2 for "poor cervical visibility"?

     ◦ d) I would suggest to add pictures, if allowed by the journal, of images leading to concordant positive, concordant negative and discordant diagnoses.

It seems that in a number of participants, visual-FGS was "positive" whereas PCR/CAA/Microscopy were negative. In addition to the fact that the lesions thought during visual inspection can be non-specific, could it also be that visual-FGS in some cases is rather the sign of past infection? And in that case can the authors discuss the added-value of visual-FGS diagnostic?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious Diseases ; Epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 21 February 2023

https://doi.org/10.21956/wellcomeopenres.20776.r54091

✓   **Susannah Colt**
    Center for International Health Research, Rhode Island Hospital, Providence, RI, USA

<u>Summary of the article:</u>
Female genital schistosomiasis (FGS) is an underreported and debilitating illness affecting women and girls and is widespread among *Schistosoma haematobium*-endemic regions throughout Africa. There are significant challenges to accurate FGS diagnosis relating to lack of a consensus method, equipment availability, technical expertise, and stigma surrounding women's reproductive health. While the proposed reference-standard for FGS diagnosis is visual inspection by colposcopy (visual-FGS), this method is subjective to expert interpretation and can be further confounded by some sexually transmitted infections. The authors report secondary findings of the cross-sectional bilharzia and HIV (BILHIV) study, in which, N=527 women from Zambia were enrolled; provided urine, cervical and vaginal self-sampled tissue swabs; and underwent clinical cervicovaginal lavage sampling and portable colposcope examination and imaging. The primary aim was to evaluate the agreement of visual-FGS by two independent experts, and the secondary aims were to compare visual-FGS with other FGS and *S. haematobium* diagnostic techniques as well as self-reported FGS sequelae. While the proportion of FGS cases—by any diagnostic method—was relatively low in this cohort, the authors report only slight agreement between the two visual-FGS experts. These results further emphasize the challenges associated with FGS diagnosis and the need for more reliable and accessible diagnostic methods. This study provides an important contribution to the FGS literature and calls attention to barriers to providing reproductive health care for African women.

<u>Introduction:</u>
1. The authors should provide some background epidemiological info on burden of FGS among women and girls in Africa, associated morbidities (relating to both physical and mental health as some of these are a feature of the secondary analysis), and emphasize that cases are underdiagnosed due to a myriad of reasons (equipment, expertise, access to women's health services, etc). If available, include any info on the burden of urogenital schistosomiasis in Zambia or offer context for why Zambia was chosen as the study site.

2. Please describe techniques for identifying FGS in the introduction (biopsy of genital tissue, visual inspection by colposcopy, PCR performed on cervicovaginal lavage material) and include the advantages and disadvantages of each.

3. It would be helpful to clarify in the introduction that this investigation reports findings of the BILHIV cohort study.

4. Suggestion to state the primary and two secondary objectives of the present study in the introduction section.

<u>Methods:</u>
1. The study benefits from the selection of two highly qualified visual-FGS experts, inclusion of advanced techniques in FGS diagnoses including qPCR of CVL material, and application of the Mobile ODT hand-held colposcope which is one of the devices recommended for FGS diagnosis in schistosomiasis-endemic areas in a 2021 systematic review of hand-held colposcopy equipment.

2. It would be helpful for the authors to clarify which participants were included for STI evaluation and the criteria for how this subset was selected.

3. Please provide detail about the regression model building approach and explanation for the inclusion for age or age and district as covariates. Have the authors considered additional potential confounding factors, such as co-infection with HIV or other STIs, other co-morbidities, sexual history variables, or pregnancy history?

Results:
1. There is a typo in the word "assessment" at the end of the "Visual FGS and Schistosoma laboratory tests" paragraph.

2. It would be interesting for authors to report the agreement between tissue-specific qPCR results (cervical swab, vaginal swab, and CVL).

Discussion:
1. A strong case is presented for the disadvantages of visual-FGS diagnosis.

2. The authors discuss the limited feasibility for PCR-based diagnosis in schistosomiasis-endemic areas and make reference to promising developments in Loop-Mediated Isothermal Amplification (LAMP) methods, which offer added accessibility in resource-limited settings. The authors might consider incorporating discussion or reference to demonstrated LAMP sensitivity for *S. japonicum* and *S. haematobium* (Wang *et al.* 2011[1] and Gandasegui *et al.*, 2015[2]).

References
1. Wang C, Chen L, Yin X, Hua W, et al.: Application of DNA-based diagnostics in detection of schistosomal DNA in early infection and after drug treatment.*Parasit Vectors*. 2011; **4**: 164 PubMed Abstract | Publisher Full Text
2. Gandasegui J, Fernández-Soto P, Carranza-Rodríguez C, Pérez-Arellano JL, et al.: The Rapid-Heat LAMPellet Method: A Potential Diagnostic Method for Human Urogenital Schistosomiasis.*PLoS Negl Trop Dis*. 2015; **9** (7): e0003963 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology and biostatistics; schistosomiasis; infectious disease immunology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 08 February 2023

https://doi.org/10.21956/wellcomeopenres.20776.r54089

? **Vanessa Christinet**
ASCRES, Lausanne, Switzerland

Thank you for inviting me to review this article on a comparative analysis of visual diagnosis of FGS between two experts in visual diagnosis of cervical images. The authors chose to explore the relationship between the diagnosis based on image analysis of the two reviewers and several objective measurements of S. Heamatobium infection diagnosis and also with symptoms described by the patients.

This study and its results are very relevant and interesting as there is little produced in this field of research. It is indeed very important to evaluate diagnosis by visual inspection of the cervix as this is the recommended method of reference but there is a lack of data on the reliability of the diagnosis.

As this study is nested within the HPTN071 study, I would recommend to briefly describe the HPTN071 study in a single sentence, even if it has already been described in other publications, in order to fully understand the profile of the participants and the general context of the study.

Indeed, the question of HIV diagnosis and the level of immunosuppression (if known) should be addressed as it is certainly an important confounding factor for the interpretation of the visualized lesions and could induce an important heterogeneity of cervical lesions making them more difficult to interpret (association between HIV and FGS, association between HIV and other STIs, association between HIV-HPV-dysplasia).

In the description of the method, it is stated that STIs were not screened for POC but later in the text it is written that a whole panel of STIs was tested. Considering that the authors state that microbiological STI diagnosis can be a confounding factor in the diagnosis and symptoms associated with FGS, it seems important to me that they are taken into account in the analysis. As this is not the case here, it would be relevant to justify why.

Cervical cancer screening with visualization of dysplasia is described in the methodology but is not included in the comparative analysis between the two reviewers despite the fact that it is described as a confounder in terms of vascularization.

In the statistical analysis section it is described that the OR was adjusted for age. I would like the authors to describe the rationale for this adjustment and why other confounding factors were not included in this adjustment (STI, HIV, cervical dysplasia). Age seems to me to be relevant for the adjustment of the association with delay of conception, but I do not understand the interest of using it for the other variables (it should be noted that the unadjusted and adjusted ORs are not very different, which speaks to a moderate interest for the inclusion of this adjustment factor).

In the method, additional detail on the number of photos per patient sent to the reviewers would be an enhancement as this could give more information on what they based their evaluation on. It would also be interesting to have more detailed information on the quality of the images to help understand the discrepancy between the two reviewers' analysis. Was a precise description of the lesions made by each of the reviewers which would allow a better understanding of the level of disagreement?

It is noteworthy that although not reaching statistical significance, both reviewers consistently identified a greater proportion of FGS lesions in individuals with objective markers of schistosome infection. If the sample size had been larger, statistical significance would probably have been reached for some variables. If the data are available, I think it would be very interesting to exploit the STI diagnoses here to see how they probably contributed to the heterogeneity of the analysis.

I find the presentation of the results not very easy to understand. The authors refer several times to the non-association without describing precisely the significance of the analysis which is actually the lack of a statistically significant difference between patients diagnosed by positive or negative visual inspection by the reviewers in relation to objective parameters of schistosome infection. Again, it should be noted that even if the differences are not statistically significant, the ORs are still positive in favour of the objective measurements of schistosomiasis. I suggest rewording the text below Table 3.

It is interesting to note that the symptom with the highest OR for both reviewers is for the association with haematuria, which is the most specific symptom of Schistosoma Heamatobium infection.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Sexual health, HIV, communicable disease, tropical medicine, public health, epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 03 February 2023

https://doi.org/10.21956/wellcomeopenres.20776.r54092

Daniela Fusco [iD]
Infectious Disease Epidemiology, Bernhard Nocht Insitute for Tropical Medicine, Hamburg, Germany

The manuscript deals with a very relevant topic in need for urgent solutions. The methods and design are adequate for the study and the results are properly described.

Some minor recommendations to the manuscript:
- Title: In the title exclusively the aspect of the agreement of experts emerges, while in the results there is a long description of a regression analysis to associate factors (i.e. symptoms) to visual FGS. This aspect should emerge in the title. Additionally in the abstract it is stated that "The agreement of expert human reviewers for visual-FGS has not been previously described". In this view, the hand-held colposcopy doesn't play a central role in the study hence it should be considered to be removed from the title.

- Methods: In the paragraph "Home and clinic-based sample collection" the type health care professionals performing the sampling should be specified. In the paragraphs "qPCR for detection of Schistosoma DNA" and "Other infections" samples storage conditions should be specified.

- Discussion: As per in the title, in this statement "This study shows "slight" agreement

between senior, highly experienced expert reviewers, highlighting the imperfect nature of human expert review of images obtained with hand-held colposcopy for FGS." the reader can have the feeling that the agreement might be different if another type of colposcope would be used. It would be advisable to reconsider the statement. From the reference 26 the statement "The Tanzanian study illustrates the limited specificity of visual techniques, since one-third of the women had cervical lesions in communities where S. haematobium is not endemic." cannot be really deduced since, i.e., travel or medical history of women is not described. In this view this statement doesn't seem to reflect the message of the reference and it should be re-considered. The conclusion statement "...we suggest caution when visual imaging is used as a stand-alone FGS diagnostic." should bring some recommendations i.e. on how to interpret or obtain better results through colposcopy since, so far, colposcopy is still the diagnostic standard for the disease and alternatives, even if urgent and needed, are not really available.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Public Health, schistosomiasis, FGS, women's health, diseases prevention, vaccinology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 31 January 2023

https://doi.org/10.21956/wellcomeopenres.20776.r54086

✔

Alexander B. Odaibo (iD)

Parasitology Research Unit, Department of Zoology, University of Ibadan, Ibadan, Nigeria

The authors evaluated the agreement of human expert reviewers on visual-FGS diagnosis with hand-held colposcopy. Two clinical experts in the diagnosis of FGS in endemic settings were engaged in the study. They assessed 527 women cervicovaginal colposcopy images independently. The study recorded a "slight" agreement between the expert reviewers of the digital images obtained during point-of-care colposcopy. The design was good and adequate but there is a need to improve on the details of some aspects of the methods to allow replication by others.

1. Urine sample was collected during home visit - the authors should provide specific time of urine collection, as this is important in the diagnosis of schistosomiasis. Was the time of collection the same for all participants?

2. Up to 50 mL of urine was centrifuged and examined - the authors should provide the specific amount of urine centrifuged, and why 50 mL and not 10 mL as recommended by WHO (2022[1]), and for ease of comparison with other studies.

3. Figure 1. Flowsheet of cervicovaginal image review - the authors should explain what "Images inaccessible" means, particularly as it was among those who attended clinic.

4. The authors use the phrase "water use" and "water contact" interchangeably. I think "water contact" is best for this study.

References

1. WHO: WHO guideline on control and elimination of human schistosomiasis. *World Health Organization*. 2022. Reference Source

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Parasitologist/Epidemiologist

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 27 January 2023

https://doi.org/10.21956/wellcomeopenres.20776.r54087

? **Louise Arenholt**
[1] Department of Clinical Medicine, Aalborg University, Aalborg, Denmark
[2] Department of Obstetrics and Gynecology and Centre for Clinical Research, North Denmark Regional Hospital, Hjørring, Denmark

Thank you for the opportunity to review this thorough and well written paper.

The paper aims to evaluate the agreement on the diagnosis of visual FGS lesions between two expert reviewers. Furthermore, the association between visual presentation of FGS and symptoms of FGS/PCR detection of FGS were evaluated.

The topic is highly relevant and only a few studies have evaluated the accuracy of visual diagnosis of FGS – even though the visual inspection is the golden standard as also highlighted by the authors.

Introduction:
No comments.

Methods:
I would recommend more detailed information on the hand-held colposcopy and image review. There is no information on how the image was captured and if they had a standard/protocol on which areas to include in the image portfolio (was it only the cervix or also the vaginal walls?). Did they have any criteria for image quality (for example the lighting or zoom level) when capturing the image? How many images were captured per woman? It seems like the reviewers have excluded images on slightly different basis – did you before the review guide the reviewers on how to include or exclude the images? For example, did the reviewers review the images on a computer and were the computers similar (for example resolution on the screen, color setting, brand etc.)? Did they have the opportunity to zoom the image or change color/brightness/contrast/saturation? My experience is that an image can look completely different if you evaluate it on two different computers. A small thing as day light or artificial light

can change the image. Could it be different settings/equipment that account for some of the difference between reviewers?

In the methods section I can't find information on how women reported the abdominal-genitourinary, and reproductive manifestations (table 4) – questionnaire, interview etc.?

Results:

I am not surprised that you do not find an association between visual-FGS and PCR-FGS. I believe that cevicovaginal lesions are results of chronic inflammation and therefore other tests should be used if an association should be found.

Discussion:

You discuss the study from Madagascar, where gridded image technique was used. And you are correct that all included women in that study were to have FGS lesions. I just need to address two things 1) the presence of rubbery papules was not an inclusion criterion for the study. Therefore, the uncertainty about whether a lesion was present or not was indeed an issue. Some of the reviewers in that study found, that more than 20% of the women had no rubbery papules. 2) the images used for the study were not from the inclusion visit, but from later visits where treatment with Praziquantel had been initiated. For that reason, some of the images could be without FGS lesions.

I agree with you that consensus on how to evaluate the images is very important. I believe that if consensus had been reached before the reviewing procedure, your results would have been very much different.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Gynecology and obstetrics. Urogynecology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.