Computer Image Analysis as a Diagnostic Tool in Female Genital Schistosomiasis

Sigve Holmen, M.D.

Thesis for the degree of Philosophiae Doctor, PhD

Norwegian Centre for Imported and Tropical Diseases

Oslo University Hospital, Ullevål

Oslo, Norway

and

Institute of Clinical Medicine

University of Oslo

Oslo, Norway

2016

Table of contents

Table of contents 2

List of figures 5

Acknowledgements 8

Definitions and abbreviations 11

List of publications 13

Related papers 14

Synopsis 15

Background 17

Schistosomiasis 17

Schistosomes 17

The life of a schistosome 18

Disease in humans 19

Female genital schistosomiasis 20

Epidemiology 20

Pathophysiology 20

Clinical presentations 21

Association with HIV 24

Diagnosing female genital schistosomiasis 26

Patient history and reported symptoms 26

Clinical investigation 27

Microscopy 27

Molecular and serological techniques 28
Diagnostics using computer image analysis 29
Colour analysis 29
Analyses of morphology 30
Research questions and aim 32
Materials and methods 33
Study design and participants 33
Study area and population 33
Ethical considerations 34
Access to and storage of data 34
Clinical investigation and questionnaire 35
Laboratory analyses 36
Statistical analyses 37
Comparisons between groups 38
Analyses of predictive accuracy 38
Regression models 38
Latent class analyses 39
Computerised image analyses 40
Selection of image material 40
Software framework 40
Identifying the region of interest 41
Specular reflections 43
Colour analyses of sandy patches 44

Abnormal blood vessels....47

Combining colour and morphological analyses 56

Results 58

Colour analyses 58

Morphological analyses.... 61

Discussion....63

The need for a new diagnostic method in FGS 63

Why computerised image analysis? 64

Sandy patches, abnormal blood vessels or both? 65

Limitations.... 66

Future research 66

References 68

List of figures

Figure 1. Distribution of human schistosomiasis by species. (Reproduced with permission from Daniel Colley, Colley et al., 2014).....17

Figure 2. Left: Schematic representation of the rice grain shapes seen in grainy sandy patches. Right: Photocolposcopic image of a grainy sandy patch at high magnification.....22

Figure 3. Left: Schematic representation of homogenous yellow sandy patch. Right: Photocolposcopic image of a homogenous yellow sandy patch. 22

Figure 4. Left: Schematic representation of the abnormal blood vessels associated with female genital schistosomiasis. Right: Photocolposcopic image of abnormal blood vessels at high magnification. 23

Figure 5. Left: Schematic representation of rubbery papules. Right: Photocolposcopic image of rubbery papules.....24

Figure 6. Digital schematic representing the ectocervix and the fornices. The clinician inserts the inner and outer delimitation of the transformation zone (yellow and blue, respectively) and colours the areas (pink) affected by lesions for future reference.....36

Figure 7. An intermediate graphic generated by one of the plugins written for ImageJ, showing the result of k-means clustering (magenta and green) as part of the process of determining the region of interest (the cervical surface). The abscissa shows the distance from the centre of the original image and the ordinate shows increasing values of the a-channel from the Lab-colour space, representing increasing values of magenta (pink).....41

Figure 8. The region of interest was refined (dashed line) by expanding the initial coarse mask (solid line) by 50 concentric rays until a drop in pixel values of the Value channel (HSV colour space). The rays were finally smoothing using a 5-point mean filter to ensure a smooth delimitation: notice the three rays protruding outside the dashed line in the upper left corner. 42

Figure 9. Perimeter gradients around specular reflections may appear yellow.....43

Figure 10. Left: the original image. Right: specular reflections have been masked (blue) by filtering on low saturation. 44

Figure 11. The colour detection algorithm was applied adaptively to perform with a wide range of exposure settings. The delimited areas represent the yellow sandy patches as detected by the algorithm. 47

Figure 12. The multiplication of the inverted green channel and the saturation channel results in an image where the blood vessels have "boosted" values, appearing more clearly than in either of the two original colour channels. 49

Figure 13. **A.** Original colour image. **B.** Equalized product of the inverted Green channel and the Saturation channel. The automatically detected region of interest (ROI) is indicated in red. **C.** The standard deviation of grey values in a 50px window calculated for each pixel and represented as relative levels of grey values (0-255). **D.** Thresholding performed using the mean grey value. **E.** Thresholding using the standard Niblack method with a k-value of -0.2. Notice the perimeter artefact as a thick white line around the ROI. **F.** Thresholding using Niblack's method with calculation of optimal k-value per 50px sliding window and elimination of the perimeter effect. 51

Figure 14. **Top**: A binary blood vessel structure after image processing. **Middle**: A log-log plot showing the number of boxes (N) required to cover the blood vessel

in relation to box size (ε). The slope of the dashed red regression line is 1.11, corresponding to the estimated fractal dimension, D. **Bottom:** A log-log plot showing the mean λ-value in relation to decreasing box size (ε). The slope of the dashed red regression line is 0.82, corresponding to the estimated lacunarity of the blood vessel. 54

Figure 15. **A.** The original colour image. **B.** After extracting the region of interest, the boosted image is thresholded by using the mean grey value. This leaves primarily blood vessel structures in the image. **C.** The circular convolution template. **D.** The result of the convolution of B and C (multiplication in the frequency domain). **E.** The result of thresholding image D on the 97.5th percentile of pixel value. Each dot represents a positive match for a circular feature. **F**. The result of the template matching (E) superimposed on the original colour image (A). 56

Figure 16. Flow chart describing the inclusion of images for validation of the colourimetric image analysis. 59

Acknowledgements

Thanks to my tutor **Eyrun Kjetland** for offering me a better alternative to military service, allowing me to come to South Africa to discover the challenges and rewards that come with fieldwork. Her lack of fear in seeking out new solutions to old problems resulted in our endeavour into new technologies. She always pushes the limits of what we think of as feasible and advocates the new order to old ears.

Thanks to my tutor **Fritz Albregtsen** for sharing his technological and innovative insights into solving problems, for giving me an office at his faculty when none was to be found at mine and for battling the bureaucracy that dwells deep within the university.

Thanks to my tutor **Mathias Onsrud** for stepping up as a gynaecological authority, helping me to understand the many issues that need to be taken into account when studying genital disease.

Thanks to **Bjørn Myrvang** and the **Norwegian Centre for Imported and Tropical Diseases** for believing in my project and me when other funders did not, allowing me to embark on this PhD.

Thanks to **Borghild Roald** and **Peter Jourdan** for introducing me to schistosomiasis through immunohistopathology and paving the way for applying computerized image analysis to solve problems in pathology.

Thanks to **Elisabeth "Lisette" van Lieshout, Birgitte Jyding Vennervald and Svein Gunnar Gundersen** for sharing their vast experience in tropical medicine and field work.

Thanks to **Patricia Ndhlovu** for challenging my views on alcohol, meat and data entry. Her down to earth and pragmatic world-view inspires me.

Thanks to **Myra Taylor** for being the local anchor at UKZN, always welcoming in her office and never shy to lend a helping hand.

Thanks to **Leiv Sandvik** for providing statistical guidance and more importantly for sharing his wisdom in life and for accompanying me in gastronomic adventures in Durban.

Thanks to **Elisabeth Kleppa** for sharing her invaluable critical thinking, allowing us to see problems from multiple angles and providing thorough feedback to manuscripts.

Thanks to **Hashini Nilushika Galappaththi-Arachchige** for housing me during field work and providing her alternative views to scientific as well as philosophical problems.

Thanks to **Pavitra Pillay** for her kind words and support and advice in life and shopping.

Thanks to **Roy Farai Manyaira** for his cheerful and tireless assistance in managing, fixing and improving the local IT-infrastructure, and for his excellent data-entering skills (which we proved to exceed those of a computer!).

Thanks to **Sipho Senkosi Zulu** for contributing with his knowledge of local issues and for good companionship on our journey to Leiden.

Thanks to **Pamela Ramkalawon, Parashti Maharaj, Fazana Karim** and **Andile Mtshali** for supporting me despite my poor laboratory skills and providing able hands and minds to solve the problems and get the analyses done. Thanks to **Prashini Moodley** for placing her laboratory at our disposition and assisting in setting up the correct assays.

Thanks to **Gordon Bailey** for his wonderful stories and his practical hands in forming the clinic to meet our every need and thanks to his wonderful wife **Molly Bailey** for helping us with MTCs and

Thanks to the staff and students at the clinic; **Adele Munsami, Andile Maphumulo, Andrea Lothe, Bongekile Dlomo, Duduzile Zwane, Elphinah Kwela, Glory Hlengwa, Jabu Msomi, Joann Goldstone, Joelyn Hargreaves, Juventus Pillay, Khantso Kolisang, Lingani Buthelezi, Londiwe Radebe, Lucky Dlomo, Manelisi Majiya, Mpume Ngwabe, Nina Kumud Henriksen, Nobuhle Mbiza, Nokubonga Lubanyana, Nosihle Maphumulo, Ntombie Lubanyana, Nqobile Dlomo, Priscilla Cele, Pumla Mkhiva, Sbongile Nzimande, Silindile Gagai, Sheshile Sibiya, Sindiswa Buthelezi, Nombeko Mpofana, Nompumelelo Ngwabe, Nozipho Mkhabela, Sibongile Khumalo, Simangele Malinga, Sylvia Sosibo, Thandeka Ziqubu, Thandekile Nxumalo, Thobekile Mqadi, Thulisile Sithole, Zama Cele, Zibuyile Chiliza, Zodwa Ngcobo** and **Zolile Ndovela** for their cheery mood, beautiful voices and relentless work in keeping the clinic running and allowing the scientific work to progress. My heart goes out to Thandeka who passed away too early.

Thanks to students, researchers and good helpers from across the globe for coming to South Africa and helping out; **Ana Randjelovic, Anna Kildemoes, Artemis Koukounari, Christine Aurlund, Debbie Clark, Erik Christensen, Gunn-Hege Karlsen, Hege Lovise Lundgreen, Ida Amelie Helgesen, Ingrid Lindström, Ingrid Hegertun, Jilna Shukla, John-Erik Armelius, Kari Klinge, Kristine Lillebø, Linnea Stråbø, Marc Baay, Mari Molvik, Susanne Klungtveit** and **Synne Grønvold.**

Definitions and abbreviations

CAA	Circulating anodic antigen
CMYK colour space	Colour space consisting of the four channels: cyan, magenta, yellow and black
Colposcopy	The use of a colposcopic microscope (a colposcope) to investigate the cervical and vaginal surfaces at magnification
DNA	Deoxyribonucleic acid
Ectocervix	The visible surface of the vaginal portion of the uterine cervix
FGS	Female genital schistosomiasis
Grainy sandy patch	Sandy patch appearing as rice shaped grains, either dispersed, clustered or a combination
HIV	Human immunodeficiency virus
Homogenous, yellow, sandy patch	Sandy patch appearing as a homogenous area of yellow discoloration, without visible grains.
HPV	Human papillomavirus
HSB colour space	Colour space consisting of the three channels: hue, saturation and brightness
HSV colour space	Colour space consisting of the three channels: hue, saturation and value
HSV	Herpes simplex virus
Lab	Colour space consisting of the three channels: luminance, a and b
Mpx	Millions of pixels (megapixels)
Mucosa	The epithelium lining certain parts of inner organs, such as the vaginal wall and the ectoverix.
Ovum / ova	Egg / eggs
PCR	Polymerase chain reaction

RGB colour space	Colour space consisting of the three channels: red, green and blue
Sandy patch	Mucosal discoloration of yellow or white colour, appearing on the urothelium of the urinary bladder and/or on the mucosal surfaces of the lower genital tract in urogenital schistosomiasis.
SDA	Strand displacement assay
SIV	Simian immunodeficiency virus
SLR	Singe lens reflex camera
SNR	Signal to noise ratio: The level of desired signal in the true digital image to the level of the noise occurring during image acquisition and compression.
STI / STD	Sexually transmitted infection / disease
Transformation zone	The cervical mucosa between the original and current squamocolumnar junctions.
UGS	Urogenital schistosomiasis: Schistosomiasis caused by *Schistosoma haematobium*, which affects the urinary and/or genital tract.
VIA	Visual inspection with acetic acid
WHO	World Health Organization

List of publications

I Colourimetric image analysis as a diagnostic tool in female genital schistosomiasis. Medical Engineering and Physics 2015;37: 309–314. doi:10.1016/j.medengphy.2014.12.007 **Holmen, S.D.,** Kjetland, E.F., Taylor, M., Kleppa, E., Lillebø, K., Gundersen, S.G., Onsrud, M., Albregtsen, F.

II The first step towards diagnosing female genital schistosomiasis by computer image analysis. American Journal of Tropical Medicine and Hygiene 2015;93: 80–86. doi:10.4269/ajtmh.15-0071

Holmen, S.D., Kleppa, E., Lilleboe, K., Moodley, J., Taylor, M., Albregtsen, F., Vennervald, B.J., Onsrud, M., Kjetland, E.F.

III Characteristics of Blood Vessels in Female Genital Schistosomiasis: Paving the Way for Objective Diagnostics at the Point of Care. PLoS Neglected Tropical Diseases 2016: e0004628. doi: 10.1371/journal.pntd.0004628

Holmen, S., Galappaththi-Arachchige, H.N., Kleppa E., Naicker, T., Taylor M., Onsrud M., Kjetland E.F., Albregtsen F.

Related papers

Jourdan, P.M., **Holmen, S.**, Gundersen, S.G., Roald, B., Kjetland, E.F., 2011. HIV target cells in Schistosoma haematobium-infected female genital mucosa. Am. J. Trop. Med. Hyg. 85, 1060–4.

Kleppa, E., **Holmen, S.**, Lillebø, K., Pillay, P., Taylor, M., Gundersen, S.G., Kjetland, E.F., Onsrud, M., 2014. Re: Al-Baghdadi O, Samarasinghe A, Wissa I. 2014. Cervical schistosomiasis. Journal of Obstetrics and Gynaecology 34:206. J. Obstet. Gynaecol. 34, 755–756.

Kleppa, E., **Holmen, S.**, Lillebø, K., Kjetland, E.F., Gundersen, S.G., Taylor, M., Moodley, P., Onsrud, M., 2014. Cervical ectopy: associations with sexually transmitted infections and HIV. A cross-sectional study of high school students in rural South Africa. Sex. Transm. Infect. 1–6.

Kjetland, E.F., Norseth, H.M., Taylor, M., Lillebø, K., Kleppa, E., **Holmen, S.**, Andebirhan, A., Yohannes, T.H., Gundersen, S.G., Vennervald, B.J., Bagratee, J., Onsrud, M., Leutscher, P.D.C., 2014. Classification of the lesions observed in female genital schistosomiasis. Int. J. Gynaecol. Obstet. 2–3.

Kleppa, E., Ramsuran, V., Zulu, S., Karlsen, G.H., Bere, A., Passmore, J.-A.S., Ndhlovu, P., Lillebø, K., **Holmen, S.**, Onsrud, M., Gundersen, S.G., Taylor, M., Kjetland, E.F., Ndung'u, T., 2014. Effect of female genital schistosomiasis and anti-schistosomal treatment on monocytes, CD4+ T-cells and CCR5 expression in the female genital tract. PLoS One 9, e98593.

Norseth, H.M., Ndhlovu, P.D., Kleppa, E., Randrianasolo, B.S., Jourdan, P.M., Roald, B., **Holmen, S.**, Gundersen, S.G., Bagratee, J., Onsrud, M., Kjetland, E.F., 2014. 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 8, e3229.

Synopsis

Female genital schistosomiasis (FGS) is a water-borne parasitic disease, which is estimated to affect more than 35 million girls and women, primarily in Sub-Saharan Africa. This disease is largely co-endemic with HIV and there is growing evidence that FGS may render populations more susceptible to HIV although the cause-effect relationship has not been definitely determined.

However, there are currently no adequate diagnostic tools for FGS in endemic areas, rendering the assessment of this disease difficult. This is a problem on the individual level but it also poses a problem to conducting studies exploring FGS.

In this thesis, I present three papers that explore an alternative diagnostic route; applying computerised image analysis to assess the characteristic sandy patches and the abnormal blood vessels associated with FGS in an attempt to devise a novel diagnostic tool that may be used at the point of care in low-resource settings. A software-based tool can be free and easy to distribute, even to remote areas. It can also be used on a wide range of devices, from laptops to cell phones with cameras capable of running third party software ("apps").

The assessment of sandy patches relies on colour analysis of the cervical mucosa with detection of areas that appear more yellow than the surrounding mucosa. Applying this method may detect 80.5 % of cases but with a limited specificity of 66.2 %.

The abnormal blood vessels associated with FGS present specific morphological features that may be used for computerised detection: circular shapes, increased fractal dimension of the overall vascular network and reduced local fractal

dimensions. By using these features, we were able to classify images with an accuracy of 78 %.

In conclusion, it is possible to create a computerised diagnostic tool for FGS but it is necessary to increase accuracy. Combination of colour analysis and vessel morphology may increase specificity but further research should also explore other imaging modalities, such as intra-vaginal spectrometry or micro-endoscopic in-vivo microscopy for more accurate readings.

Background

Schistosomiasis

Schistosomes

Schistosomes are trematodes of which more than twenty species have been described. The genus *Schistosoma (S.)* can be divided into four subgroups: *haematobium*, *mansoni*, *japonicum* and *indicum*. Each subgroup has a number of species of which some may cause pathology in humans, notably the species *S. haematobium*, *S. mansoni*, and *S. japonicum,* which are estimated to be the most commons cause of disease in humans (Colley et al., 2014).

S. mansoni and *S. haematobium* are both prevalent in Africa and the Middle-East whereas *S. japonicum* only exists in Asia (primarily China and the Philippines). *S. mansoni* is also prevalent in South America (primarily Brazil, Figure 1) (Colley et al., 2014).

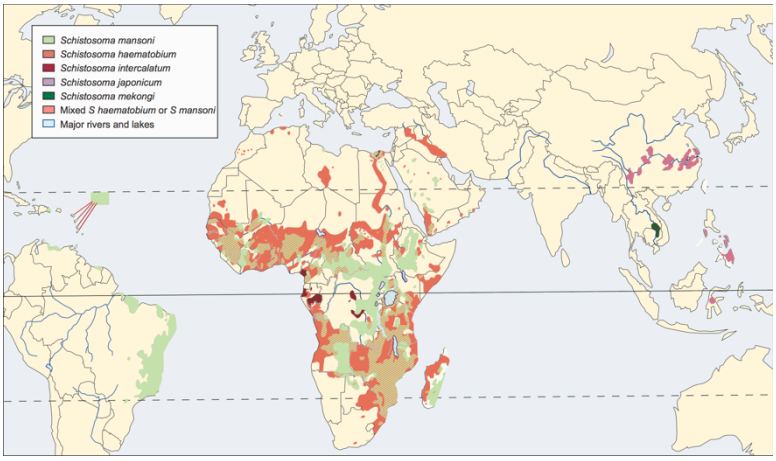

Figure 1. Distribution of human schistosomiasis by species. (Reproduced with permission from Daniel Colley, Colley et al., 2014)

The life of a schistosome

After an egg has been excreted from the main host (e.g. a human), a miracidium hatches from the egg and relies on finding its intermediate host; a snail (e.g. *Bulinus Africanus* or *Biomphalaria* for *S. haematobium* and *S. mansoni*, respectively). In the snail, the parasite undergoes asexual replication, resulting in the release of thousands of cercariae back into the water again (Colley et al., 2014; Gryseels et al., 2006).

Upon contact with the exposed epidermis of a susceptible host (e.g. a human), a cercarium will attach to the skin, shed its flagellum and make its way into the skin where it matures to the next stage; the schistosomulum (Colley et al., 2014). The schistosomulum will then penetrate a venule, which will carry it into the blood stream and to the lungs, where the schistosomulum matures further (Colley et al., 2014). Eventually it migrates to the liver sinosoids where it pairs with a partner schistosomulum of the opposite sex. The male worm encloses the female worm in a gynaecophoric slit. The pairing forms a union lasting the lifetime of the worms, which is generally 3-10 years but may be as long as 40 years (Colley et al., 2014; Gryseels et al., 2006).

Finally, the worm pair migrates from the liver to the organ where it will lay eggs. The different schistosome species have specific affinity for which organs they end up in: *S. haematobium*, tends to migrate to the venous plexa surrounding the urinary bladder and the genital tract. Here, the eggs are excreted with urine, semen and gynaecological fluids (Kjetland et al., 2005; Leutscher et al., 2008). *S. mansoni and S. japonicum* tend to migrate to the mesenteric veins or rectal veins as to shed their eggs in the intestine for excretion with fecal matter (Gryseels et al., 2006).

Disease in humans

The first manifestation of schistosome infection may occur as a urticarial rash named "cercarial dermatitis" or "swimmer's itch", caused by the penetration of the epidermis by cercaria, occurring rapidly after contact with infested water (Gryseels et al., 2006). After weeks to months, the infected patient may experience "acute schistosomiasis" or "Katayama fever" in the form of a febrile illness with fatigue, myalgia and cough as a response to the schistosomules migrating to the lungs (Gryseels et al., 2006). It is thought that infections with *S. haematobium* and *S. mansoni* generally do not cause acute schistosomiasis in patients from chronically exposed populations, possibly due to intrauterine sensitisation (Gryseels et al., 2006). However, *S. japonicum* may cause Katayama fever even in patients who have been infected previously.

Acute schistosomiasis usually resolves within weeks and about 5-7 weeks after the initial infection, the worms start laying eggs, resulting in the chronic stage of schistosomiasis.

In infections caused by *S. mansoni* or *S. japonicum*, the chronic disease is characterized by intermittent abdominal pain, loss of appetite and diarrhoea. Hepatic inflammation occurs as some eggs spill into the mesenteric veins and become trapped in liver veins (Gryseels et al., 2006). Years later, liver fibrosis may develop if the individual has the necessary immunogenetic susceptibility (Dessein et al., 1999; Gryseels et al., 2006). Chronic infections caused by *S. haematobium* are characterized by a wide range of gynaecological symptoms and signs in women, known as female genital schistosomiasis (FGS) (Kjetland et al., 2012). In men it can cause haematospermia (Leutscher et al., 2000). Eggs deposited in the bladder wall causes haematuria and may cause squamous bladder cancer (Gryseels et al., 2006).

Female genital schistosomiasis

Epidemiology

In 2003 it was estimated that 112 million people were currently infected with *S. haematobium* (Van Der Werf et al., 2003). Of the infected women, it is estimated that somewhere between 33 and 75 % have FGS (Hotez et al., 2009; Poggensee et al., 1999; Swai et al., 2006). In areas endemic of *S. haematobium*, it is likely that the first infection occurs very early in life in relation to water activities such as bathing and playing (Colley et al., 2014; Hegertun et al., 2013; Kjetland et al., 2012). However, there are major diagnostic challenges with FGS (see the chapter "Diagnosing female genital schistosomiasis"), especially in young girls where gynaecological investigations cannot be performed routinely. It is therefore difficult to estimate the exact prevalence of FGS.

Pathophysiology

As the adult worms deposit eggs, proteolytic enzymes excreted from the miracidium within the egg start degrading the surrounding tissues in an effort to exit the body and continue the schistosome life cycle (Gryseels et al., 2006). However, many eggs fail to exit the tissues before they die after 1-2 weeks (Colley et al., 2014). These eggs induce a granulomatous immune-reaction, histologically characterised by the presence of lymphocytes (Colley et al., 2014). In cervical biopsies from an endemic area, we found that women with FGS have more CD4+ T-lymphocytes in the cervical mucosa (Jourdan et al., 2011a). In women with

urogenital schistosomiasis (UGS), eggs may be deposited throughout the genitals; in the ovaries, fallopian tubes, uterus and vaginal tract (Kjetland et al., 2012).

Clinical presentations

Women and girls suffering from FGS may present with a wide range of urogenital symptoms, which manifest from early childhood (Hegertun et al., 2013). They include haematuria, dysuria, pollakisuria and/or incontinence, bloody and/or malodorous vaginal discharge, genital itch and pain (Hegertun et al., 2013; Kjetland et al., 2012). In sexually active women, there are also frequent reports of dyspareunia and post-coital bleeding (Kjetland et al., 2012).

By gynaecological investigation, lesions may be visualised as sandy patches and rubbery papules (Kjetland et al., 2012). In addition, *S. haematobium* in the genitals is associated with the presence of abnormal blood vessels and contact bleeding (Kjetland et al., 2012).

Sandy patches

In 1907, Madden described the sandy appearance of lesions due to genital schistosomiasis as "wet sea-sand deposits" (Madden, 1907). Sandy patches have since been described as pathognomonic for FGS (Badawy, 1962; Friedberg et al., 1991; Gibson, 1925; Youssef et al., 1970). In 2005 Kjetland introduced a further subdivision of sandy patches; grainy sandy patches and homogenous yellow sandy patches, based on findings in Zimbabwe (Kjetland et al., 2005).

The grainy sandy patches have rice-grain shapes in clusters of up to 300 grains, situated both superficially and deep in the mucosa (Figure 2) (Kjetland et al., 2005). The grains are estimated to measure approximately 0.005 x 0.2 mm and superficial

grains have a tactile firmness when touched with a metal spatula (Kjetland et al., 2005). These lesions were found to be significantly associated with the finding of *S haematobium* eggs in genital specimens (Kjetland et al., 2005).

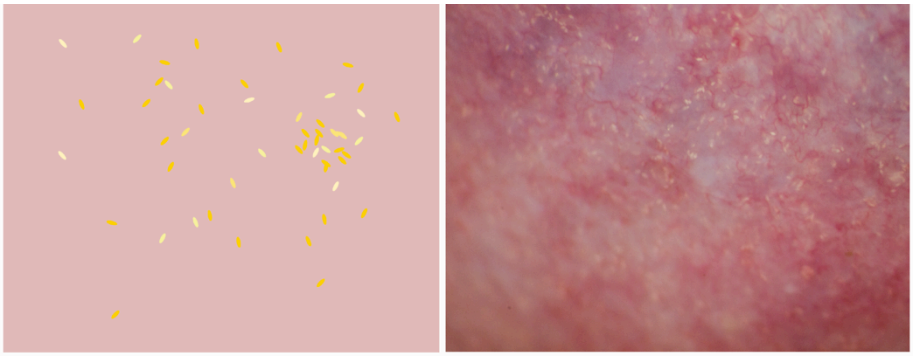

Figure 2. Left: Schematic representation of the rice grain shapes seen in grainy sandy patches. Right: Photocolposcopic image of a grainy sandy patch at high magnification.

The homogenous yellow sandy patches have been described as areas of yellow discoloration without any distinct grains (Figure 3) (Kjetland et al., 2005). These lesions were also found to be associated with the finding of *S. haematobium* eggs in genital samples but were also associated with the finding of *Herpes simplex* virus type 2 (HSV-2) (Kjetland et al., 2005).

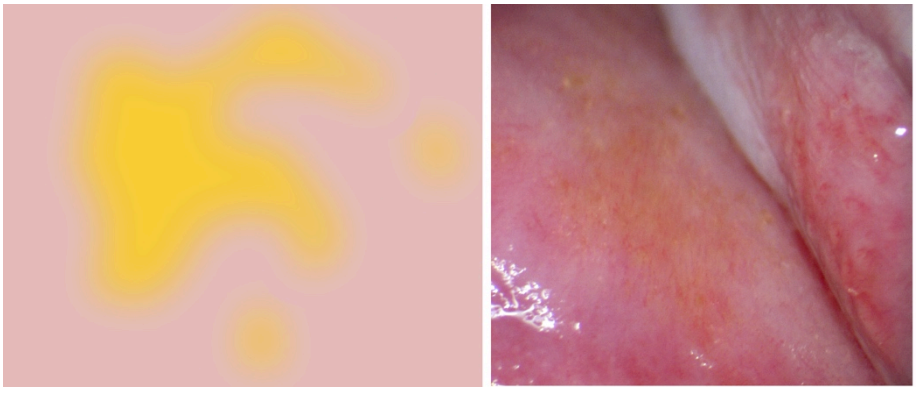

Figure 3. Left: Schematic representation of homogenous yellow sandy patch. Right: Photocolposcopic image of a homogenous yellow sandy patch.

Abnormal blood vessels

Mucosal *S. haematobium* is associated with the finding of abnormal blood vessels (Figure 4) on the cervix and it has been proposed to be a possible cause of increased contact bleeding in FGS patients (Kjetland et al., 2005, 1996). It was hypothesized that the increased vascularity seen in genital mucosa of patients suffering from FGS represents neovascularization (Kjetland et al., 1996). However, a histopathological study on cervical biopsies did not find active neovascularization around schistosome eggs (Jourdan et al., 2011b). A case report presenting clinical images alongside the histological correlates showed dilated and tortuous venules containing viable schistosome eggs surrounded by a thrombus (Jourdan et al., 2013). Clinically, the abnormal blood vessels have been described as being reticulated, branched, convoluted, circular and having uneven calibre (Kjetland et al., 2012, 2005; Norseth et al., 2014).

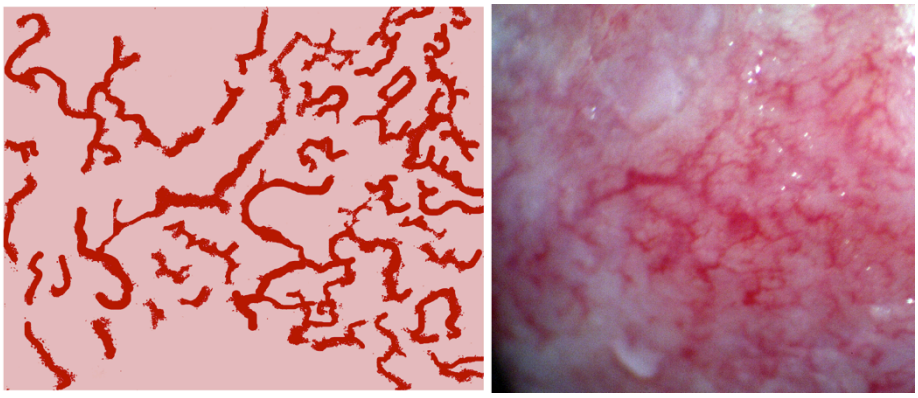

Figure 4. Left: Schematic representation of the abnormal blood vessels associated with female genital schistosomiasis. Right: Photocolposcopic image of abnormal blood vessels at high magnification.

Rubbery papules

Rubbery papules (also called tubercles or nodules) have been defined as spheroid, firm, beige, smooth papules approximately 0.3–1.2 mm large (Figure 5)

(Randrianasolo et al., 2015). This is similar to what has been described in the urinary bladder (Kirkaldy-Willis, 1946). However, rubbery papules in the female genital tract have only been observed in Madagascar (Randrianasolo et al., 2015). Histological investigation of rubbery papules reveal *S. haematobium* eggs surrounded by an intense eosinophil response (Randrianasolo et al., 2015).

Figure 5. Left: Schematic representation of rubbery papules. Right: Photocolposcopic image of rubbery papules.

Association with HIV

In 1994 It was proposed that FGS may render subjects more susceptible to STIs and more notably to HIV, based on epidemiological data and the pathophysiological properties of FGS (Feldmeier et al., 1994). However, it was not until almost 12 years later that the first cross sectional study explored the possible association between FGS and HIV and found a significant association between the presence of homogenous yellow sandy patches and having HIV (Kjetland et al., 2006). Another cross-sectional study confirmed the association in 2011 (Downs et al., 2011). A third cross-sectional study analysed compiled HIV and schistosomiasis prevalence on national levels across 43 sub-Saharan African countries and found further evidence of an association (Ndeffo Mbah et al., 2013).

Although it is not possible to conclude on the causality of this association based on cross sectional studies, a number of mechanisms have been outlined as possible factors that may contribute to both the increased susceptibility to becoming infected and to the transmission of HIV:

- Tissue damage in the form of lesions and friable blood vessels may represent a point of entry for HIV (Kjetland et al., 2005)
- Local inflammation with an abundance of HIV target cells (CD4+ T-cells) in cervical mucosa (Jourdan et al., 2011a)
- Local up regulation of the HIV co-receptor CCR5 in genital CD4+ T-cells (Kleppa et al., 2014b)
- Increased shedding of CD4+ T-cells in semen from men with UGS (Leutscher et al., 2005)

Furthermore, it is plausible that schistosome infection generally precedes HIV infection in most individuals as the former is often contracted during childhood whereas the latter is most often contracted sexually, suggesting a cause-effect association between *S. haematobium* and HIV (Mbabazi et al., 2011).

This cause-effect association is further supported by an animal study, which showed that macaques infected with *S. mansoni* required a 17-fold lower dose of Simian immunodeficiency virus (SIV) instilled rectally to acquire the disease (Chenine et al., 2008).

Diagnosing female genital schistosomiasis

Patient history and reported symptoms

Red urine (gross haematuria) is a frequent symptom in UGS, often accompanied by pollakisuria and dysuria, and in certain endemic areas, haematuria is so widespread it is even considered normal (Colley et al., 2014). Gross haematuria may be used to identify high-risk areas (WHO, 1995) but on an individual level it does not provide sufficient accuracy for the diagnosis of FGS since there may be genital involvement without urinary involvement and vice versa (Kjetland et al., 2012).

History of water contact is a theoretical requirement for infection but even patients without any recollection of freshwater contact may have schistosomiasis (Hegertun et al., 2013).

Genital symptoms such as bloody discharge, genital itch, malodorous discharge, burning sensation in the genitals, genital ulcers and genital tumours may be indicators of FGS in areas endemic of *S. haematobium* if sexually transmitted infections (STIs) can be ruled out (Hegertun et al., 2013). However, even in alleged virgins, it is impossible to rule out STIs due to sexual abuse and non-reported sexual activity based on patient history alone.

Hegertun et al showed that children with a history of having had UGS or a history of a family member having UGS were significant risk factors for genital symptoms, which may be attributable to FGS (Hegertun et al., 2013).

Clinical investigation

In patients from areas endemic of *S. haematobium*, one of three clinical findings, by visual inspection, may be diagnostic for FGS (Kjetland et al., 2014, 2012):

- Grainy sandy patches
- Homogenous yellow sandy patches
- Rubbery papules

Abnormal blood vessels are often seen concurrently and may support the diagnosis (Kjetland et al., 2014). Cervical cell atypia is an important differential diagnosis. It is therefore important to note that none of the above-mentioned lesions are aceto-white and that they are not restricted to the transformation zone (Kjetland et al., 2014). The visual inspection may be enhanced by the use of a colposcope (Norseth et al., 2014), however a study from five Sub-Saharan countries reported in 2001 that colposcopes were not available in any of the district hospitals, and only in 6 % of provincial hospitals (Chirenje et al., 2001).

Microscopy

Tissue biopsies taken directly from lesions may be crushed and inspected immediately at the point of care (Kjetland et al., 1996). The presence of schistosome eggs in genital tissue represents the gold standard for a positive diagnosis of FGS (Kjetland et al., 2012; Poggensee et al., 2001). However, in the absence of visible lesions, a random biopsy from genital tissue cannot be considered adequate for a negative diagnosis, as eggs may easily be missed (Kjetland et al., 2012). Furthermore, a biopsy of genital tissue gives the woman an iatrogenic lesion, rendering her susceptible to STIs and HIV infection during the two weeks it takes to

heal, which is especially problematic since FGS and HIV are largely co-endemic (Kjetland et al., 2014, 2012).

The Pap smear programme is widespread to local clinic level in many African countries. However, the detection of schistosome eggs in cervical smears has a sensitivity for the diagnosis of FGS of less than 15 % (Feldmeier et al., 2001; Kjetland et al., 2012).

Microscopic detection of schistosome eggs in urine is considered to be the parasitological reference test for UGS, although this technique has been shown to lack in sensitivity when compared with new molecular techniques (Ibironke et al., 2012). Furthermore, presence of eggs in urine is not a definite indication of FGS (Kjetland et al., 2012) and the absence thereof is not adequate for a negative diagnosis (Kjetland et al., 2008; Poggensee et al., 2000).

Molecular and serological techniques

PCR-techniques may be used to detect schistosome DNA in urine and cervicovaginal lavage (CVL) samples (Kjetland et al., 2009). Detection in urine constitutes the same problem as urine microscopy, in that it doesn't provide information on genital pathology, and may therefore not be used to preclude FGS (Kjetland et al., 2012). In CVL, schistosome DNA was hypothesised to be a clear indication of FGS (Kjetland et al., 2009). However, a pilot study showed that the overall sensitivity was 53 %, increasing to 67 % in a subgroup of younger women, which suggests that chronic lesions with calcified / dead eggs may not contain DNA suitable for PCR-analysis (Kjetland et al., 2012).

Detection of circulating anodic antigen (CAA) in urine has proven to be a highly sensitive and specific indicator of having live *S. haematobium* worms (van Dam et

al., 2013) but the antigen involved circulates systemically before excretion in urine and gives no indication of the site of morbidity and can therefore not be used diagnostically for FGS (Kjetland et al., 2012). The same applies to circulating cathodic antigen (CCA), which can also be detected in urine, but which has been shown to have poor sensitivity for *S. haematobium* (Ayele et al., 2008).

Neopterin- and IgA-levels in CVL have not been found to be significantly higher in cases with FGS (Poggensee et al., 2001) and detection of circulating IgG-antibodies against schistosomes can neither distinguish between past and present disease nor indicate the site of morbidity (Doenhoff et al., 2004).

Diagnostics using computer image analysis

The World Health Organization has suggested the ASSURED criteria for point-of-care tests: Affordable, Sensitive, Specific, User-friendly, Rapid and robust, Equipment-free and Deliverable to end-users (Boppart and Richards-Kortum, 2014; Drain et al., 2014; Peeling et al., 2006). It is estimated that more than 60 % of the Sub-Saharan population has access to mobile phones (Aker and Mbiti, 2010). Furthermore, camera equipped cell phones are widely available in developing countries (Etzo and Collender, 2010), which can make point-of-care diagnostic algorithms based on computerised image analysis simple and cheap to implement.

Colour analysis

Computerised image analysis has been applied in the classification of precancerous cervical lesions using colour detection (Pogue et al., 2000). Automation of such algorithms has also been described (Greenspan et al., 2009).

Analyses of morphology

Fractals and fractal dimensions

A fractal is a self similar object whose details under magnification resemble the structure as a whole (Landini et al., 1995). In nature there are many naturally occurring fractals, such as snowflakes, rivers networks and blood vessels. A fractal's complexity can be evaluated by calculating its fractal dimension, which increases with increasing complexity (Landini et al., 1995). A single point has a fractal dimension of zero, a straight line has a fractal dimension of one and a plane has a fractal dimension of two. The fractal dimension is an expression of complexity as a ratio of change in detail to change in scale. E.g. a fractal dimension of one means that, as the scale with which the fractal is observed increases, the level of detail remains exactly the same. This is the case for a straight line: regardless of the scale, it will still remain a straight line. Analysis of local connected fractal dimensions (LCFD) allows for identification of areas with higher or lower fractal dimensions within a fractal (Landini et al., 1995).

In ophthalmology, it was possible to identify patients with occlusion of the retinal artery by analysis of the retinal vessels' LCFD (Landini et al., 1995). Another study found that patients with cerebral lacunar strokes had lower fractal dimensions of the retinal vessels compared to patients with minor cortical strokes (Doubal et al., 2010).

In gynaecology, fractal dimension and fractal lacunarity have been used to classify the uterine vessels seen in hysteroscopic images of patients with endometrial cancer and abnormal uterine bleeding (Vlachokosta et al., 2013).

Template matching

A template is a predefined shape made to resemble or match a structure to be identified. In ophthalmology, several methods have been described to identify the optical disc on retinal images by various template matching algorithms (Lalonde et al., 2001; Youssif et al., 2008).

Research questions and aim

In this thesis, I wanted to explore whether it is possible to identify the characteristic features of FGS by using computerised image analysis. Furthermore, I wanted to use computerised image analyses to objectively quantify and describe vascular characteristics of FGS. This work would contribute to increased understanding of the characteristic lesions of FGS as well as opening up the scene for an alternative diagnostic approach; using non-invasive image analysis, which would be cheaper, simpler, safer and more available in endemic areas.

Materials and methods

Study design and participants

The colposcopic images were acquired from a cohort of young women included in a study exploring female genital schistosomiasis in South Africa in 2011 - 2013. The study participants were recruited from rural schools north and south of Durban, KwaZulu-Natal. They were sexually active, non-pregnant young women, attending high school and aged 16 and above. Recruitment was not based on symptoms or test results.

Study area and population

KwaZulu-Natal is endemic for urogenital schistosomiasis (Hegertun et al., 2013). Frequent contact with fresh water bodies is common in the area, both for recreational purposes and chores such as doing laundry (Morgas et al., 2010).

According to the South African antenatal sentinel survey of 2011, 37 % of women attending antenatal care in KwaZulu-Natal were HIV positive. The most recent (2009 – 2010) survey on urogenital schistosomiasis in this district reported a 32 % prevalence of *S. haematobium* egg patent infection as diagnosed by microscopy among children in primary school (Hegertun et al., 2013). According to the South African census of 2007, almost a third of the households in this area did not have access to piped water in their community, 59 % of households used a pit latrine and 9 % did not have any toilet facilities.

S. mansoni is not endemic in the area (in a subsample of the main cohort, no *S. mansoni* eggs were found in 853 stool samples, analysed by the Kato-Katz method, unpublished data).

Ethical considerations

The study was granted permissions by four ethics committees and the regional health and education boards: the European Group on Ethics in Science and New Technologies in 2011 (Ref IRSES-2010:269245), the Biomedical Research Ethics Administration, University of KwaZulu-Natal in March 2007 (Ref BF029/07, annual renewal), the Department of Health, Pietermaritzburg, KwaZulu-Natal on February 3rd 2009 (Ref HRKM010-08) and the Norwegian ethics committee, REC South East on September 17th 2007 and renewed in 2011 (Ref 469-07066a1.2007.535).

All participants signed written, informed consent forms prior to investigation and were informed of the right to withdraw at any moment. All participants were tested for HIV, abnormal cytology and STIs. Pre- and post-test HIV counselling were provided in accordance with South African regulations. Test results were given unless they did not want to know or did not attend the follow-ups. Counselling, referral and treatment were given according to South African guidelines. Anti retroviral therapy is free of charge to South African citizens. All colposcopic images are non-identifiable in that they only depict the uterine cervix.

Access to and storage of data

Raw data was encrypted, password protected and pseudonymised on two parallel non-WAN connected computer servers with AES WPA2-PSK protected wifi. The machines were under restricted access and technical control. In addition to being

held on the secure the storage server, encrypted data was backed up daily on two network attached storage devices. Paper was kept under lock and key at all times. It was only accessible to the project management, the research officer and, during the data entry period: the data entry staff. Unforeseen data (e.g. incidental findings) was referred to a clinician, nurse, social worker or psychologist. There was an encrypted track record of the referral on the encrypted servers following the same backup scheme. Any data leaving the research station, was pseudonymised, truncated and, encrypted. Data was accessed using a 128 bits encrypted https line to the server.

Clinical investigation and questionnaire

The women in the study were examined using an autoclaved metal speculum and an Olympus OCS 500 colposcope with a mounted Olympus E420 10 megapixel (Mpx) single lens reflex camera (SLR) or a Leisegang colposcope with a Canon EOS 40D 10 Mpx SLR. The image files were stored using high quality JPEG compression (the most common file format used in digital cameras). Lesion size, appearance and location were recorded by digital sketching on a simplified model of the cervix (Figure 6).

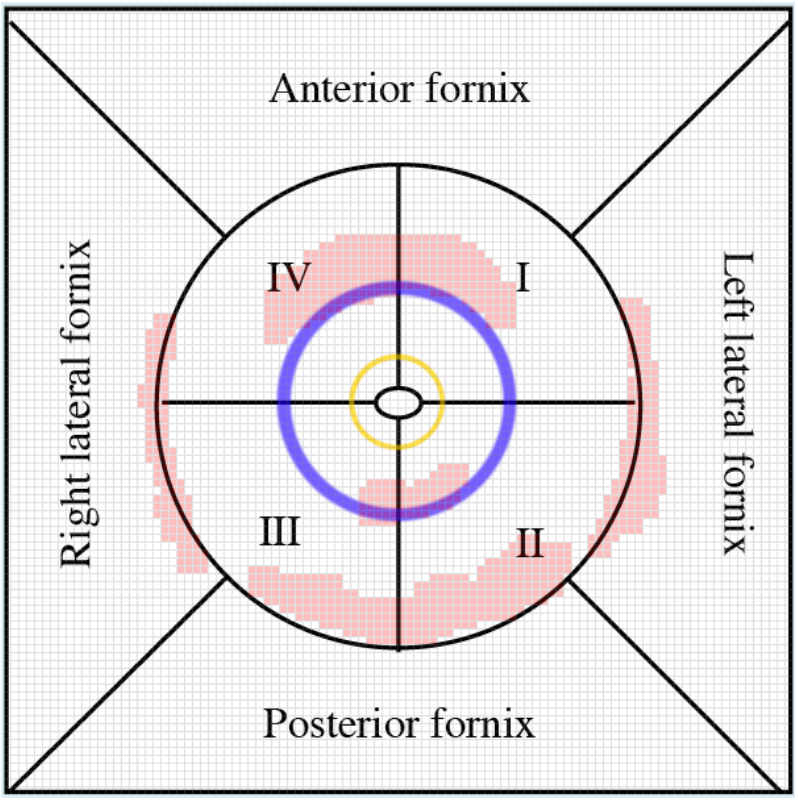

Figure 6. Digital schematic representing the ectocervix and the fornices. The clinician inserts the inner and outer delimitation of the transformation zone (yellow and blue, respectively) and colours the areas (pink) affected by lesions for future reference.

Trained research assistants interviewed all participants in the local language (isiZulu). The questionnaire included questions concerning sexual behaviour, pregnancy, contraception and urogenital symptoms.

Laboratory analyses

A single urine sample was collected between 10 a.m. and 2 p.m. on the day of the clinical examination. Two slides were made from every sample and two independent technicians read each of the slides by light microscopy (Morgas et al., 2010).

Pap smears were prepared on glass slides. An ectocervical sample was acquired using a wooden spatula and an endocervical sample using a cytobrush. The slides

were spray-fixed and read later by light microscopy. Cell atypia was classified according to the Bethesda System as atypical squamous cells of unknown significance (ASCUS), low- and high-grade squamous cell intraepithelial lesions (LSIL and HSIL). Number of neutrophils and degree of inflammation were graded as none, mild, moderate or marked.

Cervico-vaginal lavage (CVL) was collected by spraying 10 mL of saline four times onto the ectocervical mucosa followed by withdrawal back into a syringe. CVL was analysed for *Trichomonas vaginalis* using an in-house real-time PCR technique (Laboratory of Infection, Prevention and Control, University of KwaZulu-Natal, Durban, South Africa). *Chlamydia trachomatis* and *Neisseria gonorrhoea* were analysed using strand displacement assay on a ProbeTec machine (Becton, Dickinson and Company [BD], Franklin Lakes, New Jersey, USA). *Treponema pallidum* was analysed using rapid plasma reagin on serum (Macro Vue 110, BD, Franklin Lakes, New Jersey, USA). *Herpes simplex* type 2 antibodies were detected in serum using ELISA (Ridascreen HSV 2 IgG, Davies Diagnostics, Randburg, South Africa). HIV was detected in serum using a rapid antibody test (Core One Step HIV 1/2 test kit, Kendon laboratories, Durban, South Africa).

Statistical analyses

Statistical analyses and graphs were produced using IBM SPSS Statistics Version 19 (IBM Company, Chicago, USA), STATA (StataCorp, College Station, Texas, USA) and the LCA STATA plugin (The Methodology Center, Penn State, Pennsylvania, USA).

All analyses were presented with statistical significance defined within a 95 % interval of confidence.

Comparisons between groups

Comparison of means was done using the Student's T-test if data were normally distributed. Otherwise, distributions were compared using the Mann-Whitney-U test. If comparing the means of more than three groups, analysis of variance (ANOVA) was used if the data was normally distributed. Otherwise the Kruskal-Wallis analysis of variance was used.

Comparison of frequency between groups was done using the Chi-square test, unless the expected number of observations in one sub-group was less than five, in which case Fisher's exact test was used.

Analyses of predictive accuracy

Receiver operating characteristics (ROC) curves were made to evaluate the ability of a continuous variable to predict a dichotomous variable by varying the cut-off value (Metz, 1978). The area under the curve (AUC) was calculated as an indication of the overall predictive value (Hanley and McNeil, 1982). The optimal cut-off value was determined by identifying the point on the ROC curve closest to the upper left corner and this value was then used to calculate the corresponding sensitivity and specificity.

Regression models

Multivariable logistic regression models were constructed using an inclusion threshold corresponding to an 85 % interval of confidence ($p < 0.15$). Variables were eliminated from the model one by one (backwards elimination) based on a minimum significance level of 0.15 and only if the model's likelihood ratio did not change significantly within a confidence interval of 95 % ($p < 0.05$).

Latent class analyses

Latent class analysis (LCA) enables the analysis of multiple observed dichotomous variables that imperfectly measure a non-observable (latent) variable (Formann and Kohlmann, 1996; Goodman, 1974). The latent variable may consist of two or more classes. The LCA estimates the probability of having a positive response in each of the observed variables conditional on class adherence. E.g. a simple LCA analysis on a condition with two possible states (healthy, not healthy) as determined by two tests that each give an imperfect measure of the condition could look like this:

	Class 1: Healthy	Class 2: Not healthy
Test 1	**0.9**	0.1
Test 2	**0.7**	0.3

We can read from the table that persons allocated to class 1 (healthy) has 90 and 70 % probability of testing positive on Test 1 and 2, respectively.

The general procedure of LCA is to generate competing models with different number of classes and determine how well they fit with the data, as evaluated by parsimony, maximum likelihood and the Akaike and Bayesian information criterions. A good model will be able to "classify" cases in the dataset into a latent class with relatively high probability of class adherence, thus yielding lower values of the information criterions than competing models. Furthermore, simpler models are preferred to complex models if they perform equally well. This is termed the principle of parsimony (Collins and Lanza, 2013).

Latent class statistics were applied in paper II, using urine microscopy for schistosome ova, clinical investigation for sandy patches, schistosoma PCR of urine and CVL, school prevalence by urine microscopy and the output from the computer

analysis to model a latent variable to describe the FGS diagnosis. This latent variable was then used as a surrogate gold standard to evaluate the computerised diagnosis.

Computerised image analyses

Selection of image material

We manually searched the documentation database for images that fulfilled the following criteria: the cervix should be in the field of view, the sandy patch (if present) should be located on the cervical mucosa and visible in the image (but not necessarily the central element). There should be no foreign material in the field of view (swab, spatula, acetic acid etc.) and the exposure and focus should be adequate for visualisation of the lesion, though not necessarily perfect.

Software framework

All image processing and analyses were performed using ImageJ (open source, free software from National Institutes of Health, US). Plugins for ImageJ were written in Java (Oracle Corporation, Redwood Shores, US) to perform the specific analyses. Scripts (macros) were written to automate the loading of batches of images and executing the analyses and recording the results. Numeric results were recorded in text files and intermediate images were generated to allow for visual verification of each stage of the analyses (Figure 7).

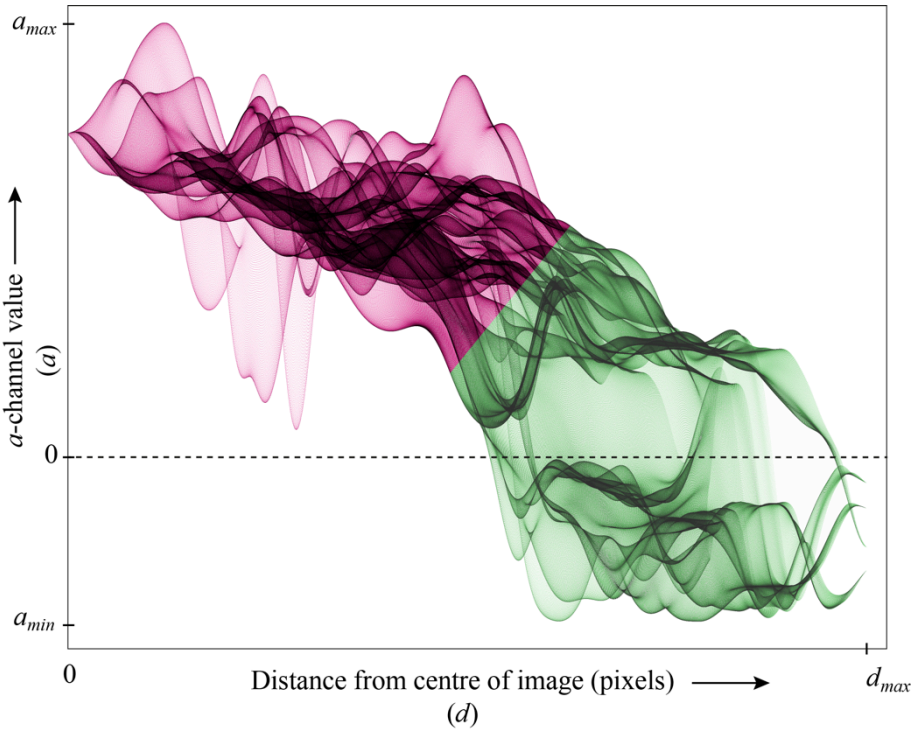

Figure 7. An intermediate graphic generated by one of the plugins written for ImageJ, showing the result of k-means clustering (magenta and green) as part of the process of determining the region of interest (the cervical surface). The abscissa shows the distance from the centre of the original image and the ordinate shows increasing values of the a-channel from the Lab-colour space, representing increasing values of magenta (pink).

Identifying the region of interest

Colposcopic images often contain non-cervical elements such as parts of the speculum, medical instruments and skin. To perform a reliable image analysis, it is therefore important to first identify the region of interest (ROI). We developed a method of automated detection of the ROI, drawing from a previously described method (Greenspan et al., 2009). The cervix is characterised by colours in the tones of magenta and it usually occupies the centre of the image. The colour channel *a* of the *Lab* colour space, represents increasing intensity of magenta for increasing positive values. The ectocervical mucosa can therefore be characterised by high

values of *a* in combination with a short distance from the centre of the image. Figure 7 shows how pixels were clustered in two groups: those that were likely and unlikely to be part of the cervical surface.

This procedure identified the central area of the cervix and the mask needed to be refined to follow the cervical curvature precisely. The mask was therefore expanded by reading pixel values in the Value channel of the HSV colour space in concentric rays emanating from the initial coarse mask. The Value channel represents darkness/lightness, and due to the cervix' convex shape, has lower values around the perimeter where light falls on a tangential angle to the cervical surface. A ray's end point was hence defined when a reduction in mean value of 20 % or more for two consecutive runs of 50 pixels was found (i.e. 50 pixels were 20 % darker than the former 50 pixels, Figure 8).

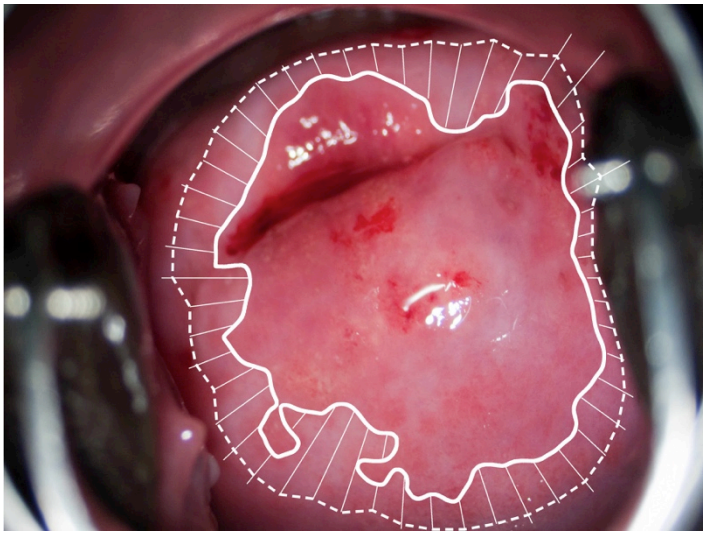

Figure 8. The region of interest was refined (dashed line) by expanding the initial coarse mask (solid line) by 50 concentric rays until a drop in pixel values of the Value channel (HSV colour space). The rays were finally smoothing using a 5-point mean filter to ensure a smooth delimitation: notice the three rays protruding outside the dashed line in the upper left corner.

Specular reflections

A convex, reflective surface will appear with specular reflections regardless of the angle of inspection. This is the case of a moist cervix illuminated by a colposcope. The specular reflections appear as bright white spots on the images and may perturb the image analyses in a number of ways: The colour values in these areas represent local maxima (over-exposed areas) that may skew calculations on pixel values in an area. Furthermore, the perimeter pixels of specular reflections may appear yellow (Figure 9), which may be a problem if the purpose of the image analysis in question is to detect yellow pixels. It is therefore necessary to exclude these areas from the image analyses, e.g. by masking them.

Figure 9. Perimeter gradients around specular reflections may appear yellow.

The specular reflections have very low values in the Saturation channel of the HSV colour space. This feature allowed us to mask reflections by filtering pixels

with values 1.5 standard deviations or more below the mean value in the Saturation channel. To ensure complete coverage of perimeter gradients, the masks were expanded radially by 20 pixels (Figure 10).

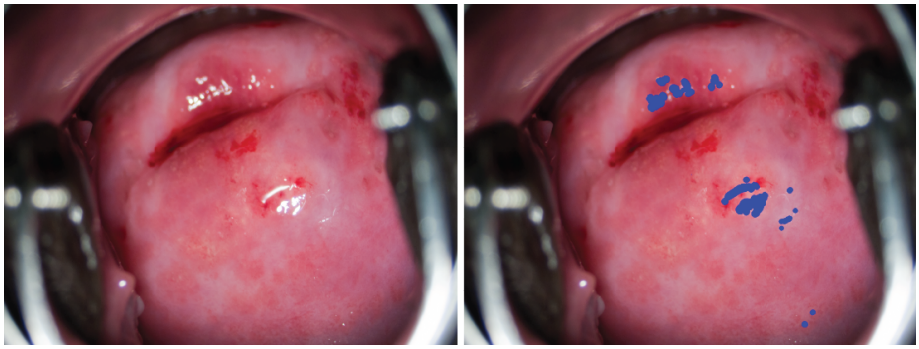

Figure 10. Left: the original image. Right: specular reflections have been masked (blue) by filtering on low saturation.

Colour analyses of sandy patches

Grainy sandy patches (Figure 2) and homogenous sandy patches (Figure 3) have the characteristic quality of appearing yellow on visual inspection. This may be used in colour image analysis to identify lesions based on degree of "yellowness". However, there are a few challenges to this problem:

What is "yellow"?

There are a number of different colour spaces (ways to represent colour) and only one of them has a "yellow-channel" (CMYK colour space). However, this is a colour space primarily intended for printing and the yellow-colour represented by this channel is unlikely to conform to the subtle variations of yellow seen in sandy patches. It was therefore necessary to evaluate a number of colour spaces in order to identify the colour channels that were most likely to be able to measure the yellow seen in sandy patches. This was done by measuring mean intensity in each colour channel ($\Delta_{\bar{c}}$) of the colour spaces RGB, HSV and Lab (a total of 9 channels) in pixels

that were verified by the clinicians as certain to be part of sandy patches and selecting those with the strongest ability to discriminate these from the surrounding mucosa (E1).

$$\Delta_{\bar{c}} = \left|\bar{c}_{sandy\ patch} - \bar{c}_{mucosa}\right| \quad \text{(E1)}$$

This allowed us to choose seven colour channels that were useful in discriminating sandy patches from surrounding mucosa; Red, Green, Blue, Hue, Saturation, Value and b.

Every image is different

The colours change from image to image due to different cameras, exposures, colour balances etc. Therefore, "yellow" is not a specific set of colour values in different colour channels. It is a relative entity that must be analysed in comparison to the surrounding mucosa. By measuring the mean mucosal colour value ($\bar{x}$), we could define the value windows that were most likely to contain the sandy patches for each colour channel (Table 1).

Channel	Lower threshold	Upper threshold
Red	$\bar{x} + \Delta_{\bar{c}} - 2\sigma\Delta_{\bar{c}}$	$\bar{x} + \Delta_{\bar{c}} + 2\sigma\Delta_{\bar{c}}$
Green	$\bar{x} + \Delta_{\bar{c}} - 1.5\sigma\Delta_{\bar{c}}$	$\bar{x} + \Delta_{\bar{c}} + 2\sigma\Delta_{\bar{c}}$
Blue	x_{min}	$\bar{x} + \Delta_{\bar{c}} + 2\sigma\Delta_{\bar{c}}$
Hue	$\bar{x} + \Delta_{\bar{c}} - 0.625\sigma\Delta_{\bar{c}}$	$\bar{x} + \Delta_{\bar{c}} + 4\sigma\Delta_{\bar{c}}$
Saturation	$\bar{x} + \Delta_{\bar{c}} - 2\sigma\Delta_{\bar{c}}$	$\bar{x} + \Delta_{\bar{c}} + 2\sigma\Delta_{\bar{c}}$
Value	$\bar{x} + \Delta_{\bar{c}} - 2\sigma\Delta_{\bar{c}}$	$\bar{x} + \Delta_{\bar{c}} + 2\sigma\Delta_{\bar{c}}$
b	$\bar{x} + \Delta_{\bar{c}} - \sigma\Delta_{\bar{c}}$	x_{max}

Table 1. The mean colour value ($\bar{x}$) is measured within the region of interest. The predicted colour value of a sandy patch corresponds to the addition of the previously calculated difference in colour values between sandy patches and healthy mucosa ($\Delta_{\bar{c}}$). The upper and lower values of the range are calculated by addition and subtraction, respectively, of various multiples of its standard deviation ($\sigma\Delta_{\bar{c}}$).

Since the calculated window depends on the measured mean mucosal value in each image, this constitutes an adaptive approach that will compensate for varying exposure and colour balance from image to image (Figure 11).

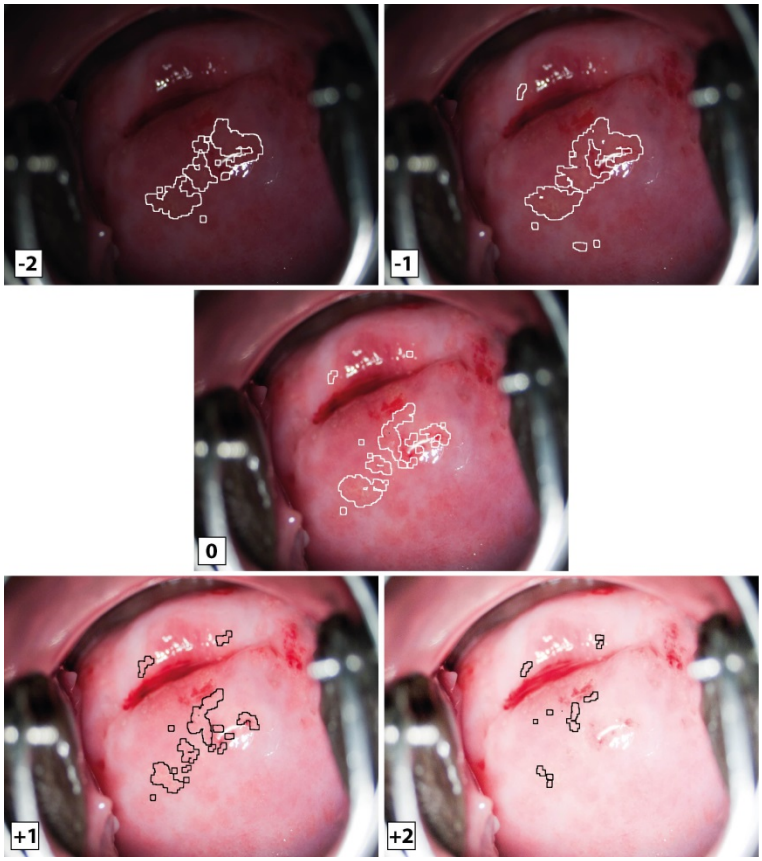

Figure 11. The colour detection algorithm was applied adaptively to perform with a wide range of exposure settings. The delimited areas represent the yellow sandy patches as detected by the algorithm.

Abnormal blood vessels

Isolating the blood vessels

The abnormal blood vessels typical of FGS (Figure 4) have been described as having specific morphology on visual inspection. In order to do morphological analyses on the cervical blood vessels, the blood vessel structures had to be isolated. It is well known from photocolposcopy that the application of a green filter on the light source may render blood vessels easier to see because they appear almost black as they reflect very little green light (Mayeaux and Cox, 2011). Likewise, the vessels

appear very dark in the Green channel in a regular RGB colour image acquired in normal (white) light. Furthermore, the blood vessels appear in a strong red colour, representing high values in the Saturation channel of the HSV colour space. It was therefore possible to use two different colour channels for the isolation of cervical blood vessels.

However, the ectocervix has a convex shape, reflecting the light unevenly across the surface. It was therefore necessary to equalize the colour channels before analysis. This was done by generating an inverted image for each channel, which was subsequently smoothed using a Gaussian blur filter and then calculating the sum (addition) with the original image. The inverted Green channel could then be multiplied with the saturation channel to produce an image with "boosted" values of the pixels with low values of green and high values of saturation (Figure 12). This image is henceforth referred to as the boosted image.

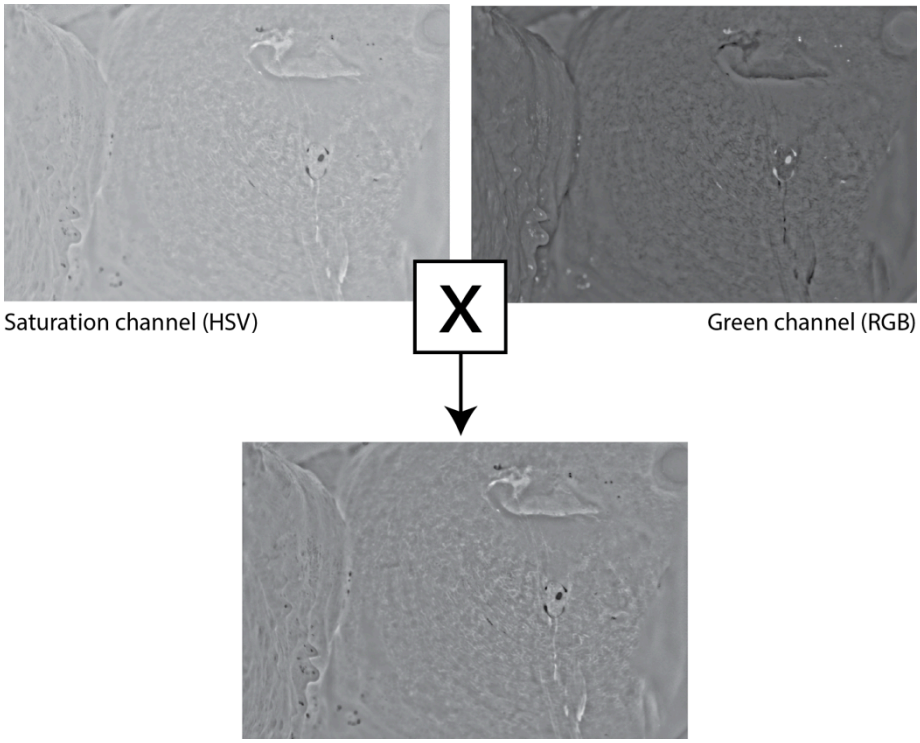

Figure 12. The multiplication of the inverted green channel and the saturation channel results in an image where the blood vessels have "boosted" values, appearing more clearly than in either of the two original colour channels.

Thresholding

The product obtained from the multiplication of the inverted Green channel and the Saturation channel is still not suitable for fractal analyses. The blood vessels need to be represented in a binary image with only two colour values; black and white (0 and 1). This can be achieved by defining a threshold value, T, of grey-level (E2). This is known as thresholding.

$$pixel = (pixel \geq T)\ ?\ object : background \quad \text{(E2)}$$

However, determining an optimal threshold value that works in a large number of images obtained from different cameras under different conditions is not trivial. It is therefore more useful to calculate a different threshold per image, such as using the

mean grey value (Figure 13D). Although we equalized the colour channels used to generate the boosted image, there will always be local variations in illumination and contrast within an image. Therefore, we applied the local adaptive thresholding Niblack method (Niblack, 1986) to ensure local optimal thresholding of each pixel (Figure 13E). It is applied per pixel using the mean (μ) and standard deviation (σ) calculated within a window surrounding each pixel (E3).

$$pixel = (pixel \geq \mu + k * \sigma)\ ?\ object : background \quad \text{(E3)}$$

The k-value (E3) represents a constant of which the sign (positive or negative) determines whether to keep dark or light objects and the magnitude determines how "strict" the thresholding should be. This can be set to a default value of -0.2 or 0.2 depending on whether the background is dark or light (Sezgin and Sankur, 2004). Otherwise it can be found empirically by trial and error (Chaki et al., 2014). Neither of these approaches will yield an optimal k-value for each image. However, it is possible to determine the optimal k-value for each individual image if the distributions of foreground and background pixels are known (Albregtsen, F., 2009, Selecting an optimal parameter value for Nieblack's adaptive thresholding method, personal communication). These can be estimated by expectation-maximization (EM), which is an iterative algorithm to find distributions with the highest likelihood (Redner and Walker, 1982). The k-value can then be expressed as:

$$k = \frac{d\left(\frac{1}{2} - F\right) - \frac{1}{d} ln\left(\frac{F}{1-F}\right)}{\sqrt{1 - d^2\ (F-1)F}} \quad \text{(E4)}$$

Where d is the number of standard deviations (σ) separating the mean background and foreground values and F is the probability of the foreground distribution.

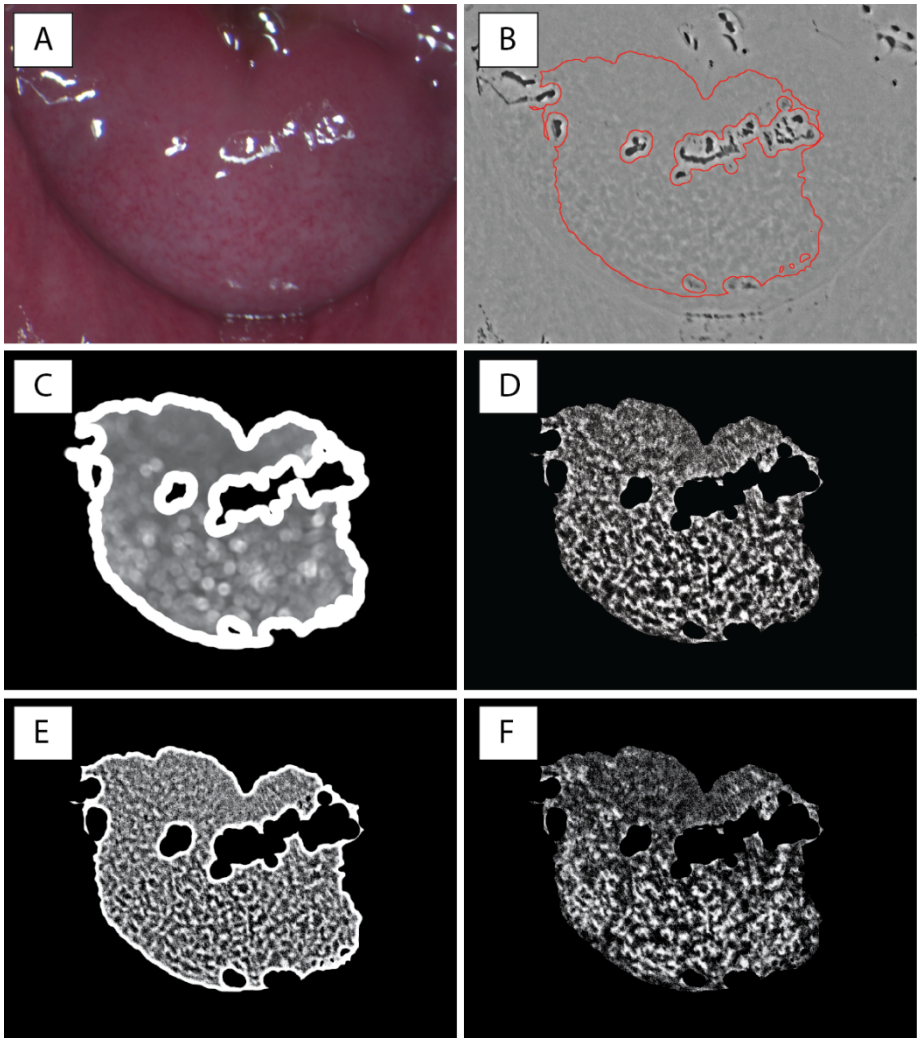

Figure 13. **A.** Original colour image. **B.** Equalized product of the inverted Green channel and the Saturation channel. The automatically detected region of interest (ROI) is indicated in red. **C.** The standard deviation of grey values in a 50px window calculated for each pixel and represented as relative levels of grey values (0-255). **D.** Thresholding performed using the mean grey value. **E.** Thresholding using the standard Niblack method with a k-value of -0.2. Notice the perimeter artefact as a thick white line around the ROI. **F.** Thresholding using Niblack's method with calculation of optimal k-value per 50px sliding window and elimination of the perimeter effect.

Local adaptive thresholding techniques have a tendency to generate noise in areas with low contrast. This has been solved in a number of ways, such as Bernsen's method (Sezgin and Sankur, 2004) of setting a lower threshold value (lc_t) of local

contrast (lc) below which the threshold value will default to 128 (E5). Local contrast (lc) is defined as the difference between maximum and minimum grey values within a sliding window.

$$if\ (lc < lc_t)\ pixel = (midgrey \geq 128)\ ?\ object : background$$
$$else\ pixel = (pixel \geq midgrey)\ ?\ object : background \quad \text{(E5)}$$

We used a similar approach, but instead of evaluating the local contrast ad modum Bernsen, we evaluated the local standard deviation (σ) of grey values in reference to the mean standard deviation ($\bar{\sigma}$) within the ROI as a more robust measure of contrast (E6). If the contrast was low, Niblack's method was not applied and the threshold value was set to the estimated mean foreground value (μ_F).

$$if\ (\sigma < \bar{\sigma})\ pixel = (pixel \geq \mu_F)\ ?\ object : background$$
$$else\ pixel = (pixel \geq \mu + k * \sigma)\ ?\ object : background \quad \text{(E6)}$$

Furthermore, the Niblack method has an inherent flaw when working within a ROI. If the standard deviation within a sliding window is very high, as it will be close to the perimeter of the ROI (Figure 13C), the resulting threshold value will be either very low or very high, producing a perimeter artefact (Figure 13E). This can be avoided by extending the abovementioned procedure to include windows with very high contrast; exceeding the 97.5th percentile of contrast (E7). The result is elimination of the perimeter artefact (Figure 13F).

$$if\ (\sigma < \bar{\sigma}\ ||\ \sigma > \bar{\sigma} + 1.96 * \bar{\sigma}_\sigma)\ pixel = (pixel \geq \mu_F)\ ?\ object : background$$
$$else\ pixel = (pixel \geq \mu + k * \sigma)\ ?\ object : background \quad \text{(E7)}$$

Reticular, branched and convoluted

Blood vessels may be considered fractals; objects whose details under magnification resemble the structure as a whole (Landini et al., 1995). Increased branching, reticularity and convolution are all factors representing increased

morphological complexity. The complexity of fractals may be estimated by calculating the fractal dimension, either of the entire fractal or locally as local connected fractal dimensions (LCFDs) (Landini et al., 1995).

To estimate the blood vessels' fractal dimensions we used a simple method that consists of estimating the above-mentioned scaling phenomenon by counting the number of rectangular boxes required to cover the fractal (Smith et al., 1996). By gradually decreasing the box size, more boxes are required to cover the fractal. However, if the fractal has a high degree of complexity, the number of boxes required will increase more with scaling than it would for a fractal of less complexity, thus representing a higher fractal dimension, D (Figure 14). This can be expressed as the log-log relationship between the numbers of boxes, N, divided by the inverse box size, ε (E8).

$$D = \lim_{\varepsilon \to 0} \frac{log\ N_{\varepsilon}}{log(1/\varepsilon)} \quad \text{(E8)}$$

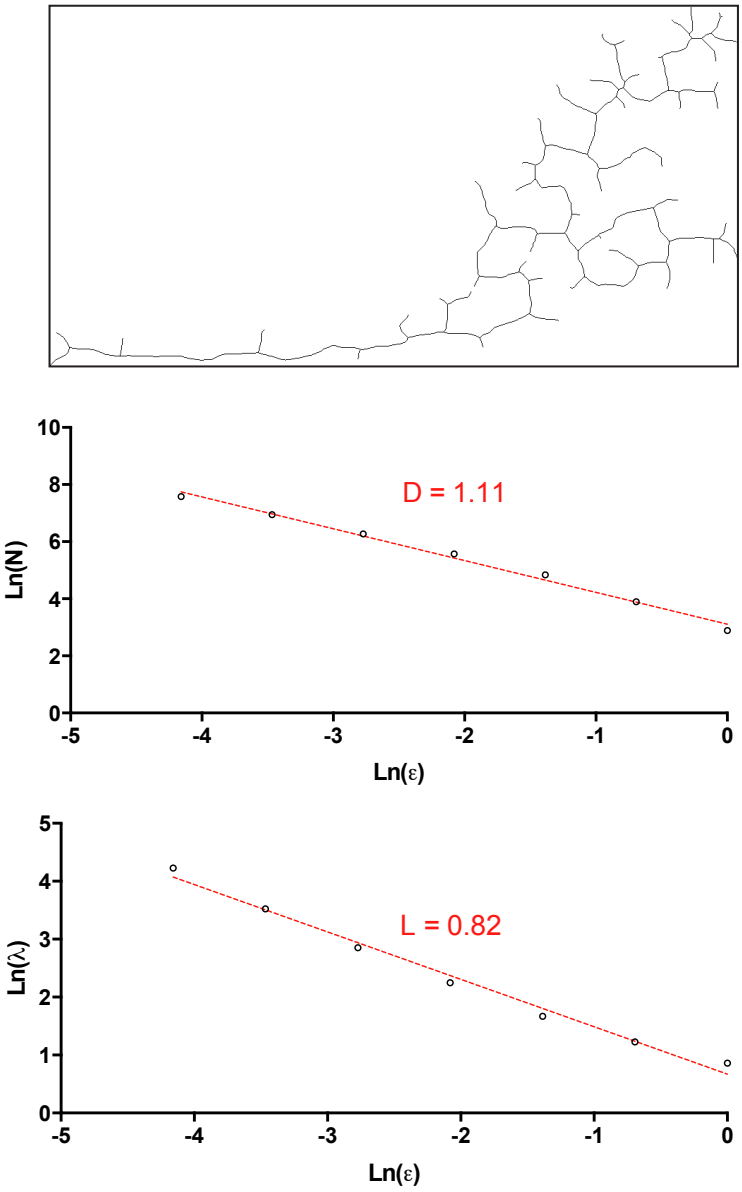

Figure 14. **Top**: A binary blood vessel structure after image processing. **Middle**: A log-log plot showing the number of boxes (N) required to cover the blood vessel in relation to box size (ε). The slope of the dashed red regression line is 1.11, corresponding to the estimated fractal dimension, D. **Bottom:** A log-log plot showing the mean λ-value in relation to decreasing box size (ε). The slope of the dashed red regression line is 0.82, corresponding to the estimated lacunarity of the blood vessel.

Lacunarity is another characteristic of fractals that describes their complexity in terms of heterogeneity. A fractal with many gaps and spaces is said to have high

lacunarity. This can also be estimated by using the box counting algorithm (Smith et al., 1996). However, for lacunarity, the interesting measure, denoted λ, is the variation in number of pixels present from box to box. This can be expressed as the standard deviation (σ) divided by the mean (μ) number of pixels per box (E9).

$$\lambda = \left(\frac{\sigma}{\mu}\right)^2 \quad \text{(E9)}$$

Using a sliding window, it is therefore possible to calculate a λ-value per pixel for each of the box sizes (ε). The lacunarity of the fractal can be expressed as the log-log relationship between the mean λ-value and box size, ε (Figure 14).

Circular features

The abnormal blood vessels typical of FGS have been described as having circular features (Kjetland et al., 2012, 2005; Norseth et al., 2014). The vascular structures in the boosted image (Figure 12) were isolated by removing all pixels below the mean grey level (Figure 15B). A circular template was generated to resemble the characteristic circular shape of the vessels as indicated by the clinicians (Figure 15C). Template matching is a very processor-intensive process if applied as a pixel-by-pixel approach in the image domain. With mobile applications in developing countries in mind, we sought to find a less processor intensive method that performed adequately. Multiplication in the frequency domain corresponds to convolution in the image domain but multiplication is a much simpler operation. Therefore, template matching was performed by converting the template and the processed image to the frequency domain by the fast Fourier transform, multiplying them and finally converting the result back to the image domain (Figure 15D). In the resulting image, areas with circular configurations have higher intensity due to the intersection of convoluted circles. A threshold level was set to remove pixels below

the 97.5th percentile of intensity. The final image contains clustered pixels representing the centres of the matched circular structures (Figure 15E). The number of circles identified was recorded for each image.

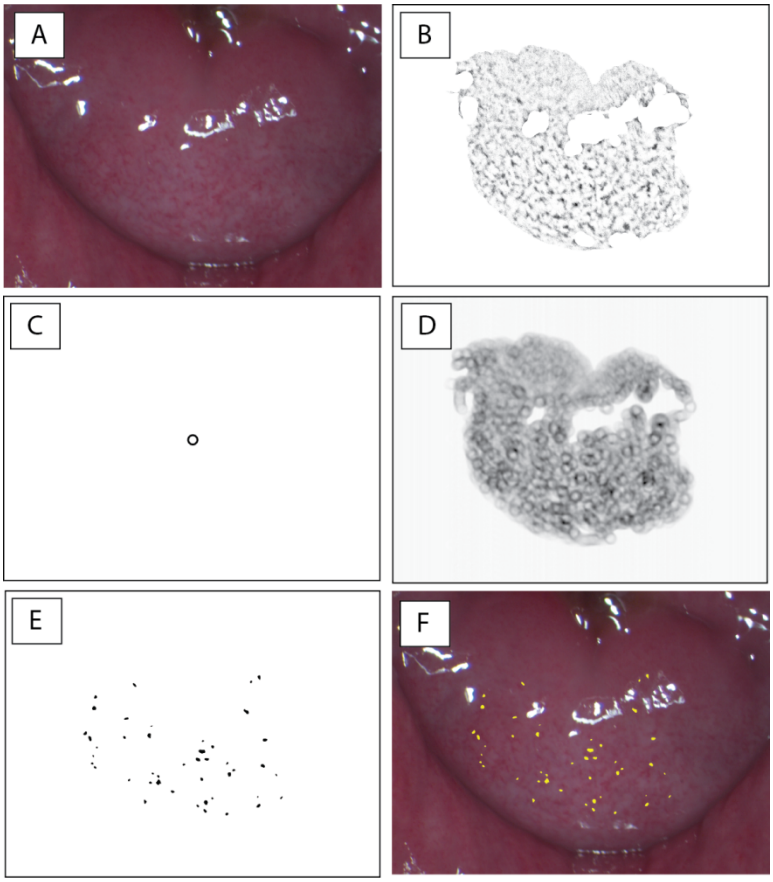

Figure 15. **A.** The original colour image. **B.** After extracting the region of interest, the boosted image is thresholded by using the mean grey value. This leaves primarily blood vessel structures in the image. **C.** The circular convolution template. **D.** The result of the convolution of B and C (multiplication in the frequency domain). **E.** The result of thresholding image D on the 97.5th percentile of pixel value. Each dot represents a positive match for a circular feature. **F**. The result of the template matching (E) superimposed on the original colour image (A).

Combining colour and morphological analyses

Although sandy patches and abnormal vessels might represent different aspects of the pathology in FGS and the one might be present without the other, combining

the two analyses could be diagnostically useful. However, it is important not to include the same images that have already been used in the development of the respective methods. This removes 240 otherwise eligible images from the analyses, of which 110 had been diagnosed with FGS clinically. This selection bias is difficult to adjust for, since it will dramatically reduce the prevalence of FGS in the remaining image material, thus also reducing the estimated sensitivity and artificially boosting the specificity.

Results

Colour analyses

Developing the method

The first results from development of the method (Paper I) showed that the colourimetric analysis might be able to identify sandy patches with an accuracy ranging from 60 - 87 %, depending on image quality (Table 2).

Quality	AUC[a]	Sensitivity	Specificity
10 Mpx, no noise added	0.77	83 %	73 %
5 Mpx, no noise added	0.78	87 %	73 %
5 Mpx, noise added (10 SD)[b]	0.76	80 %	70 %
5 Mpx, noise added (15 SD)[b]	0.67	60 %	70 %

a Area under receiver operating characteristics curve

b Noise was generated using a Gaussian distribution with 10 or 15 standard deviations.

Table 2. The first results from the development of the method, validated on a small set of 60 images, modified to simulate various image qualities.

Validating the method

The colourimetric diagnosis was validated using a dataset consisting of 1074 women visiting the clinic (Paper II). The clinicians examined 821 women gynaecologically and indicated the presence of sandy patches in 26 % (n = 212). Sandy patches appearing as grains were described in 112 (14 %) participants and sandy patches appearing as homogenous, yellow areas were described in 128 (16 %). Presence of both lesion types was described in 28 (3 %). Digital images were

acquired in 781 of these women and of these, 683 cases had images that were found adequate for image analysis (Figure 16).

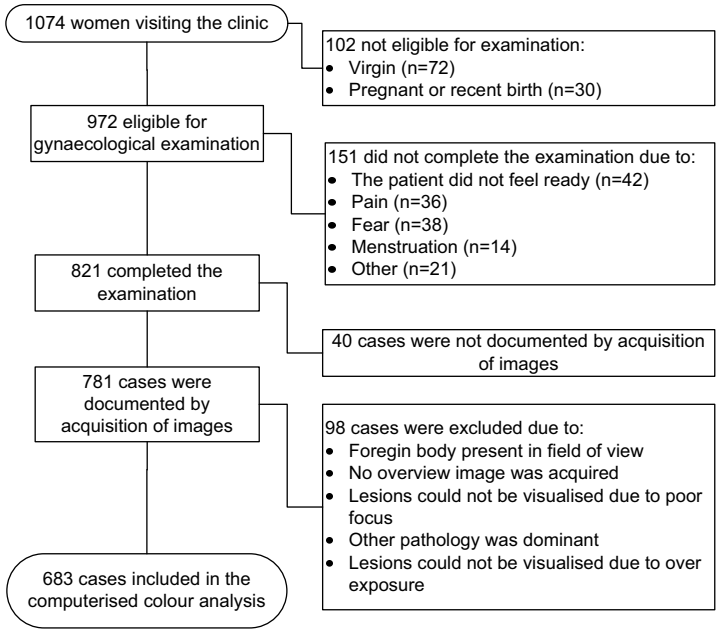

Figure 16. Flow chart describing the inclusion of images for validation of the colourimetric image analysis.

In the absence of a gold standard to use for validation of the method, we fitted a latent statistics model dividing the population into three classes. Table 3 shows a contingency table with the probabilities of having a positive finding for an observed variable, conditional on each of the three latent classes.

	Probability of having a positive variable conditional on class adherence		
Observed variables	Schistosomiasis negative	Urinary schistosomiasis	FGS positive with eggs in urine or lavage
Computer colour analysis	0.38	0.00	**0.70**
Schistosoma PCR in CVL	0.02	0.27	0.40
Schistosoma PCR in urine	0.04	**0.89**	**0.88**
Urine microscopy	0.01	**1.00**	**0.74**
Clinical finding of sandy patch	0.09	0.00	0.46
School prevalence 0-9 %	0.17	0.00	0.06
School prevalence 10-19 %	0.36	0.18	0.17
School prevalence 20-29 %	0.34	**0.55**	0.35
School prevalence ≥ 30 %	0.14	0.27	0.43

Probabilities exceeding 0.5 are indicated in bold.

Table 3. Latent class analysis was used to classify the population (n = 1074) in three classes based on all available information.

As shown in Table 3, the first class (1), which represented 80 % of the population, was characterised by having none of the schistosomiasis-related findings. This class was labelled "Schistosomiasis negative". (2) Individuals in the second class were found to have high risk for positive urine microscopy and PCR for presence of schistosome ova as well as attending a school with medium-high or high prevalence of schistosomiasis (combined risk of 82 %). However, they had very low risk of positive findings by clinical investigation and computer colour analysis. This

class, which represented 6 % of the population, was labelled "Urinary schistosomiasis without FGS". (3) The final, third class, which represented 14 % of the population, was characterised by individuals who were likely to have positive findings in urine (both microscopy and PCR) as well as by computer analysis. In addition, they had a 46 % chance of having a positive finding by clinical investigation and a 40 % chance of having a positive PCR result in CVL. Furthermore, the combined risk of attending a medium-high or high prevalence school was 78 %. This class was labelled "FGS positive with eggs in urine or lavage". This final class was used as a surrogate gold standard in order to estimate the predictive quality of the computer colour analysis. Sensitivity was calculated to 80.5 % and specificity 66.2 % using the optimal cut-off level for defining a positive diagnosis, as determined by using a ROC curve (AUC = 0.76, 95 % CI = 0.71 – 0.81, $P < 0.001$).

Morphological analyses

The morphological analyses (Paper III) showed that circular template matches, distance between vessels and LCFD were most useful in classifying the images. These variables were used to fit a regression model, constituting the diagnostic predictor (Table 4). We showed that abnormal blood vessels typical of FGS could be detected with an accuracy of 78 %.

	Normal cervix	Blood vessels typical of FGS	Other cervical pathology
Negative	42 (84 %)	11 (22 %)	38 (76 %)
Positive	8 (16 %)	39 (78 %)	12 (24 %)
Total	50	50	50

Table 4. Sensitivity and specificity of the regression model consisting of circular template matches, mean distance between vessels and mean LCFD.

Discussion

In this thesis, I have presented new methods for diagnosing female genital schistosomiasis (FGS), relying on computerised image analysis of sandy patches and abnormal blood vessels typical of FGS. I've shown that the accuracy of these methods may range from 60-80 %, which is promising as a concept but underlines the requirement for further research to increase accuracy, mainly in terms of specificity.

The need for a new diagnostic method in FGS

Diagnosing FGS is an unresolved problem for which a safe, reliable and affordable solution is long overdue (Kjetland et al., 2012). The diagnostic challenges lie partly in the nature of the lesions. The calcified (dead) eggs may persist in genital tissues long after the live parasites have been eradicated. Therefore, techniques aimed at detecting egg excretion are of little help in diagnosing lesions. However, the inflammatory processes surrounding the calcified eggs may still persist and it has been shown that genital mucosa with calcified eggs have a higher density of $CD4^+$ T-lymphocytes, creating a setting which may increase the risk of HIV infection (Jourdan et al., 2011a). Furthermore, these lesions have been found to persist after both standard and high-dose treatment with praziquantel (Kjetland et al., 2005).

Clinical diagnosis of schistosomiasis by visual inspection using a colposcope has been deemed a safe and acceptable approach in endemic areas (Kjetland et al., 2012). However, a study performed in five sub-Saharan countries in 2001, showed that colposcopes were not available in any of the district hospitals, and in only 6 % of provincial hospitals (Chirenje et al., 2001). Furthermore, colposcopic equipment is

expensive with the cheapest models costing around USD 2,000 and it requires a high level of medical training to use. It is therefore necessary to find alternative means of diagnosing FGS that do not require a colposcope or a medical expert at the point of care. The methods that I propose in this thesis, using computer colour analysis, represents an alternative pathway to develop a tool for safe and objective assessment of the lesions as an aid in a clinical diagnosis, possibly in combination with image analysis for cancer diagnosis (Dickman et al., 2001; Ji et al., 2000).

Why computerised image analysis?

A computerised diagnostic method is ideal for use in remote areas with limited access to electricity and infrastructure as it may run on battery powered mobile devices. It has been shown that images acquired with a cell phone camera were adequate in visualising the cervix for evaluation of pre-cancerous, gynaecological lesions (Quinley et al., 2011). By using simple electronic devices such as hand held cameras, mobile phone cameras or other, there will be no need for expensive investments in new equipment. There are a number of cancer screening programs in Sub-Saharan Africa where small handheld cameras are currently used for documenting findings after application of acetic acid (VIA) instead of using a colposcope (Cremer et al., 2010; Ford, 2011). Implementing a new, software-based analysis in such a setting comes at no additional cost. Together with standardised protocols for acquiring the images, this tool might even be used for quantification of the lesion size to monitor treatment effect over time.

Furthermore, distribution of a software tool to remote areas would be simple and cheap. If made available online, it could easily be downloaded to imaging devices. Instructions and training material could equally be made available, either bundled in

the software or as an online resource. Furthermore, online availability allows for upgrades of the software's diagnostic and management algorithms as scientific progress is made and accuracy improved.

Sandy patches, abnormal blood vessels or both?

Sandy patches are characteristic lesions of FGS (Kjetland et al., 2005; Poggensee et al., 2001). Biopsies show that they contain schistosome eggs, both viable and calcified (Jourdan et al., 2011a; Kjetland et al., 1996). The presence of sandy patches as observed visually is diagnostic of FGS in endemic areas (Kjetland et al., 2014).

Abnormal blood vessels are associated with FGS (Kjetland et al., 2005) but the exact pathogenetic processes causing these characteristic blood vessels are unknown (Jourdan et al., 2013, 2011b). It has been hypothesized that worms and/or eggs obstructing the vascular lumen may cause dilatation of vessels or that perivascular granuloma and oedema caused by eggs can cause compression of the vessel lumen resulting in occlusion and proximal dilatation (Jourdan et al., 2013). Egg-related thrombosis has also been described in a case report and it has been hypothesized that this may contribute to causing the abnormal blood vessels in FGS (Jourdan et al., 2013). Although the characteristic abnormal blood vessels are associated with FGS they are not considered pathognomonic (Kjetland et al., 2005). Detection of abnormal blood vessels is therefore not sufficient to diagnose FGS in itself (Kjetland et al., 2014). It could however be a valuable supplement in an algorithm using detection of both sandy patches and abnormal blood vessels.

Limitations

The image material used in this thesis was obtained from young women. The cervical mucosa changes with age; columnar epithelium is gradually covered by squamous epithelium (Kleppa et al., 2014a). This may alter the appearance of the ectoverix and it may limit the method's applicability in older women. Furthermore, there is a higher prevalence of cervical malignancy with increasing age, which makes it even more important to focus on this differential diagnosis when approaching older age groups.

We did not include pregnant women in this study. It is likely that the ectocervical appearance is altered during pregnancy. Furthermore, there are no studies published on the appearance of lesions specific to FGS during pregnancy.

Future research

The ultimate goal of medical research is clinical implementation to improve health care. In this thesis, I've presented new methods for diagnosing FGS based on analysis of gynaecological images. However, more research is required before this can be implemented in a clinical setting at the point of care: How acceptable is this method for the patient? Will women be comfortable with the health care worker acquiring gynaecological images for analysis? And will this be harder to accept if the imaging device in question is a mobile phone? Furthermore, clinical implementation will require that the image analysis works on a wide range of imaging devices available in endemic areas. Quality control must be done in real life settings with scrutiny of user friendliness and reliability in different settings.

The methods described in this thesis rely on gynaecological photos acquired by an imaging device such as a handheld camera. However, this adds a layer of difficulty to the problem, as it is difficult to standardize the process of acquiring the image: light source, distance from the cervix, image resolution etc. Other possibilities should be sought; perhaps analysis of the colour changes in the gynaecological mucosa could be more reliably assessed by using colour spectrometry? Or maybe eggs in the mucosa could be visualized in vivo using microendoscopy? This would however require the concept and design followed by production and distribution of such specialized devices. This will most likely be an expensive process, making it difficult to implement in low-resource settings.

References

Aker, J.C., Mbiti, I.M., 2010. Mobile Phones and Economic Development in Africa. J. Econ. Perspect. 24, 207–232.

Ayele, B., Erko, B., Legesse, M., Hailu, A., Medhin, G., 2008. Evaluation of circulating cathodic antigen (CCA) strip for diagnosis of urinary schistosomiasis in Hassoba school children, Afar, Ethiopia. Parasite 15, 69–75.

Badawy, A.H., 1962. Schistosomiasis of the cervix. Br. Med. J. 5275, 369–372.

Boppart, S.A., Richards-Kortum, R., 2014. Point-of-care and point-of-procedure optical imaging technologies for primary care and global health. Sci. Transl. Med. 6.

Chaki, N., Shaikh, S.H., Saeed, K., 2014. Exploring Image Binarization Techniques. Springer.

Chenine, A.L., Shai-Kobiler, E., Steele, L.N., Ong, H., Augostini, P., Song, R., Lee, S.J., Autissier, P., Ruprecht, R.M., Secor, W.E., 2008. Acute Schistosoma mansoni Infection Increases Susceptibility to Systemic SHIV Clade C Infection in Rhesus Macaques after Mucosal Virus Exposure. PLoS Negl Trop Dis 2, e265.

Chirenje, Z.M., Rusakaniko, S., Kirumbi, L., Ngwalle, E.W., Makuta-Tlebere, P., Kaggwa, S., Mpanju-Shumbusho, W., Makoae, L., 2001. Situation analysis for cervical cancer diagnosis and treatment in east, central and southern African countries. Bull. World Health Organ. 79, 127–132.

Colley, D.G., Bustinduy, A.L., Secor, W.E., King, C.H., 2014. Human

schistosomiasis. Lancet 383, 2253–64.

Collins, L.M., Lanza, S.T., 2013. Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences. John Wiley & Sons.

Cremer, M.L., Peralta, E.I., Dheming, S.G., Jimenez, M.E., Davis-Dao, C. a, Alonzo, T. a, Blumenthal, P.D., Felix, J.C., 2010. Digital assessment of the reproductive tract versus colposcopy for directing biopsies in women with abnormal Pap smears. J. Low. Genit. Tract Dis. 14, 5–10.

Dessein, a J., Hillaire, D., Elwali, N.E., Marquet, S., Mohamed-Ali, Q., Mirghani, a, Henri, S., Abdelhameed, a a, Saeed, O.K., Magzoub, M.M., Abel, L., 1999. Severe hepatic fibrosis in Schistosoma mansoni infection is controlled by a major locus that is closely linked to the interferon-gamma receptor gene. Am. J. Hum. Genet. 65, 709–721.

Dickman, E.D., Doll, T.J., Chiu, C.K., Ferris, D.G., 2001. Identification of cervical neoplasia using a simulation of human vision. J. Low. Genit. Tract Dis. 5, 144–52.

Doenhoff, M.J., Chiodini, P.L., Hamilton, J. V, 2004. Specific and sensitive diagnosis of schistosome infection: can it be done with antibodies? Trends Parasitol. 20, 35–9.

Doubal, F.N., MacGillivray, T.J., Patton, N., Dhillon, B., Dennis, M.S., Wardlaw, J.M., 2010. Fractal analysis of retinal vessels suggests that a distinct vasculopathy causes lacunar stroke. Neurology 74, 1102–1107.

Downs, J.A., Mguta, C., Kaatano, G.M., Mitchell, K.B., Bang, H., Simplice, H., Kalluvya, S.E., Changalucha, J.M., Johnson Jr, W.D., Fitzgerald, D.W., 2011.

Urogenital Schistosomiasis in Women of Reproductive Age in Tanzania's Lake Victoria Region. Am. J. Trop. Med. Hyg. 84, 364–369.

Drain, P.K., Hyle, E.P., Noubary, F., Freedberg, K. a., Wilson, D., Bishai, W.R., Rodriguez, W., Bassett, I. V., 2014. Diagnostic point-of-care tests in resource-limited settings. Lancet Infect. Dis. 14, 239–249.

Etzo, S., Collender, G., 2010. The mobile phone "revolution" in Africa: Rhetoric or reality? Afr. Aff. (Lond). 109, 659–668.

Feldmeier, H., Helling-Giese, G., Poggensee, G., 2001. Unreliability of PAP smears to diagnose female genital schistosomiasis. Trop. Med. Int. Heal. 6, 31–33.

Feldmeier, H., Krantz, I., Poggensee, G., 1994. Female genital schistosomiasis as a risk-factor for the transmission of HIV. AIDS 5, 368–372.

Ford, C., 2011. Implementation of cervical cancer prevention services for HIV-infected women in Zambia: measuring program effectiveness. J. Fam. Plann. Reprod. Health Care 37, 186.

Formann, A.K., Kohlmann, T., 1996. Latent class analysis in medical research. Stat. Methods Med. Res. 5, 179–211.

Friedberg, D., Berry, A. V, Schneider, J., 1991. Schistosomiasis of the female genital tract. Med. J. South Africa 8, S1–S16.

Gibson, R.W.B., 1925. Bilharziasis of the female genital tract. Med. J. South Africa 21, 44–45.

Goodman, L.A., 1974. The analysis of systems of qualitative variables when some of the variables are unobservable. Part IA modified latent structure approach. Am. J. Sociol. 79, 1179–1259.

Greenspan, H., Gordon, S., Zimmerman, G., Lotenberg, S., Jeronimo, J., Antani, S., Long, R., 2009. Automatic detection of anatomical landmarks in uterine cervix images. IEEE Trans. Med. Imaging 28, 454–68.

Gryseels, B., Polman, K., Clerinx, J., Kestens, L., 2006. Human schistosomiasis. Lancet.

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 29–36.

Hegertun, I.E.A., Sulheim Gundersen, K.M., Kleppa, E., Zulu, S.G., Gundersen, S.G., Taylor, M., Kvalsvig, J.D., Kjetland, E.F., 2013. S. haematobium as a Common Cause of Genital Morbidity in Girls: A Cross-sectional Study of Children in South Africa. PLoS Negl. Trop. Dis. 7, e2104.

Hotez, P.J., Fenwick, A., Kjetland, E.F., 2009. Africa's 32 cents solution for HIV/AIDS. PLoS Negl. Trop. Dis. 3, e430.

Ibironke, O., Koukounari, A., Asaolu, S., Moustaki, I., Shiff, C., 2012. Validation of a new test for schistosoma haematobium based on detection of dra1 DNA fragments in urine: evaluation through latent class analysis. PLoS Negl. Trop. Dis. 6, e1464.

Ji, Q., Engel, J., Craine, E., 2000. Texture analysis for classification of cervix lesions. IEEE Trans. Med. Imaging 19, 1144–9.

Jourdan, P.M., Holmen, S.D., Gundersen, S.G., Roald, B., Kjetland, E.F., 2011a. HIV target cells in Schistosoma haematobium-infected female genital mucosa. Am. J. Trop. Med. Hyg. 85, 1060–4.

Jourdan, P.M., Randrianasolo, B.S., Feldmeier, H., Chitsulo, L., Ravoniarimbinina,

P., Roald, B., Kjetland, E.F., 2013. Pathologic mucosal blood vessels in active female genital schistosomiasis: new aspects of a neglected tropical disease. Int. J. Gynecol. Pathol. 32, 137–40.

Jourdan, P.M., Roald, B., Poggensee, G., Gundersen, S.G., Kjetland, E.F., 2011b. Increased Vascularity in Cervicovaginal mucosa with Schistosoma haematobium infection. PLoS Negl. Trop. Dis. 5, e1170.

Kirkaldy-Willis, W.H., 1946. Cystoscopy in the diagnosis and treatment of Bilharzia haematobium infection. Br. J. Surg. 34, 189–194.

Kjetland, E.F., Kurewa, E.N., Ndhlovu, P.D., Gwanzura, L., Mduluza, T., Gomo, E., Midzi, N., Mason, P.R., Friis, H., Gundersen, S.G., 2008. Female genital schistosomiasis - a differential diagnosis to sexually transmitted disease: Genital itch and vaginal discharge as indicators of genital S. haematobium morbidity in a cross-sectional study in endemic rural Zimbabwe. Trop. Med. Int. Heal. 13, 1509–1517.

Kjetland, E.F., Leutscher, P.D.C., Ndhlovu, P.D., 2012. A review of female genital schistosomiasis. Trends Parasitol. 28, 58–65.

Kjetland, E.F., Ndhlovu, P.D., Gomo, E., Mduluza, T., Midzi, N., Gwanzura, L., Mason, P.R., Sandvik, L., Friis, H., Gundersen, S.G., 2006. Association between genital schistosomiasis and HIV in rural Zimbabwean women. AIDS 20, 593–600.

Kjetland, E.F., Ndhlovu, P.D., Mduluza, T., Gomo, E., Gwanzura, L., Mason, P.R., Kurewa, E.N., Midzi, N., Friis, H., Gundersen, S.G., 2005. Simple clinical manifestations of genital Schistosoma haematobium infection in rural

Zimbabwean women. Am. J. Trop. Med. Hyg. 72, 311–319.

Kjetland, E.F., Norseth, H.M., Taylor, M., Lillebø, K., Kleppa, E., Holmen, S.D., Andebirhan, A., Yohannes, T.H., Gundersen, S.G., Vennervald, B.J., Bagratee, J., Onsrud, M., Leutscher, P.D.C., 2014. Classification of the lesions observed in female genital schistosomiasis. Int. J. Gynaecol. Obstet. 2–3.

Kjetland, E.F., Poggensee, G., Helling-Giese, G., Richter, J., Sjaastad, A., Chitsulo, L., Kumwenda, N., Gundersen, S.G., Krantz, I., Feldmeier, H., 1996. Female genital schistosomiasis due to Schistosoma haematobium. Clinical and parasitological findings in women in rural Malawi. Acta Trop. 62, 239–255.

Kjetland, E.F., ten Hove, R.J., Gomo, E., Midzi, N., Gwanzura, L., Mason, P.R., Friis, H., Verweij, J.J., Gundersen, S.G., Ndhlovu, P.D., Mduluza, T., van Lieshout, L., 2009. Schistosomiasis PCR in vaginal lavage as an indicator of genital Schistosoma haematobium infection in rural Zimbabwean women. Am. J. Trop. Med. Hyg. 81, 1050–1055.

Kleppa, E., Holmen, S.D., Lillebø, K., Kjetland, E.F., Gundersen, S.G., Taylor, M., Moodley, P., Onsrud, M., 2014a. Cervical ectopy: associations with sexually transmitted infections and HIV. A cross-sectional study of high school students in rural South Africa. Sex. Transm. Infect. 1–6.

Kleppa, E., Ramsuran, V., Zulu, S., Karlsen, G.H., Bere, A., Passmore, J.-A.S., Ndhlovu, P., Lillebø, K., Holmen, S.D., Onsrud, M., Gundersen, S.G., Taylor, M., Kjetland, E.F., Ndung'u, T., 2014b. Effect of female genital schistosomiasis and anti-schistosomal treatment on monocytes, CD4+ T-cells and CCR5 expression in the female genital tract. PLoS One 9, e98593.

Lalonde, M., Beaulieu, M., Gagnon, L., 2001. Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching. IEEE Trans. Med. Imaging 20, 1193–1200.

Landini, G., Murray, P.I., Misson, G.P., 1995. Local connected fractal dimensions and lacunarity analyses of 60° fluorescein angiograms. Investig. Ophthalmol. Vis. Sci. 36, 2749–2755.

Leutscher, P.D., Pedersen, M., Raharisolo, C., Jensen, J.S., Hoffman, S., Lisse, I., Ostrowski, S.R., Reimert, C.M., Mauclere, P., Ullum, H., 2005. Increased prevalence levels of leukocytes and elevated cytokine levels in semen from Schistosoma haematobium-infected individuals. J. Infect. Dis. 191, 1639–1647.

Leutscher, P.D., Ramarokoto, C.E., Reimert, C., Feldmeier, H., Esterre, P., Vennervald, B.J., 2000. Community-based study of genital schistosomiasis in men from Madagascar [letter]. Lancet 355, 117–118.

Leutscher, P.D.C., Van Dam, G.T.J., Reimert, C.M., Ramarakoto, C.E., Deelder, A.M., Ørnbjerg, N., 2008. Eosinophil cationic protein, soluble egg antigen, circulating anodic antigen, and egg excretion in male urogenital schistosomiasis. Am. J. Trop. Med. Hyg. 79, 422–426.

Madden, F.C., 1907. Bilharziosis. Cassell & Company, Limited, London, UK.

Mayeaux, E.., Cox, J.T., 2011. Modern Colposcopy Textbook and Atlas. Lippincott Williams & Wilkins.

Mbabazi, P.S., Andan, O., Fitzgerald, D.W., Chitsulo, L., Engels, D., Downs, J.A., 2011. Examining the relationship between urogenital schistosomiasis and HIV

infection. PLoS Negl. Trop. Dis. 5, e1396.

Metz, C.E., 1978. Basic principles of ROC analysis. Semin. Nucl. Med. 8, 283–298.

Morgas, D.T., Kvalsvig, J.D., Gundersen, S.G., Taylor, M., Kjetland, E.F., Thomassen Morgas, D.E., 2010. Schistosomiasis and water-related practices in school girls in rural KwaZulu-Natal, South Africa. South. African J. Epidemiol. Infect. 25, 30–33.

Ndeffo Mbah, M.L., Poolman, E.M., Drain, P.K., Coffee, M.P., van der Werf, M.J., Galvani, A.P., 2013. HIV and Schistosoma haematobium prevalences correlate in sub-Saharan Africa. Trop. Med. Int. Health 00.

Niblack, W., 1986. An Introduction to Digital Image Processing. Prentice-Hall.

Norseth, H.M., Ndhlovu, P.D., Kleppa, E., Randrianasolo, B.S., Jourdan, P.M., Roald, B., Holmen, S.D., Gundersen, S.G., Bagratee, J., Onsrud, M., Kjetland, E.F., 2014. 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 8, e3229.

Peeling, R.W., Holmes, K.K., Mabey, D., Ronald, a, 2006. Rapid tests for sexually transmitted infections (STIs): the way forward. Sex. Transm. Infect. 82 Suppl 5, v1–v6.

Poggensee, G., Feldmeier, H., Krantz, I., 1999. Schistosomiasis of the female genital tract: public health aspects. Parasitol. Today 15, 378–381.

Poggensee, G., Kiwelu, I., Weger, V., Goppner, D., Diedrich, T., Krantz, I., Feldmeier, H., Göppner, D., 2000. Female genital schistosomiasis of the lower genital tract: prevalence and disease-associated morbidity in Northern

Tanzania. J. Infect. Dis. 181, 1210–1213.

Poggensee, G., Sahebali, S., Van Marck, E., Swai, B., Krantz, I., Feldmeier, H., 2001. Diagnosis of genital cervical schistosomiasis: comparison of cytological, histopathological and parasitological examination. Am. J. Trop. Med. Hyg. 65, 233–236.

Pogue, B.W., Mycek, M.A., Harper, D., 2000. Image analysis for discrimination of cervical neoplasia. J. Biomed. Opt. 5, 72–82.

Quinley, K.E., Gormley, R.H., Ratcliffe, S.J., Shih, T., Szep, Z., Steiner, A., Ramogola-Masire, D., Kovarik, C.L., 2011. Use of mobile telemedicine for cervical cancer screening. J. Telemed. Telecare 17, 203–209.

Randrianasolo, B.S., Jourdan, P.M., Ravoniarimbinina, P., Ramarokoto, C.E., Rakotomanana, F., Ravaoalimalala, V.E., Gundersen, S.G., Feldmeier, H., Vennervald, B.J., van Lieshout, L., Roald, B., Leutscher, P., Kjetland, E.F., Jourdan, M., Ravoniarimbinina, P., Ramarokoto, C.E., Rakotomanana, F., Ravaoalimalala, V.E., Gundersen, S.G., Feldmeier, H., Vennervald, B.J., Lieshout, L. Van, Roald, B., Leutscher, P., 2015. Gynecological Manifestations, Histopathological Findings, and Schistosoma-Specific Polymerase Chain Reaction Results Among Women With Schistosoma haematobium Infection: A Cross-sectional Study in Madagascar. J. Infect. Dis. 1–10.

Redner, R.A., Walker, H.F., 1982. Mixture densities, maximum likelihood, and the EM algorithm 26, 195–239.

Sezgin, M., Sankur, B., 2004. Survey over image thresholding techniques and

quantitative performance evaluation. J. Electron. Imaging 13, 146.

Smith, T.G., Lange, G.D., Marks, W.B., 1996. Fractal methods and results in cellular morphology - Dimensions, lacunarity and multifractals. J. Neurosci. Methods 69, 123–136.

Swai, B., Poggensee, G., Mtweve, S., Krantz, I., 2006. Female genital schistosomiasis as an evidence of a neglected cause for reproductive ill-health: a retrospective histopathological study from Tanzania. BMC Infect. Dis.

van Dam, G.J., de Dood, C.J., Lewis, M., Deelder, A.M., van Lieshout, L., Tanke, H.J., van Rooyen, L.H., Corstjens, P.L. a M., 2013. A robust dry reagent lateral flow assay for diagnosis of active schistosomiasis by detection of Schistosoma circulating anodic antigen. Exp. Parasitol. 135, 274–282.

Van Der Werf, M.J., De Vlas, S.J., Brooker, S., Looman, C.W.N., Nagelkerke, N.J.D., Habbema, J.D.F., Engels, D., 2003. Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. In: Acta Tropica. pp. 125–139.

Vlachokosta, A. a., Asvestas, P. a., Gkrozou, F., Lavasidis, L., Matsopoulos, G.K., Paschopoulos, M., 2013. Classification of hysteroscopical images using texture and vessel descriptors. Med. Biol. Eng. Comput. 51, 859–867.

WHO, 1995. Idenitification of high risk communities for schistosomiasis in Africa: a multicountry study. Geneva, Switzerland.

Youssef, A.F., Fayad, M.M., Shafeek, M.A.E.D.A.-E.-D., 1970. Bilharziasis of the cervix uteri. J. Obstet. Gynaecol. Br. Commonw. 77, 847–851.

Youssif, A.A.H.A.R., Ghalwash, A.Z., Ghoneim, A.A.S.A.R., 2008. Optic disc

detection from normalized digital fundus images by means of a vessels' direction matched filter. IEEE Trans. Med. Imaging 27, 11–18.