

South Africa

Professor Frank Tanser, UCL; AHRI

Mr Sipho Zulu, UKZN; AHRI

**Distribution of urogenital schistosomiasis and HIV infection in
a typical rural South African population with rapid increase in
access to piped water and sanitation: Implications for public health**

Study Documentation

February 11, 2019

Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Identification	DDI.AHRI.Schisto.ClinData.2018.v1

Table of Contents

Overview.....	4
Scope & Coverage.....	4
Producers & Sponsors.....	4
Sampling.....	4
Data Collection.....	5
Data Processing & Appraisal.....	5
Accessibility.....	5
Files Description.....	6
AHRI.Schisto.ClinData.2018.v1.....	6
Variables List.....	7
AHRI.Schisto.ClinData.2018.v1.....	7
Variables Description.....	9
AHRI.Schisto.ClinData.2018.v1.....	10

Distribution of urogenital schistosomiasis and HIV infection in a typical rural South African population with rapid increase in access to piped water and sanitation:Implications for public health

Overview	
Identification	AHRI.Schisto.ClinData.2018.v1
Version	V1.0.0
Abstract The dataset generated will be used to determine the distribution of schistosomiasis in the population (aged 15-50 years) and also used with the other datasets, will locate the hotspots for schisto-HIV coinfection in the population and also identify causal factors for the distribution.	
Kind of Data	Survey Data
Unit of Analysis	Demographic data, each participant was asked a set of questions on the history of bilharzia and also the understanding of it as well as frequency of water contact Clinical data, each participant has one urine test results. None of the participants were tested twice.

Scope & Coverage	
Keywords	Tuberculosis
Topics	Schistosomiasis-hiv coinfection distribution
Time Period(s)	2017-2018
Countries	South Africa
Geographic Coverage uMkhanyakude district in northern KwaZulu-Natal	
Universe Individuals aged 15-50 years (both males and females) who are resident members of households in the AHRI DSS area	

Producers & Sponsors	
Primary Investigator(s)	Professor Frank Tanser, UCL; AHRI Mr Sipho Zulu, UKZN; AHRI
Other Producer(s)	Africa Health Research Institute (AHRI)
Funding Agency/ies	Research Development Grant (UKZN) (RDG) , Core funding College of Health Sciences scholarship (CHS) , Core funding University Capacity Development Program (UCDP) , Core funding
Other Acknowledgment(s)	Dr. Kobus, Herbst , Chief Information Officer , AHRI Dickman, Gareta , Head of Research Data Management , AHRI Njabulo Dayi , Data management, data cleaning , AHRI

Sampling
Sampling Procedure All individuals aged 15-50 years who are resident members of households in the DSS area were eligible to participate in the schistosomiasis study.

Data Collection	
Data Collection Dates	start 2017-08-01 end 2018-03-29

Data Processing & Appraisal	
Data Editing	NA

Accessibility	
Citation Requirements	https://doi.org/10.23664/AHRI.Schisto.ClinData.2018.v1

Files Description

Dataset contains 1 file(s)

AHRI.Schisto.ClinData.2018.v1	
# Cases	3701
# Variable(s)	30

Variables List

Dataset contains 30 variable(s)

File AHRI.Schisto.ClinData.2018.v1							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	IintID	Individual Identifier	continuous	numeric-12.0	3701	0	-
2	DoB	Date Of Birth	discrete	character-11	3701	-	-
3	ca_sex	Sex	discrete	numeric-12.0	3701	0	-
4	ca_age	Age	continuous	numeric-12.0	3701	0	-
5	ca_attmp..	Conducting Contact Attempt - One	discrete	numeric-12.0	3701	0	-
6	ca_attmp..	Contact Attempt Date and Time	discrete	character-11	3701	-	-
7	ca_attmp..	Conducting Contact Attempt - Two	discrete	numeric-12.0	1594	2107	-
8	ca_attmp..	Contact Attempt Date and Time	discrete	character-11	1121	-	-
9	ca_attmp..	Conducting Contact Attempt - Three	discrete	numeric-12.0	408	3293	-
10	ca_attmp..	Contact Attempt Date and Time	discrete	character-11	202	-	-
11	ca_prema..	Was this a Pre-mature completion?	discrete	numeric-12.0	3208	493	-
12	ca_prema..	Reason for pre-mature completion	discrete	numeric-12.0	357	3344	-
13	ca_parti..	Was participation refused?	discrete	numeric-12.0	2851	850	-
14	ca_refus..	Who refuse participation?	discrete	numeric-12.0	219	3482	-
15	ca_visit..	Visit Date	discrete	character-11	2964	-	-
16	emancipa..	Reason emancipated,	discrete	numeric-12.0	292	3409	-
17	emancipa..	Is participant emancipated	discrete	numeric-12.0	3	3698	-
18	specimen..	To have my urine specimen stored	discrete	numeric-12.0	1860	1841	-
19	samples..	To have my stored urine specimen be used by Africa Health Research Institute rese	discrete	numeric-12.0	1860	1841	-
20	urine_mi..	To have my urine tested by microscopy and be informed of the results should I tes	discrete	numeric-12.0	1860	1841	-
21	agree_to..	To answer questions on self- reported general health	discrete	numeric-12.0	1860	1841	-
22	knowledg..	To answer questions about knowledge and history of bilharzia	discrete	numeric-12.0	1857	1844	-
23	future_c..	To be contacted in future if certain parts of the information I provide today are	discrete	numeric-12.0	1860	1841	-

File AHRI.Schisto.ClinData.2018.v1							
#	Name	Label	Type	Format	Valid	Invalid	Question
24	data_lin..	To allow Africa Health Research Institute to link my Department of Health clinic	discrete	numeric-12.0	1860	1841	-
25	know_bil..	Do you know what bilharzia is?	discrete	numeric-12.0	1862	1839	-
26	ever_had..	Have you ever had bilharzia?	discrete	numeric-12.0	1862	1839	-
27	ever_had..	Have you ever had red urine?	discrete	numeric-12.0	1861	1840	-
28	had_bilh..	Have you been treated for bilharzia in the last 12 months?	discrete	numeric-12.0	1846	1855	-
29	frequent..	How often (on average) do you have contact with rivers, dams, lakes or other wate	discrete	numeric-12.0	1851	1850	-
30	Result	Results	discrete	numeric-12.0	1857	1844	-

Variables Description

Dataset contains 30 variable(s)

File : AHRI.Schisto.ClinData.2018.v1

IintID: Individual Identifier

Information	[Type= continuous] [Format=numeric] [Range= 140-209213] [Missing=*]
Statistics [NW/ W]	[Valid=3701 /-] [Invalid=0 /-] [Mean=78153.105 /-] [StdDev=50447.087 /-]

DoB: Date Of Birth

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=3701 /-]

ca_sex: Sex

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=3701 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Male	1742	47.1%
2	Female	1959	52.9%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_age: Age

Information	[Type= continuous] [Format=numeric] [Range= 15-51] [Missing=*]
Statistics [NW/ W]	[Valid=3701 /-] [Invalid=0 /-] [Mean=28.934 /-] [StdDev=10.132 /-]

ca_attmp1_fieldattmp: Conducting Contact Attempt - One

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=3701 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Yes	3701	100.0%
2	No	0	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_attmp1_fielddatetime: Contact Attempt Date and Time

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=3701 /-]

ca_attmp2_fieldattmp: Conducting Contact Attempt - Two

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1594 /-] [Invalid=2107 /-]

Value	Label	Cases	Percentage
1	Yes	1123	70.5%
2	No	471	29.5%
Sysmiss		2107	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_attmp2_fielddatetime: Contact Attempt Date and Time

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=1121 /-]

ca_attmp3_fieldattmp: Conducting Contact Attempt - Three

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=408 /-] [Invalid=3293 /-]

File : AHRI.Schisto.ClinData.2018.v1

ca_attmp3_fielddatmp: Conducting Contact Attempt - Three

Value	Label	Cases	Percentage
1	Yes	202	<div><div></div></div> 49.5%
2	No	206	<div><div></div></div> 50.5%
Sysmiss		3293	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_attmp3_fielddatetime: Contact Attempt Date and Time

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=202 /-]

ca_premature_comp: Was this a Pre-mature completion?

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=3208 /-] [Invalid=493 /-]

Value	Label	Cases	Percentage
1	Yes	357	<div><div></div></div> 11.1%
2	No	2851	<div><div></div></div> 88.9%
Sysmiss		493	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_premature_reas: Reason for pre-mature completion

Information	[Type= discrete] [Format=numeric] [Range= 1-96] [Missing=*]
Statistics [NW/ W]	[Valid=357 /-] [Invalid=3344 /-]

Value	Label	Cases	Percentage
1	Reported dead	19	<div><div></div></div> 5.3%
2	Outmigrated outside PIPSA	293	<div><div></div></div> 82.1%
96	Other	45	<div><div></div></div> 12.6%
Sysmiss		3344	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_particip_refused: Was participation refused?

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=2851 /-] [Invalid=850 /-]

Value	Label	Cases	Percentage
1	Yes	219	<div><div></div></div> 7.7%
2	No	2632	<div><div></div></div> 92.3%
Sysmiss		850	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_refused_by: Who refuse participation?

Information	[Type= discrete] [Format=numeric] [Range= 1-96] [Missing=*]
Statistics [NW/ W]	[Valid=219 /-] [Invalid=3482 /-]

Value	Label	Cases	Percentage
1	Self	211	<div><div></div></div> 96.3%
2	Partner	1	<div><div></div></div> 0.5%
3	HH Head	4	<div><div></div></div> 1.8%
4	BS Owner	1	<div><div></div></div> 0.5%

File : AHRI.Schisto.ClinData.2018.v1

ca_refused_by: Who refuse participation?

Value	Label	Cases	Percentage
96	Other	2	0.9%
Sysmiss		3482	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ca_visitdate: Visit Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=2964 /-]

emancipation: Reason emancipated,

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=292 /-] [Invalid=3409 /-]

Value	Label	Cases	Percentage
0	No	283	96.9%
1	Yes	9	3.1%
Sysmiss		3409	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

emancipated: Is participant emancipated

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
Statistics [NW/ W]	[Valid=3 /-] [Invalid=3698 /-]

Value	Label	Cases	Percentage
1	Child headed household	1	33.3%
2	Participant has a child	2	66.7%
3	Participant is married	0	
Sysmiss		3698	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

specimen_stored: To have my urine specimen stored

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1860 /-] [Invalid=1841 /-]

Value	Label	Cases	Percentage
1	I consent	1860	100.0%
2	I do not consent	0	
Sysmiss		1841	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

samples_used_in_future: To have my stored urine specimen be used by Africa Health Research Institute rese

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1860 /-] [Invalid=1841 /-]

Value	Label	Cases	Percentage
1	I consent	1859	99.9%
2	I do not consent	1	0.1%
Sysmiss		1841	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

File : AHRI.Schisto.ClinData.2018.v1

urine_microscopy: To have my urine tested by microscopy and be informed of the results should I tes

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1860 /-] [Invalid=1841 /-]

Value	Label	Cases	Percentage
1	I consent	1860	100.0%
2	I do not consent	0	
Sysmiss		1841	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

agree_to_answer_questions: To answer questions on self-reported general health

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1860 /-] [Invalid=1841 /-]

Value	Label	Cases	Percentage
1	I consent	1859	99.9%
2	I do not consent	1	0.1%
Sysmiss		1841	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

knowledge_and_history_of_bilharz: To answer questions about knowledge and history of bilharzia

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1857 /-] [Invalid=1844 /-]

Value	Label	Cases	Percentage
1	I consent	1857	100.0%
2	I do not consent	0	
Sysmiss		1844	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

future_communication: To be contacted in future if certain parts of the information I provide today are

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1860 /-] [Invalid=1841 /-]

Value	Label	Cases	Percentage
1	I consent	1860	100.0%
2	I do not consent	0	
Sysmiss		1841	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

data_linkage: To allow Africa Health Research Institute to link my Department of Health clinic

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=1860 /-] [Invalid=1841 /-]

Value	Label	Cases	Percentage
1	I consent	1860	100.0%
2	I do not consent	0	
Sysmiss		1841	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

know_bilharzia: Do you know what bilharzia is?

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
--------------------	--

File : AHRI.Schisto.ClinData.2018.v1

know_bilharzia: Do you know what bilharzia is?

Statistics [NW/ W] [Valid=1862 /-] [Invalid=1839 /-]

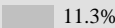


Value	Label	Cases	Percentage
1	Yes	1337	 71.8%
2	No	475	 25.5%
3	Don't Know	50	 2.7%
Sysmiss		1839	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ever_had_bilharzia: Have you ever had bilharzia?

Information [Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]

Statistics [NW/ W] [Valid=1862 /-] [Invalid=1839 /-]


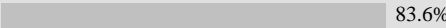

Value	Label	Cases	Percentage
1	Yes	211	 11.3%
2	No	1579	 84.8%
3	Don't Know	72	 3.9%
Sysmiss		1839	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

ever_had_urine: Have you ever had red urine?

Information [Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]

Statistics [NW/ W] [Valid=1861 /-] [Invalid=1840 /-]

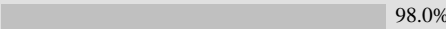
Value	Label	Cases	Percentage
1	Yes	279	 15.0%
2	No	1556	 83.6%
3	Don't Know	26	 1.4%
Sysmiss		1840	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

had_bilharzia_in_last_year: Have you been treated for bilharzia in the last 12 months?

Information [Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]

Statistics [NW/ W] [Valid=1846 /-] [Invalid=1855 /-]

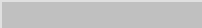
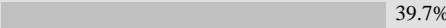
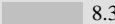
Value	Label	Cases	Percentage
1	Yes	35	 1.9%
2	No	1810	 98.0%
3	Don't Know	1	 0.1%
Sysmiss		1855	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

frequent_water_contact: How often (on average) do you have contact with rivers, dams, lakes or other wate



Information [Type= discrete] [Format=numeric] [Range= 1-6] [Missing=*]

Statistics [NW/ W] [Valid=1851 /-] [Invalid=1850 /-]

Value	Label	Cases	Percentage
1	Never	392	 21.2%
2	Very occasionally (less than once a month)	735	 39.7%
3	Once a month	154	 8.3%

File : AHRI.Schisto.ClinData.2018.v1

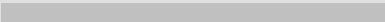

frequent_water_contact: How often (on average) do you have contact with rivers, dams, lakes or other wate

Value	Label	Cases	Percentage
4	Once a week	146	 7.9%
5	More than once a week (not daily)	182	 9.8%
6	Daily	242	 13.1%
Sysmiss		1850	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

Result: Results

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=1857 /-] [Invalid=1844 /-]

Value	Label	Cases	Percentage
0	Negative	1773	 95.5%
1	Positive	84	 4.5%
Sysmiss		1844	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.