

South Africa

Njabulo Dayi, Africa Health Research Institute
Dickman Gareta, Africa Health Research Institute
Nompumelelo Mkwanazi, Africa Health Research Institute
Shahmanesh M, Africa Health Research Institute
Francis SC, Africa Health Research Institute
Baisley K, Africa Health Research Institute
Mchunu SL, Africa Health Research Institute
Ferguson J, Africa Health Research Institute
Smit T, Africa Health Research Institute
Crucitti T, Africa Health Research Institute
Dlamini S, Africa Health Research Institute
Seeley J, Africa Health Research Institute
Pillay Deenan, Africa Health Research Institute
McGrath N, Africa Health Research Institute
Mthiyane TN, Africa Health Research Institute
, Africa Health Research Institute
, Africa Health Research Institute

iGugu_Main_Study_Data_PLOS Medicine

Study Documentation

September 11, 2026

Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Identification	AHRI.iGugu.Main.Study.Data.PLOS.Medicine

Table of Contents

Overview.....	4
Scope & Coverage.....	4
Producers & Sponsors.....	5
Sampling.....	5
Accessibility.....	5
Files Description.....	6
Igugu_main_study_data_PLOS_Medicine.....	6
Variables List.....	7
Igugu_main_study_data_PLOS_Medicine.....	7
Variables Description.....	9
Igugu_main_study_data_PLOS_Medicine.....	10

iGugu_Main_Study_Data_PLOS Medicine

Overview

Identification	AHRI.iGugu.Main.Study.Data.PLOS.Medicine
Version	v1.0.0

Abstract

Igugu_main_study_data_PLOS Medicine:

This dataset is the main study data for: Francis SC, Mthiyane TN, Baisley K, Mchunu SL, Ferguson J, Smit T, Crucitti T, Gareta D, Dlamini S, Mutevedzi T, Seeley J, Pillay D, McGrath N, Shahmanesh M. Prevalence of Sexually Transmitted Infections among Young People in South Africa: A Nested Survey in a Health and Demographic Surveillance Site. PLOS Medicine 2017

In this paper, we report on the results of an acceptability, feasibility and prevalence study of home-based sampling for STI/BV among young men and women aged 15-24 years old in the Africa Health Research Institute (AHRI) health and demographic surveillance site (DSS) in rural KwaZulu-Natal, South Africa. A total of 1342 young people, stratified by age (15-19 and 20-24 years) and sex were selected from the HDSS sampling frame, and 447 were enrolled (70% of those contacted). Participants were interviewed and blood, self-collected urine (men) and vaginal swabs (women) were tested for herpes simplex virus type-2 (HSV-2), chlamydia, gonorrhoea, trichomoniasis, and BV. Further details on study methods are available in the published work. The dataset provided here is the main analytical dataset for the study. It is a de-identified dataset for the 477 study participants, with only the variables necessary to replicate the analysis presented in the paper. To preserve confidentiality, we report ages in 5-yr intervals, all dates are removed, and education has been recoded into 2 categories. A dummy identification number has been added (i.e. not the actual study ID).

Igugu_study_flow_chart_PLOS Medicine:

This dataset is a supplementary dataset for: Francis SC, Mthiyane TN, Baisley K, Mchunu SL, Ferguson J, Smit T, Crucitti T, Gareta D, Dlamini S, Mutevedzi T, Seeley J, Pillay D, McGrath N, Shahmanesh M. Prevalence of Sexually Transmitted Infections among Young People in South Africa: A Nested Survey in a Health and Demographic Surveillance Site. PLOS Medicine 2017

In this paper, we report on the results of an acceptability, feasibility and prevalence study of home-based sampling for STI/BV among young men and women aged 15-24 years old in the Africa Health Research Institute (AHRI) health and demographic surveillance site (HDSS) in rural KwaZulu-Natal, South Africa. A total of 1342 young people, stratified by age (15-19 and 20-24 years) and sex were selected from the HDSS sampling frame, and 447 were enrolled (70% of those contacted). Participants were interviewed and blood, self-collected urine (men) and vaginal swabs (women) were tested for herpes simplex virus type-2 (HSV-2), chlamydia, gonorrhoea, trichomoniasis, and BV. Further details on study methods are available in the published work.

The dataset provided here is a supplementary dataset containing information on the 1342 young people who were sampled. It is a de-identified dataset with only the variables necessary to replicate the analysis presented in Supplementary Table 1 in the paper. To preserve confidentiality, we report ages in 5-yr intervals, all dates are removed, and education has been recoded into 2 categories. A dummy identification number has been added (i.e. not the actual study ID). The dataset was constructed using information collected in the annual survey of the AHRI HDSS; the survey data and full documentation are available on the AHRI data repository.

Scope & Coverage

Countries	South Africa
Geographic Coverage	
Universe	

Producers & Sponsors	
Primary Investigator(s)	Njabulo Dayi, Africa Health Research Institute Dickman Gareta, Africa Health Research Institute Nompumelelo Mkwanazi, Africa Health Research Institute Shahmanesh M, Africa Health Research Institute Francis SC, Africa Health Research Institute Baisley K, Africa Health Research Institute Mchunu SL, Africa Health Research Institute Ferguson J, Africa Health Research Institute Smit T, Africa Health Research Institute Crucitti T, Africa Health Research Institute Dlamini S, Africa Health Research Institute Seeley J, Africa Health Research Institute Pillay Deenan, Africa Health Research Institute McGrath N, Africa Health Research Institute Mthiyane TN, Africa Health Research Institute , Africa Health Research Institute , Africa Health Research Institute
Other Producer(s)	Africa Health Research Institute (AHRI)

Sampling
<u>Sampling Procedure</u> Simple random sample of DSS residents aged 15-24 years, stratified by age group (15-19 and 20-24) and sex

Accessibility
<u>Access Conditions</u> Access to the data requires accurate completion of the online data access application form accessible on the AHRI Data repository(< https://data.ahri.org/ >). Data users are required to abide by the data use conditions stipulated on the application for access to the data. Failure to do so may result in their data access privileges being revoked by the Data Custodian. In order to recognise the effort and intellectual contributions of AHRI investigators in producing and curating the data, users of AHRI data must acknowledge the source of the data and abide by the terms and conditions under which the data is accessed and must cite the dataset in publication using the citation provided as part of this documentation. All analytical datasets published on the AHRI Data Repository are assigned digital object identifier (DOIs) and the DOIs can be found on the Data Repository under Study Description tab - Access policy. AHRI data users are required to always cite the dataset using the relevant DOI.
<u>Citation Requirements</u> Dayi, N., Gareta, D., Mkwanazi, N., Shahmanesh, M., Francis, S. C., Baisley, K., Mchunu, S. L., Ferguson, J., Smit, T., Crucitti, T., Dlamini, S., Seeley, J., Deenan, P., McGrath, N., & Mthiyane, T. N. (2026). iGugu_Main_Study_Data_PLOS Medicine [Dataset]. Africa Health Research Institute. DOI: https://doi.org/10.23664/AHRI.IGUGU.MAIN.STUDY.DATA.PLOS.MEDICINE

Files Description

Dataset contains 1 file(s)

Igugu_main_study_data_PLOS Medicine	
# Cases	447
# Variable(s)	58

Variables List

Dataset contains 58 variable(s)

File Igugu_main_study_data_PLOS Medicine							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	idno	-	continuous	numeric-9.0	447	0	-
2	gender	Sex	discrete	numeric-9.0	447	0	-
3	agecat	Age group	discrete	numeric-9.0	447	0	-
4	inschool	In school?	discrete	numeric-9.0	446	1	-
5	inschool ..	Inschool or working	discrete	numeric-9.0	447	0	-
6	work	Working?	discrete	numeric-9.0	446	1	-
7	educ	Highest level of education completed	discrete	numeric-16.0	447	0	-
8	ses_cat	Socio-economic status	discrete	numeric-9.0	445	2	-
9	ever_smoke	Ever smoked cigarette	discrete	numeric-21.0	445	2	-
10	everdrun ..	Ever drunk one drink of alcohol	discrete	numeric-20.0	445	2	-
11	used_can ..	Ever used cannabis/dagga	discrete	numeric-21.0	446	1	-
12	otherdru ..	Use other drugs	discrete	numeric-9.0	447	0	-
13	circumci ..	Ever been circumcised	discrete	numeric-21.0	188	259	-
14	vaginain ..	Do you ever cleanse inside vagina?	discrete	numeric-20.0	258	189	-
15	hormonal	Use hormonal contraception	discrete	numeric-9.0	259	188	-
16	contrace ..	Currently using contraceptives	discrete	numeric-23.0	254	193	-
17	everhadsex	Have you ever had sex	discrete	numeric-20.0	443	4	-
18	usedacon ..	Used condom last time you had sex?	discrete	numeric-20.0	220	227	-
19	genital ..	Ever touched genitals	discrete	numeric-9.0	447	0	-
20	oralsex ..	Oral sex (provide)	discrete	numeric-20.0	436	11	-
21	oralsex ..	Oral sex (receive)	discrete	numeric-20.0	441	6	-
22	everhada ..	Ever had anal sex	discrete	numeric-20.0	440	7	-
23	sex_part ..	Number of current sexual partners	discrete	numeric-11.0	414	33	-
24	life_par ..	Number of lifetime partners	discrete	numeric-20.0	407	40	-
25	sextrans ..	In sexual relationship for material	discrete	numeric-21.0	223	224	-
26	violence ..	Violence-victim	discrete	numeric-20.0	441	6	-
27	doneviol ..	Violence-perpetrator	discrete	numeric-20.0	440	7	-
28	discuss ..	Did partner discuss his/her HIV status with you?	discrete	numeric-20.0	222	225	-
29	know_par ..	Do you know his/her HIV status	discrete	numeric-32.0	222	225	-
30	discuss ..	Have you discussed your HIV status with your partner	discrete	numeric-20.0	220	227	-

File Igugu_main_study_data_PLOS Medicine							
#	Name	Label	Type	Format	Valid	Invalid	Question
31	understa ..	Understand the study	continuous	numeric-8.0	432	15	-
32	understa ..	Ease of understanding instructions for sample collection	continuous	numeric-9.0	417	30	-
33	sampleco ..	Ease of sample collection	continuous	numeric-9.0	416	31	-
34	feltcomf ..	Felt comfortable	continuous	numeric-8.0	431	16	-
35	feltinco ..	Felt in control	continuous	numeric-8.0	431	16	-
36	feltrela ..	Felt relaxed	continuous	numeric-8.0	430	17	-
37	feltdoin ..	Was doing it correctly	continuous	numeric-8.0	425	22	-
38	location ..	Location was private	continuous	numeric-8.0	430	17	-
39	wasanxious	Was enxious	continuous	numeric-8.0	418	29	-
40	wasembar ..	Was embarrassed	continuous	numeric-8.0	431	16	-
41	waspainful	Was painful	continuous	numeric-8.0	430	17	-
42	ct_res	Chlamydia	discrete	numeric-13.0	436	11	-
43	ng_res	Gonorrhoea	discrete	numeric-13.0	436	11	-
44	tv_res	Trichomonas vaginalis	discrete	numeric-13.0	436	11	-
45	hsv_res	HSV-2	discrete	numeric-13.0	419	28	-
46	bv_res	Bacterial vaginosis	discrete	numeric-12.0	239	208	-
47	syp_res	Syphilis	discrete	numeric-24.0	422	25	-
48	ng_ct	Gonorrhoea or Chlamydia	discrete	numeric-9.0	436	11	-
49	sti	Any STI	discrete	numeric-9.0	436	11	-
50	hiv2016	HIV status (2016 serosurvey)	discrete	numeric-9.0	230	217	-
51	anysymptom	Any symptoms	discrete	numeric-9.0	447	0	-
52	resid	Residence location	discrete	numeric-10.0	447	0	-
53	strata	Stratum	discrete	numeric-18.0	447	0	-
54	Ntotwt	Sampling & non-response weight (normalised)	continuous	numeric-9.0	447	0	-
55	Ndwt	Sampling weight (normalised)	continuous	numeric-9.0	447	0	-
56	contacti ..	Preferred method of contact if negative	discrete	numeric-14.0	390	57	-
57	contacti ..	Preferred method of contact if positive	discrete	numeric-14.0	390	57	-
58	anyphone	Has access to telephone?	discrete	numeric-9.0	447	0	-

Variables Description

Dataset contains 58 variable(s)

File : Igugu_main_study_data_PLOS Medicine

idno

Information [Type= continuous] [Format=numeric] [Range= 1001-1447] [Missing=*]

Statistics [NW/ W] [Valid=447 /-] [Invalid=0 /-] [Mean=1224 /-] [StdDev=129.182 /-]

gender: Sex

Information [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]

Statistics [NW/ W] [Valid=447 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Male	188	42.1%
2	Female	259	57.9%

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.

agecat: Age group

Information [Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]

Statistics [NW/ W] [Valid=447 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	15-19	248	55.5%
2	20-24	199	44.5%

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.

inschool: In school?

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=446 /-] [Invalid=1 /-]

Value	Label	Cases	Percentage
0	No	143	32.1%
1	Yes	303	67.9%
Sysmiss		1	

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.

inschool_work: Inschool or working

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=447 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	131	29.3%
1	Yes	316	70.7%

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.

work: Working?

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=446 /-] [Invalid=1 /-]

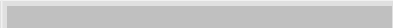
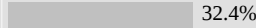
Value	Label	Cases	Percentage
0	No	412	92.4%
1	Yes	34	7.6%
Sysmiss		1	

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 : Igugu_main_study_data_PLOS Medicine

educ: Highest level of education completed

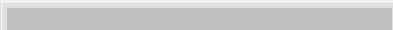
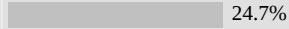
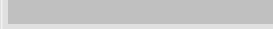
Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=447 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Less than matric	302	 67.6%
2	Matric or higher	145	 32.4%

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.

ses_cat: Socio-economic status

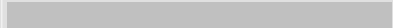
Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
Statistics [NW/ W]	[Valid=445 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
1	Low	197	 44.3%
2	Middle	110	 24.7%
3	High	138	 31.0%
Sysmiss		2	

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_smoke: Ever smoked cigarette

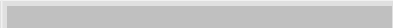
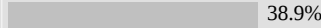
Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=445 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
0	No	406	 91.2%
1	Yes	38	 8.5%
3	Prefer not to answer	1	 0.2%
Sysmiss		2	

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.

everdrunk_alcohol: Ever drunk one drink of alcohol

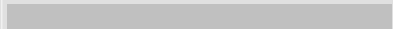
Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=445 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
0	No	266	 59.8%
1	Yes	173	 38.9%
3	Prefer not to answer	6	 1.3%
Sysmiss		2	

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.

used_cannabis: Ever used cannabis/dagga

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=446 /-] [Invalid=1 /-]

Value	Label	Cases	Percentage
0	No	427	 95.7%
1	Yes	19	 4.3%
3	Prefer not to answer	0	
Sysmiss		1	

File : Igugu_main_study_data_PLOS Medicine

used_cannabis: Ever used cannabis/dagga

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.

otherdrugs_ever: Use other drugs

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=447 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	442	98.9%
1	Yes	5	1.1%

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.

circumcised: Ever been circumcised

Information [Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]

Statistics [NW/ W] [Valid=188 /-] [Invalid=259 /-]

Value	Label	Cases	Percentage
0	No	92	48.9%
1	Yes	96	51.1%
3	Prefer not to answer	0	
Sysmiss		259	

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.

vaginainside_cleanse: Do you ever cleanse inside vagina?

Information [Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]

Statistics [NW/ W] [Valid=258 /-] [Invalid=189 /-]

Value	Label	Cases	Percentage
0	No	214	82.9%
1	Yes	41	15.9%
3	Prefer not to answer	3	1.2%
Sysmiss		189	

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.

hormonal: Use hormonal contraception

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=259 /-] [Invalid=188 /-]

Value	Label	Cases	Percentage
0	No	219	84.6%
1	Yes	40	15.4%
Sysmiss		188	

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.

contraception_use: Currently using contraceptives

Information [Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]

Statistics [NW/ W] [Valid=254 /-] [Invalid=193 /-]

Value	Label	Cases	Percentage
0	No	159	62.6%
1	Yes	94	37.0%
3	Prefer not to answer	1	0.4%

File : Igugu_main_study_data_PLOS Medicine

contraception_use: Currently using contraceptives

Value	Label	Cases	Percentage
Sysmiss		193	

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.

everhadsex: Have you ever had sex

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=443 /-] [Invalid=4 /-]

Value	Label	Cases	Percentage
0	No	215	48.5%
1	Yes	216	48.8%
3	Prefer not to answer	12	2.7%
Sysmiss		4	

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.

usedacondom: Used condom last time you had sex?

Information	[Type= discrete] [Format=numeric] [Range= 0-8] [Missing=*]
Statistics [NW/ W]	[Valid=220 /-] [Invalid=227 /-]

Value	Label	Cases	Percentage
0	No	116	52.7%
1	Yes	98	44.5%
3	Prefer not to answer	5	2.3%
8	Don't remember	1	0.5%
Sysmiss		227	

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.

genital_touch: Ever touched genitals

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=447 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	No	264	59.1%
1	Yes	183	40.9%

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.

oralsex_give: Oral sex (provide)

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=436 /-] [Invalid=11 /-]

Value	Label	Cases	Percentage
0	No	365	83.7%
1	Yes	65	14.9%
3	Prefer not to answer	6	1.4%
Sysmiss		11	

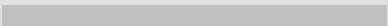
Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

oralsex_receive: Oral sex (receive)

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=441 /-] [Invalid=6 /-]

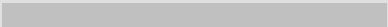
File : Igugu_main_study_data_PLOS Medicine

oralsex_receive: Oral sex (receive)

Value	Label	Cases	Percentage
0	No	334	 75.7%
1	Yes	100	 22.7%
3	Prefer not to answer	7	 1.6%
Sysmiss		6	

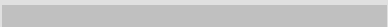
Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

everhadanalsex: Ever had anal sex

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]		
Statistics [NW/ W]	[Valid=440 /-] [Invalid=7 /-]		
Value	Label	Cases	Percentage
0	No	419	 95.2%
1	Yes	14	 3.2%
3	Prefer not to answer	7	 1.6%
Sysmiss		7	

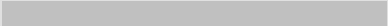
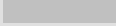
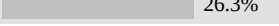
Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

sex_partners: Number of current sexual partners

Information	[Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]		
Statistics [NW/ W]	[Valid=414 /-] [Invalid=33 /-]		
Value	Label	Cases	Percentage
0	None	215	 51.9%
1	One	177	 42.8%
2	Two or more	22	 5.3%
Sysmiss		33	

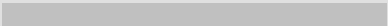
Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

life_partners: Number of lifetime partners

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]		
Statistics [NW/ W]	[Valid=407 /-] [Invalid=40 /-]		
Value	Label	Cases	Percentage
0	None	215	 52.8%
1	One	85	 20.9%
2	Two or more	107	 26.3%
3	Prefer not to answer	0	
Sysmiss		40	

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.

sextransact: In sexual relationship for material

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]		
Statistics [NW/ W]	[Valid=223 /-] [Invalid=224 /-]		
Value	Label	Cases	Percentage
0	No	189	 84.8%
1	Yes	18	 8.1%
3	Prefer not to answer	16	 7.2%

File : Igugu_main_study_data_PLOS Medicine

sextransact: In sexual relationship for material

Value	Label	Cases	Percentage
Sysmiss		224	

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.

violencetoyou: Violence-victim

Information	[Type= discrete] [Format=numeric] [Range= 0-93] [Missing=*]
Statistics [NW/ W]	[Valid=441 /-] [Invalid=6 /-]

Value	Label	Cases	Percentage
0	No	396	89.8%
1	Yes	37	8.4%
93	Prefer not to answer	8	1.8%
Sysmiss		6	

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.

doneviolence: Violence-perpetrator

Information	[Type= discrete] [Format=numeric] [Range= 0-93] [Missing=*]
Statistics [NW/ W]	[Valid=440 /-] [Invalid=7 /-]

Value	Label	Cases	Percentage
0	No	409	93.0%
1	Yes	19	4.3%
93	Prefer not to answer	12	2.7%
Sysmiss		7	

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.

discuss_partnerstatus: Did partner discuss his/her HIV status with you?

Information	[Type= discrete] [Format=numeric] [Range= 0-8] [Missing=*]
Statistics [NW/ W]	[Valid=222 /-] [Invalid=225 /-]

Value	Label	Cases	Percentage
0	No	104	46.8%
1	Yes	106	47.7%
3	Prefer not to answer	8	3.6%
8	Don't remember	4	1.8%
Sysmiss		225	

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_partnerstatus: Do you know his/her HIV status

Information	[Type= discrete] [Format=numeric] [Range= 0-4] [Missing=*]
Statistics [NW/ W]	[Valid=222 /-] [Invalid=225 /-]

Value	Label	Cases	Percentage
0	No	107	48.2%
1	Yes	109	49.1%
3	Don't remember	1	0.5%
4	Prefer not to answer	5	2.3%
Sysmiss		225	

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 : Igugu_main_study_data_PLOS Medicine

discuss_yourstatus: Have you discussed your HIV status with your partner

Information	[Type= discrete] [Format=numeric] [Range= 0-8] [Missing=*]
Statistics [NW/ W]	[Valid=220 /-] [Invalid=227 /-]

Value	Label	Cases	Percentage
0	No	80	36.4%
1	Yes	116	52.7%
3	Prefer not to answer	19	8.6%
8	Don't remember	5	2.3%
Sysmiss		227	

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.

understandthestudy: Understand the study

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=432 /-] [Invalid=15 /-] [Mean=12.197 /-] [StdDev=28.949 /-]

understandinstruct: Ease of understanding instructions for sample collection

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=417 /-] [Invalid=30 /-] [Mean=12.859 /-] [StdDev=28.893 /-]

samplecollection: Ease of sample collection

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=416 /-] [Invalid=31 /-] [Mean=14.055 /-] [StdDev=30.905 /-]

feltcomfortable: Felt comfortable

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=431 /-] [Invalid=16 /-] [Mean=10.109 /-] [StdDev=25 /-]

feltincontrol: Felt in control

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=431 /-] [Invalid=16 /-] [Mean=10.183 /-] [StdDev=24.053 /-]

feltrelaxed: Felt relaxed

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=430 /-] [Invalid=17 /-] [Mean=10.251 /-] [StdDev=23.9 /-]

feltdoingcorrectly: Was doing it correctly

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=425 /-] [Invalid=22 /-] [Mean=7.078 /-] [StdDev=18.416 /-]

locationwasprivate: Location was private

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=430 /-] [Invalid=17 /-] [Mean=6.391 /-] [StdDev=20.649 /-]

wasanxious: Was enxious

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
Statistics [NW/ W]	[Valid=418 /-] [Invalid=29 /-] [Mean=67.722 /-] [StdDev=41.22 /-]

wasembarrassed: Was embarrassed

Information	[Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]
-------------	--

File : Igugu_main_study_data_PLOS Medicine

wasembarrassed: Was embarrassed

Statistics [NW/ W] [Valid=431 /-] [Invalid=16 /-] [Mean=68.8 /-] [StdDev=41.461 /-]

waspainful: Was painful

Information [Type= continuous] [Format=numeric] [Range= 0-100] [Missing=*]

Statistics [NW/ W] [Valid=430 /-] [Invalid=17 /-] [Mean=86.398 /-] [StdDev=29.465 /-]

ct_res: Chlamydia

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=436 /-] [Invalid=11 /-]

Value	Label	Cases	Percentage
0	Negative	399	91.5%
1	Positive	37	8.5%
Sysmiss		11	

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.

ng_res: Gonorrhoea

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=436 /-] [Invalid=11 /-]

Value	Label	Cases	Percentage
0	Negative	427	97.9%
1	Positive	9	2.1%
Sysmiss		11	

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.

tv_res: Trichomonas vaginalis

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=436 /-] [Invalid=11 /-]

Value	Label	Cases	Percentage
0	Negative	421	96.6%
1	Positive	15	3.4%
Sysmiss		11	

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.

hsv_res: HSV-2

Information [Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]

Statistics [NW/ W] [Valid=419 /-] [Invalid=28 /-]

Value	Label	Cases	Percentage
0	Negative	311	74.2%
1	Positive	108	25.8%
2	Indeterminate	0	
Sysmiss		28	

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.

bv_res: Bacterial vaginosis

Information [Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]

Statistics [NW/ W] [Valid=239 /-] [Invalid=208 /-]

File : Igugu_main_study_data_PLOS Medicine

bv_res: Bacterial vaginosis

Value	Label	Cases	Percentage
0	Normal	112	46.9%
1	Intermediate	25	10.5%
2	BV	102	42.7%
Sysmiss		208	

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.

syp_res: Syphilis

Information	[Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]		
Statistics [NW/ W]	[Valid=422 /-] [Invalid=25 /-]		
Value	Label	Cases	Percentage
0	Negative	416	98.6%
1	Early/previous infection	5	1.2%
2	Active	1	0.2%
Sysmiss		25	

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.

ng_ct: Gonorrhoea or Chlamydia

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]		
Statistics [NW/ W]	[Valid=436 /-] [Invalid=11 /-]		
Value	Label	Cases	Percentage
0	Neg	391	89.7%
1	Pos	45	10.3%
Sysmiss		11	

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.

sti: Any STI

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]		
Statistics [NW/ W]	[Valid=436 /-] [Invalid=11 /-]		
Value	Label	Cases	Percentage
0	No	376	86.2%
1	Yes	60	13.8%
Sysmiss		11	

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.

hiv2016: HIV status (2016 serosurvey)

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]		
Statistics [NW/ W]	[Valid=230 /-] [Invalid=217 /-]		
Value	Label	Cases	Percentage
0	Negative	194	84.3%
1	Positive	36	15.7%
Sysmiss		217	

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.

# anysymptom: Any symptoms			
Information		[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]	
Statistics [NW/ W]		[Valid=447 -/] [Invalid=0 -/]	
Value	Label	Cases	Percentage
0	No	372	83.2%
1	Yes	75	16.8%
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.			
# resid: Residence location			
Information		[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]	
Statistics [NW/ W]		[Valid=447 -/] [Invalid=0 -/]	
Value	Label	Cases	Percentage
1	Urban	19	4.3%
2	Peri-urban	155	34.7%
3	Rural	273	61.1%
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.			
# strata: Stratum			
Information		[Type= discrete] [Format=numeric] [Range= 1-4] [Missing=*]	
Statistics [NW/ W]		[Valid=447 -/] [Invalid=0 -/]	
Value	Label	Cases	Percentage
1	Males aged 15-19	124	27.7%
2	Females aged 15-19	124	27.7%
3	Males aged 20-24	64	14.3%
4	Females aged 20-24	135	30.2%
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.			
# Ntotwt: Sampling & non-response weight (normalised)			
Information		[Type= continuous] [Format=numeric] [Range= 0.53085458278656-5.57635259628296] [Missing=*]	
Statistics [NW/ W]		[Valid=447 -/] [Invalid=0 -/] [Mean=1.016 -/] [StdDev=0.465 -/]	
# Ndwt: Sampling weight (normalised)			
Information		[Type= continuous] [Format=numeric] [Range= 0.820998728275299-1.17371797561646] [Missing=*]	
Statistics [NW/ W]		[Valid=447 -/] [Invalid=0 -/] [Mean=1 -/] [StdDev=0.156 -/]	
# contactifnegative: Preferred method of contact if negative			
Information		[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]	
Statistics [NW/ W]		[Valid=390 -/] [Invalid=57 -/]	
Value	Label	Cases	Percentage
1	Telephone call	222	56.9%
2	SMS	151	38.7%
3	WhatsApp	17	4.4%
Sysmiss		57	
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.			
# contactifpositive: Preferred method of contact if positive			
Information		[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]	
Statistics [NW/ W]		[Valid=390 -/] [Invalid=57 -/]	

# contactifpositive: Preferred method of contact if positive			
Value	Label	Cases	Percentage
1	Telephone call	230	59.0%
2	SMS	144	36.9%
3	WhatsApp	16	4.1%
Sysmiss		57	
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.			
# anyphone: Has access to telephone?			
Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]		
Statistics [NW/ W]	[Valid=447 /-] [Invalid=0 /-]		
Value	Label	Cases	Percentage
0	No	57	12.8%
1	Yes	390	87.2%
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.			