



# Evaluation Protocol Data Manual

<b>Acronyms and Abbreviations</b> .....	3
<b>Enumeration Data</b> .....	4
<b>Plots</b> .....	4
<b>Households</b> .....	6
<b>Household Members</b> .....	11
<b>Survey Data</b> .....	13
<b>Participant and Community Variables</b> .....	13
<b>Residency and Mobility</b> .....	15
<b>Community Engagement</b> .....	16
<b>Demographics and Education</b> .....	18
<b>Education and Employment</b> .....	20
<b>HIV Testing History, Test Review, and Result Documentation</b> .....	22
<b>HIV Tested and Untested</b> .....	25
<b>HIV Care and Adherence and Medical Care</b> .....	27
<b>Circumcision</b> .....	31
<b>Sexual Behavior</b> .....	34
<b>Sexual Behavior: Three most recent partners, past 12 months</b> .....	39
<b>Reproductive Health and Pregnancy</b> .....	43
<b>Medical Diagnoses</b> .....	44
<b>Substance Use</b> .....	47
<b>Stigma and Stigma Opinion</b> .....	48

**Access to Care**..... 52

**Primary Endpoint Data**..... 54

## Acronyms and Abbreviations

<b>AHS</b>	Annual household survey
<b>ART</b>	Antiretroviral treatment (or therapy, treatment)
<b>BCPP</b>	Botswana Combination Prevention Project
<b>BHS</b>	Baseline household survey
<b>ESS</b>	End of study survey
<b>HIC</b>	HIV incidence cohort
<b>HIV</b>	Human immunodeficiency virus
<b>HIV-1 RNA</b>	HIV-1 ribonucleic acid load in plasma (or viral load)
<b>HTC</b>	HIV testing and counseling
<b>IPMS</b>	Integrated Patient Management System
<b>LTC</b>	Linkage to care
<b>MC</b>	Male circumcision
<b>PIMS</b>	Patient Information Management System
<b>UNAIDS</b>	Joint United Nations Programme on HIV/AIDS
<b>WHO</b>	World Health Organization

## Enumeration Data

### Plots

Variable Name	Description	Additional detail	Availability of Data	
			20% random sample at Year 1 (all pairs)	Remaining 80% sample at Year 3 (pairs 1, 6, and 9)
de_plot_idC	Unique plot identifier	De-identified, unique codes were randomly assigned to each plot	✓	✓
region	Geographical region of community	Classified as 'Northern', 'Southern', or 'Central'	✓	✓
community_rndmN	Community number (1-30)	De-identified	✓	✓
pair_rndmN	Assigned pair number (1-15)	De-identification preserved the 1-15 range (randomly scrambled the pairs) and maintain link to community	✓	✓
intervention	Randomization arm		✓	✓
plot_residency_status	<p>Categorical variable summarizing plot status following all Evaluation Protocol study activities in a given year.</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "Residential/habitable"</li> <li>- "Residential/inhabitable"</li> <li>- "Non-residential"</li> <li>- "Inaccessible"</li> </ul>	Plots are classified as residential and habitable if the geographic area identified by the research team appears to be potentially used for living and/or sleeping by at least one person (e.g. house or hut with roof, and some evidence of being inhabited, such as curtains in windows, livestock). In contrast, plots determined to be residential, but not habitable included those with only unfinished or incomplete structures, animal pens/kraals, or demolished/deserted households. Non-residential plots also included areas with only commercial shops, health facilities, schools, etc.	✓	✓

		Only residential and habitable plots were eligible for household assessments and enumerations.		
residential_habitable	Yes/No indicator for the plot designated residential and habitable.		✓	✓
household_count	Number of households on the plot.		✓	✓
reshab_plot_type	<p>Categorization of the number of households on residential and habitable plots.</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "Single: 1 household"</li> <li>- "Double: 2 households"</li> <li>- "Multi: &gt;2 households"</li> </ul>	Variable set to missing if (1) plot not residential and habitable or (2) plot residential and habitable, but no households found	✓	✓
plot_enumerated	Yes/No indicator for the enumeration of ≥1 household on plot		✓	✓
plot_enrolled	Yes/No indicator for ≥1 enrolled household member on ≥1 household on plot		✓	✓
used_for_ess	Yes/No indicator for ≥1 enrolled household member participating in ESS on a plot originally selected for 20% sample in year	Variable coded to 'Yes' if ≥1 household members participated in the ESS in year 3, and 'No' if (1) no household members participated in the ESS or (2) if no household members participated in either survey	✓	✗

## Households

Variable Name	Description	Details	Availability of Data	
			20% random sample at Year 1 (all pairs)	Remaining 80% sample at Year 3 (pairs 1, 6, and 9)
de_hh_idC	Unique household identifier	De-identified, unique codes were randomly assigned to each household	✓	✓
de_plot_idC	Unique plot identifier	De-identified, unique codes were randomly assigned to each plot	✓	✓
region	Geographical region of community	Classified as 'Northern', 'Southern', or 'Central'	✓	✓
community_rndmN	Community number (1-30)	De-identified	✓	✓
pair_rndmN	Assigned pair number (1-15)	De-identification preserved the 1-15 range (randomly scrambled the pairs) and maintain link to community	✓	✓
intervention	Randomization arm		✓	✓
final_household_status	<p>Categorical variable summarizing household status following all Evaluation Protocol study activities in a given year.</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "Enumerated not enrolled"</li> <li>- "Household not visited"</li> <li>- "Household refused"</li> <li>- "No eligible representative"</li> <li>- "No one home"</li> <li>- If 20% sample: "HH members enrolled in evaluation protocol"</li> <li>- If 80% sample: "HH members participated in ESS"</li> </ul>		✓	✓

<p>hh_enumerated hh_enrolled</p>	<p>Yes/No indicator variables for enumerated and/or enrolled households, respectively.</p>		<p>✓</p>	<p>✓</p>
<p>used_for_ess</p>	<p>Yes/No indicator for ≥1 enrolled household member participating in ESS</p>	<p>Variable coded to ‘Yes’ if ≥1 household member participated in the ESS in year 3, and ‘No’ if (1) no household members participated in the ESS or (2) if no household members participated in either survey</p>	<p>✓</p>	<p>✗</p>
<p>occupancy_status</p>	<p>Occupancy status of households in the 20% sample in year 1 (community pairs 2-15).</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- “Regularly occupied”</li> <li>- “Seasonally occupied”</li> <li>- “Rarely or never occupied”</li> </ul>	<p>For plots determined to be residential and habitable, research staff made up to three attempts to enumerate members of each household on the plot. If, after three attempts, nobody was found at home, the research staff member approached a neighbor or a representative at the Village Development Committee for assistance. This third-party informant, who was required to be 18 years of age or older, was asked if members of the household stayed there regularly, temporarily/seasonally, rarely or never.</p> <p>Information from third-party informants on occupancy was not available for Pair 1 communities but was available for all subsequent community pairs.</p> <p>All enumerated and/or enrolled households were considered regularly occupied. Irrespective of enumeration or enrollment status, if research staff encountered someone at home the</p>	<p>✓</p>	<p>✗</p>

		household was considered regularly occupied.		
member_count enrolled_count	Number of enumerated and enrolled household members, respectively.	Variable member_count set to 0 if household was not enumerated.  Variable enrolled_count set to 0 if household was enumerated but no household members were enrolled.  Both variables set to missing if household neither enumerated nor enrolled.	✓	✓
children_under5 children_5to15	Number of enumerated household members aged younger than 5 years or aged 5 to 15 years, respectively.	Variable set to missing if household was not enumerated.	✓	✓
flooring_type	Categorical variable summarizing the main type of flooring in the household as reported by the household respondent.  Responses: - "Tile wood parquet carpet" - "Cement" - "Dirt/earth"	Derived variable incorporating any responses provided in the free text field option, "Other (please specify)."	✓	✓
water_source	Categorical variable summarizing the main source of water in the household as reported by the household respondent.  Responses: - "Piped indoors" - "Standpipe/tap within plot" - "Neighbor/relative tap or standpipe" - "Communal tap" - "Borehole or protected well" - "Bowser tanker or buy from others"	Derived variable incorporating any responses provided in the free text field option, "Other (please specify)."	✓	✓



	- "River lake or unprotected well"			
toilet_facility	<p>Categorical variable summarizing the main type of toilet facility in the household as reported by the household respondent.</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "Flush toilet within plot"</li> <li>- "Pit latrine within plot"</li> <li>- "Neighbor/community toilet/pit latrine"</li> <li>- "Bush"</li> </ul>	Derived variable incorporating any responses provided in the free text field option, "Other (please specify)."	✓	✓
energy_source	<p>Categorical variable summarizing the main source of energy in the household as reported by the household respondent.</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "Electricity"</li> <li>- "Gas"</li> <li>- "Charcoal/wood/paraffin"</li> <li>- "No cooking done"</li> </ul>	Derived variable incorporating any responses provided in the free text field option, "Other (please specify)."	✓	✓
living_rooms	Continuous variable summarizing the total number of rooms in the household as reported by the household respondent.		✓	✓
cattle_owned goats_owned sheep_owned	Continuous variables summarizing the total number of cattle, goats, and sheep owned by members of the household as reported by the household respondent.		✓	✓
elec_cellphone elec_computer elec_internet elec_landline	Yes/No indicator variables representing household ownership of various electrical appliances as reported by household representative.		✓	✓

elec_radio elec_refrig elec_tv				
transport_animal transport_bicycle transport_cart transport_motorcycle transport_motorveh transport_tractor	Yes/No indicator variables representing household ownership of various forms of transport as reported by household representative.		✓	✓
smaller_meals	Yes/No indicator variable for any household member needing to eat smaller meals due to lack of food.		✓	✓

## Household Members

Variable Name	Description	Details	Availability of data	
			20% random sample at Year 1 (all pairs)	Remaining 80% sample at Year 3 (pairs 1, 6, and 9)
de_subj_idC	Unique household member identifier	De-identified, unique codes were randomly assigned to each enumerated household member, the same identifier is used if the household member enrolls	✓	✓
de_hh_idC	Unique household identifier	De-identified, unique codes were randomly assigned to each household	✓	✓
de_plot_idC	Unique plot identifier	De-identified, unique codes were randomly assigned to each plot	✓	✓
region	Geographical region of community	Classified as 'Northern', 'Southern', or 'Central'	✓	✓
community_rndmN	Community number (1-30)	De-identified	✓	✓
pair_rndmN	Assigned pair number (1-15)	De-identification preserved the 1-15 range (randomly scrambled the pairs) and maintain link to community	✓	✓
intervention	Randomization arm		✓	✓
final_member_status	<p>Derived categorical variable summarizing household member status following all Evaluation Protocol study activities in a given year.</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "Survey participant"</li> <li>- "Age ineligible"</li> <li>- "Refused"</li> <li>- "Absent"</li> <li>- "Unknown"</li> </ul>		✓	✓

	<ul style="list-style-type: none"> <li>- If 20% sample, "Incapacitated or non-citizen"</li> <li>If 80% sample: "Ineligible by other criteria"</li> </ul>			
relation_to_head	Categorical variable summarizing the relationship between the household member and head of the household as reported by household representative.		✓	✓
gender	Gender of household member as reported by household representative.		✓	✓
age_in_years	Age (in years) of household member as reported by household representative (if not enrolled) or member (if enrolled).	If enumerated household member subsequently enrolled in BHS/AHS or ESS, variable was replaced with age based on confirmed date of birth.	✓	✓
age_eligible	Yes/No indicator variable for meeting age eligibility criteria (16-64 years) for Evaluation Protocol study participation.		✓	✓
eligible_member	Yes/No indicator variable for meeting all eligibility criteria (including age, residency, citizenship, and competency) for Evaluation Protocol study participation.		✓	✓
inelig_other	<p>Yes/No indicator variable for being age eligible, but not all meeting one or more additional eligibility criteria.</p> <p>-</p>	Note: For BHS/AHS in year 1, enumerated household members who were not a Botswana citizen or married to a citizen were ineligible for Evaluation Protocol study participation. This requirement was lifted in year 3 for ESS.	✓	✓
inelig_other_reason	Reason for ineligibility among age-eligible enumerated household members	Note: For BHS/AHS in year 1, enumerated household members who were not a Botswana citizen or married to a citizen were ineligible for Evaluation Protocol study participation. This requirement was lifted in year 3 for ESS.	✓	✓

## Survey Data

### Participant and Community Variables

Plot, household, and participant identifiers are available in all datasets.

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
de_subj_idC	Unique participant identifier	De-identified, unique codes were randomly assigned to each enrolled participant	✓	✓	✓
de_hh_idC	Unique household identifier	De-identified, unique codes were randomly assigned to each household	✓	✓	✓
de_plot_idC	Unique plot identifier	De-identified, unique codes were randomly assigned to each plot	✓	✓	✓
region	Geographical region of community	Classified as 'Northern', 'Southern', or 'Central'	✓	✓	✓
community_rndmN	Community number (1-30)	De-identified	✓	✓	✓
pair_rndmN	Assigned pair number (1-15)	De-identification preserved the 1-15 range (randomly scrambled the pairs) and maintain link to community	✓	✓	✓
intervention	Randomization Arm		✓	✓	✓
gender	Gender		✓	✓	✓
age_at_interview	Continuous age at household visit	Derived from date of birth provided on consent form	✓	✓	✓
age5cat	Categorical age variable - 16-24 years - 25-34 years - 35-44 years - 45-54 year		✓	✓	✓

	- 55-64 years				

## Residency and Mobility

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
length_residence	Q.2. How long have you lived in this community?	"Don't want to answer" responses coded to missing.	✓	✓	✓
permanent_resident	Q.3. In the past 12 months, have you typically spent 14 or more nights per month in this community?		✓	✓	✓
intend_residency	Q.4. Do you intend to move out of the community in the next 12 months?		✓	✓	x
nights_away	Q.5. In the past 12 months, in total how many nights did you spend away from this community, including visits to cattle post and lands? [If you don't know exactly, give your best guess]	"Don't want to answer" and "I am not sure" responses coded to missing.	✓	✓	✓
cattle_postlands	Q.6. In the past 12 months, during the times you were away from this community, where were you primarily staying?	<p>Responses provided in the free text field option, "... if other, specify" were either incorporated into existing response categories or as new indicator variables.</p> <p>"Don't want to answer" responses coded to missing.</p> <p>Variable coded to missing if participant previously reported zero nights away during the past 12 months</p>	✓	✓	✓

## Community Engagement

The Community Engagement section was only administered in year 1 for community pairs 2-15.

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
community_engagment	Q.2. How active are you in community activities such as burial society, Motshelo, Syndicate, PTA, VDC (Village Development Committee), Mophato and development of the community that surrounds you?	“Don't want to answer” responses coded to missing.	✓	✗	✗
vote_engagement	Q.3. Did you vote in the last local government election?	“Don't want to answer” and “Not applicable (no election, can't vote)” responses coded to missing.	✓	✗	✗
prob_housing prob_malaria prob_roads prob_schools prob_sewer prob_unemployment prob_water prob_hiv  prob_alcohol_drug prob_crime prob_electricity prob_develop	Q.4. What are the major problems in this neighbourhood?	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is ‘Yes’ for each response selected.  Responses provided in the free text field option, “Q.5. ... if other, specify” were either incorporated into existing response categories or as new indicator variables.  “Don't want to answer” responses coded to missing.	✓	✗	✗



prob_govt prob_healthcare prob_youth prob_teenpreg prob_neighbor prob_environment prob_poverty					
solve_engagement	Q.6. If there is a problem in this neighbourhood, do the adults work together in solving it?	“Don't want to answer” and “Don't know” responses coded to missing.	✓	✗	✗

### Demographics and Education

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
religion_name	Q.2. What is your religion affiliation?	Derived variable incorporating responses provided in the free text field option, "Q.3. ... if 'Other', specify" which matched a pre-specified response provided in the survey.	✓	✗	✓
religious_affill	Yes/No indicator variable set to "Yes" if participant reported any religious affiliation on Q.2. and "No" otherwise.		✓	✗	✓
relig_theogrp	<p>Categorical variable summarizing individual religious affiliations into broad theological groups based on historical and cultural characteristics</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- "African Independent"</li> <li>- "Catholic"</li> <li>- "Eastern churches"</li> <li>- "Pentacostal"</li> <li>- "Small denominations"</li> <li>- "Non-religious"</li> </ul>	Derived variable incorporating responses provided in the free text field option, "Q.3. ... if 'Other', specify" which matched a pre-specified response provided in the survey.	✓	✗	✓
ethnicity	Q.4. What is your ethnic group?	Derived variable incorporating responses provided in the free text field option, "Q.5. ... if 'Other', specify" which matched a pre-specified response provided in the survey.	✓	✗	✗

		“Don't want to answer” responses coded to missing.			
marital_status	Q.6. What is your current marital status?	“Don't want to answer” responses coded to missing.	✓	✓	✓
num_wives	Q.7. WOMEN: How many wives does your husband have (including traditional marriage), including yourself?	Note: There are >0 responses by male participants even though the question was intended to be answered by female participants only.	✓	✓	✓
husband_wives	Q.8. MEN: How many wives do you have, including traditional marriage?	Note: There are >0 responses by female participants even though the question was intended to be answered by male participants only.	✓	✓	✓
live_alone livewith_family livewith_partner livewith_siblings livewith_other	Q.9. Who do you currently live with?	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is ‘Yes’ for each response selected and missing otherwise.  “Don't want to answer” responses coded to missing.	✓	✓	✓

Education and Employment

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
education	Q.2. What is your highest level of education attainment?		✓	✓	✓
job_description	Q.6. Describe the work that you do or did in your most recent job. If you have more than one profession, choose the one you spend the most time doing.	Variable coded to missing if participant is unemployed.	✓	✓	✓
employment_status	<p>Derived variable based on:                      Q.3. Are you currently working?                      Q.5. What is the reason why you are not working?</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- Employed</li> <li>- Unemployed and looking for work</li> <li>- Unemployed, but not looking for work</li> </ul>	<p>Variable coded to “Employed” if working = “Yes”</p> <p>Variable coded to “Unemployed and looking for work” if working = “No” and reason_unemployed = “Waiting to continue agricultural work” or “Unemployed (looking for work)” or “Unemployed (waiting to start new work)”</p> <p>Variable coded to “Unemployed, but not looking for work” if working = “No” and reason_unemployed = “Unable to work (permanently sick or injured)” or “Student/ Apprentice/ Volunteer” or “Housewife/</p>	✓	✓	✓

		Homemaker (not looking for work)” or “Retired” or “Not looking for work”.			
monthly_income	<p>Q.7. In the past month, how much money did you earn from work you did or received in payment [retirement benefits, child maintenance, food basket, etc]?</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- No income</li> <li>- 1-199 pula</li> <li>- 200-499 pula</li> <li>- 500-999 pula</li> <li>- 1,000-4,999 pula</li> <li>- 5,000-10,000 pula</li> <li>- More than 10,000 pula</li> </ul>	<p>Variable coded to ‘No income’ if participant is unemployed and monthly income is not reported.</p> <p>“Don't want to answer” responses coded to missing.</p>	✓	✓	✓

## HIV Testing History, Test Review, and Result Documentation

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
hiv_status_current	<p>Derived variable summarizing participant HIV infection status at each study visit; constructed based on responses and information gathered from the following survey items (in addition to results of rapid in-home HIV testing):</p> <p><i>HIV Testing History:</i>            Q.2. Have you ever been tested for HIV before?            Q.4. Is a record of last [most recent] HIV test [OPD card, Tebelopele, other] available to review?            Q.5. Please tell me the results of your last [most recent] HIV test?            Q.6. Do you have any other available documentation of positive HIV status?</p> <p><i>HIV Test Review:</i>            Q.2. What was the recorded date of the last HIV test?            Q.3. What was the recorded HIV test result?</p>	<p>HIV-positive = participant tested HIV-positive during rapid in-home testing or provided documentation of a previously known HIV-positive status.</p> <p>HIV-negative = participant tested HIV-negative during rapid in-home testing.</p> <p>Missing = participant could not provide documentation of a previously known HIV-positive status and refused rapid in-home testing.</p>	✓	✓	✓

	<p><i>HIV result documentation:</i>                  Q.2. What is the recorded date of this previous HIV test (or of the document that provides supporting evidence of HIV infection)?                  Q.3. What is the type of document used?</p>				
hiv_status_time	<p>Derived variable summarizing HIV infection status at each study visit and timing of participant awareness of status</p>	<p>HIV-positive (newly diagnosed) = participant could not provide documentation of a previously known HIV-positive status and tested HIV-positive during rapid in-home testing.</p> <p>HIV-positive (previously diagnosed) = participant provided documentation of a previously known HIV-positive status</p> <p>HIV-negative = participant tested HIV-negative during rapid in-home testing.</p> <p>Missing = participant could not provide documentation of a previously known HIV-positive status and refused rapid in-home testing.</p>	✓	✓	✓
selfreport_has_tested	<p>Q.2. Have you ever been tested for HIV before?</p>	<p>Note: Self-reported history HIV test irrespective of available supporting documentation (if any) of test.</p> <p>“Don't want to answer” responses coded to missing.</p>	✓	✓	✓
selfreport_has_testrecord	<p>Q4. Record of last test available?</p>		✓	✓	✓

selfreport_when_hivtest	Q.3. When was the last [most recent] time you were tested for HIV?	Note: Self-reported timing of most recent HIV test irrespective of available supporting documentation (if any) of test date.  “Don't want to answer” and “I am not sure” responses coded to missing.	✓	✓	✓
selfreporte_hiv_result	Q.5. Please tell me the results of your last [most recent] HIV test?	Note: Self-reported result of most recent HIV test irrespective of available supporting documentation (if any) of test result.  “Don't want to answer” and “I am not sure” responses coded to missing.	✓	✓	✓
prev_hivresult_days	Derived variable of the number of days from enrollment date to date of most recent HIV test with supporting documentation.		✓	✓	✓
prev_hivresult	Derived variable of result of most recent HIV test with supporting documentation.		✓	✓	✓
prev_hivpos_doctype	Type of documentation of previous HIV result	From ‘result_doc_type’ on survey	✓	✓	✓
viral_load	Continuous viral load result from blood draw at study visit.	Note: Per protocol, viral load was measured in all HIV-positive participants at enrollment and study end.  Viral load was also measured in year 2 for participants newly testing positive (i.e. seroconverters) between enrollment and AHS-1.	✓	✓	✓
cd4_value	Continuous CD4 cell count from in-home point-of-care CD4 testing at study visit.	Note: Per protocol, CD4 cell count was measured in (1) HIV-positive participants not on ART at enrollment and (2) participants newly testing positive during follow-up (i.e. seroconverters) without evidence of ART initiation.	✓	✓	✓



## HIV Tested and Untested

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
prev_hivtest_num	Q.2. How many times before today have you had an HIV test?		✓	✗	✗
prev_hivresult_locale	Q.3. Where were you tested for HIV, the last [most recent] time you were tested?	Derived variable incorporating responses provided in the free text field option, "Q.4. ... if 'Other', specify" which matched a pre-specified response provided in the survey.  "Don't want to answer" and "I am not sure" responses coded to missing.	✓	✓	✓
prev_hivtest_reason	Q.5. Not including today's HIV test, which of the following statements best describes the reason you were tested the last [most recent] time you were tested before today?	"Don't want to answer" and "I am not sure" responses coded to missing.	✓	✗	✗
hiv_pills	Q.6. Have you ever heard about treatment for HIV with pills called antiretroviral therapy or ARVs [or HAART]?	"Not sure" responses coded to missing.	✓	✗	✗
arvs_hiv_test	Q.7. Do you believe that treatment for HIV with antiretroviral therapy (or ARVs) can help HIV-positive people to live longer?	"Not sure" responses coded to missing.	✓	✗	✗
hivuntested_reason	Q.2. If you were not tested for HIV in the 12 months prior to	"Don't want to answer" and "I am not sure" responses coded to missing.	✓	✗	✗

	today, what is the main reason why not?				
--	---	--	--	--	--

## HIV Care and Adherence and Medical Care

HIV Care and Adherence and Medical Care questions were administered to HIV-positive participants who knew of their HIV-positive status before enrollment into the Evaluation Protocol.

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
arv_status_current	<p>Derived variable summarizing participant ART status (among HIV-positive participants) at each study visit; constructed based on responses and information gathered from the following survey items (in addition to results of rapid in-home HIV testing):</p> <p><i>HIV Care and Adherence:</i> Q.7. Have you ever taken any antiretroviral therapy (ARVs) for your HIV infection? [For women: Do not include treatment that you took during pregnancy to protect your baby from HIV]</p> <p>Q.11. Are you currently taking antiretroviral therapy (ARVs)?</p> <p>Q.12. Is there evidence [OPD card, tablets, masa number]</p>	<p>On ARVs = Documented evidence of current ARV use.</p> <p>ARV-naïve = (1) Participant reports no current or past use of ART or (2) participant reports current use, but unable to provide supporting documentation.</p> <p>ARV-defaulter = Participant reports past use of ART, but no current use.</p> <p>Variable coded to missing for HIV-negative participants.</p>	✓	✓	✓

	<p>that the participant is on therapy?</p> <p>Q.13. Which clinic facility are you already receiving therapy from?</p> <p>Q.10. When did you first start taking antiretroviral therapy (ARVs)?</p>				
arv_duration_days	<p>Derived variable (among HIV-positive participants with documented current use of ARVs) summarizing the total number of days since initiation of ARVs to current study visit.</p> <p>Q.10. When did you first start taking antiretroviral therapy (ARVs)?</p>		✓	✓	✓
medical_care_hiv	<p>Q.3. Have you ever received HIV-related medical or clinical care, for such things as a CD4 count (masole), IDCC/ PMTCT registration, additional clinic-based counseling?</p>	<p>“Don't want to answer” responses coded to missing.</p>	✓	✓	✓
no_medical_care	<p>Q.4. What is the main reason you have not received HIV-related medical or clinical care?</p>	<p>Derived variable incorporating responses provided in the free text field option, “ ... if ‘Other’, specify” which matched a pre-specified response provided in the survey.</p> <p>“Don't want to answer” and “I am not sure” responses coded to missing.</p>	✓	✓	✓

last_hivcare_days	Days from enrollment to date captured by: Q.3. When did you last (most recently) receive HIV-related medical care for such things as a CD4 count (masole), IDCC/ PMTCT registration, additional clinic-based counseling?	Note: Data are self-reported by participant and were not confirmed or verified by supporting documentation  “Don't want to answer” responses coded to missing.	✓	✓	✓
lowest_cd4	Q.4. What was your lowest CD4 (masole) count that was ever measured?	“Don't want to answer” and “I am not sure” responses coded to missing.	✓	✓	✓
ever_recommended_arv	Q.6. Have you ever been recommended by a doctor/nurse or other healthcare worker to start antiretroviral therapy (ARVs), a combination of medicines to treat your HIV infection? [common medicines include: combivir, truvada, atripla, nevirapine]	“Don't want to answer” responses coded to missing.	✓	✓	✓
ever_taken_arv	Q.7. Have you ever taken any antiretroviral therapy (ARVs) for your HIV infection? [For women: Do not include treatment that you took during pregnancy to protect your baby from HIV]	“Don't want to answer” responses coded to missing.	✓	✓	✓
why_no_arv	Q.8. What was the main reason why you have not started ARVs?	Derived variable incorporating responses provided in the free text field option, “ ... if ‘Other’, specify” which matched a pre-specified response provided in the survey.	✓	✓	✓

		“Don't want to answer” and “I am not sure” responses coded to missing.			
on_arv	Q.11. Are you currently taking antiretroviral therapy (ARVs)?	“Don't want to answer” responses coded to missing.	✓	✓	✓
arv_evidence	Q.12. Is there evidence [OPD card, tablets, masa number] that the participant is on therapy?	Variable coded to missing for (1) HIV-negative participants and (2) HIV-positive participants who report no current or past use of ARVs.	✓	✓	✓
clinic_receiving_from	Q.13. Which clinic facility are you already receiving therapy from?		✓	✓	✓
arv_stop_days	Days from enrollment to date captured by: Q.15. When did you stop taking ARV's?	Note: Data are self-reported by participant and were not confirmed or verified by supporting documentation	✓	✓	✓
arv_stop	Q.16. What was the main reason why you stopped taking ARVs?	Note: Variable contains a high level of missingness (only intended for HIV-positive participants currently defaulted from ARVs).	✓	✓	✓
adherence_4_day	Q.18. During the past 4 days, on how many days have you missed taking all your doses of antiretroviral therapy (ART)?	“Don't want to answer” responses coded to missing.	✓	✓	✓
adherence_4_wk	Q.19. Thinking about the past 4 weeks, on average, how would you rate your ability to take all your medications as prescribed?	“Don't want to answer” responses coded to missing.	✓	✓	✓

## Circumcision

Circumcision questions administered to male participants only.

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
circumcised	Q.2. Are you circumcised?	“Not sure” responses coded to missing.	✓	✓	✓
circumcision_days	Days from enrollment to reported date of circumcision		✓	✓	✓
circ_benefit_hygiene circ_benefit_reducedhiv circ_benefit_reducedcancer circ_benefit_reducedsti	Q.3. What do you believe are the health benefits of male circumcision? (Indicate all that apply.)	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is ‘Yes’ for each response selected and missing otherwise  Variable coded to missing if participant reported being uncircumcised.  “Don't want to answer” and “I am not sure” responses coded to missing.	✓	✓	✓
circumcised_age	Q.4. At what age were you circumcised?	Derived from ‘when_circ’ and ‘age_unit_circ’, separated into categories. - If numeric age given but units missing then value is missing	✓	✓	✓
circumcised_location	Q.6. Where were you circumcised?	Derived variable incorporating responses provided in the free text field option, “... if ‘Other’, specify” which matched a pre-specified response provided in the survey.  “Don’t want to answer” and “I am not sure” responses coded to missing.	✓	✓	✓

circumcised_reason	Q.8. What was the main reason why you were circumcised?	Derived variable incorporating responses provided in the free text field option, "... if 'Other', specify" which matched a pre-specified response provided in the survey.  "Don't want to answer" and "I am not sure" responses coded to missing.	✓	✓	✓
uncirc_benefit_hygiene uncirc_benefit_reducedhiv uncirc_benefit_reducedcancer uncirc_benefit_reducedsti	Q.3. What do you believe are the health benefits of male circumcision? (Indicate all that apply.)	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is 'Yes' for each response selected and missing otherwise  Variable coded to missing if participant reported being circumcised.  "Don't want to answer" and "I am not sure" responses coded to missing.	✓	✓	✓
uncirc_reason	Q.4. What is the main reason that you have not yet been circumcised?.	Derived variable incorporating responses provided in the free text field option, "... if 'Other', specify" which matched a pre-specified response provided in the survey.  "Don't want to answer" and "I am not sure" responses coded to missing.	✓	✓	✓
uncirc_future	Q.6. Would you ever consider being circumcised in the future?	Variable coded to missing if participant reported being circumcised.  "Not sure" responses coded to missing.	✓	✓	✓
uncirc_future_reason	Q.7. Which of the following might increase your willingness to be circumcised the most?	Variable coded to missing if participant reported being circumcised.  "Don't want to answer" and "I am not sure" responses coded to missing.	✓	✓	✓



<p>uncirc_services_free</p>	<p>Q.8. Were you aware that circumcision services are provided free of charge at most health facilities?</p>	<p>Variable coded to missing if participant reported being circumcised.  “Don't want to answer” responses coded to missing.</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>
<p>uncirc_aware_free</p>	<p>Q.9. Where did you learn that circumcision services were available free at most health facilities?</p>	<p>Variable coded to missing if participant reported being circumcised.  “Don't want to answer” and “I am not sure” responses coded to missing.</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>

Sexual Behavior

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
ever_sex	Q.2. In your lifetime, have you ever had sex with anyone (including your spouse, friends, or someone you have just met)?	Note: Variable based on Q.2. and any relevant information provided in other or subsequent survey questions, including all partner-specific information collected on the participant’s three most recent partners.  “Don't want to answer” and “I am not sure” responses coded to missing.	✓	✓	✓
partners_lifetime	Continuous variable of number of sexual partners during participant’s lifetime.  Q.3. In your lifetime, how many different people have you had sex with? Please remember to include casual and once-off partners (prostitutes and truck drivers) as well as long-term partners (spouses, boyfriends /girlfriends) [If you can't recall the exact number, please give a best guess]	Note: Variable based on Q.3. and any relevant information provided in other or subsequent survey questions, including all partner-specific information collected on the participant’s three most recent partners.  Variable coded to missing if participant reported no prior sexual activity (lifetime).  “Don't want to answer” responses coded to missing.	✓	✓	✓
partners_12mos	Continuous variable of number of sexual partners during the past 12 months	Note: Variable based on Q.4. and any relevant information provided in other or subsequent survey questions, including all partner-specific information collected on the participant’s three most recent partners.	✓	✓	✓

	<p>Q.4. In the past 12 months, how many different people have you had sex with? Please remember to include casual and once-off partners (prostitutes and truck drivers) as well as long-term partners (spouses, boyfriends /girlfriends) [If you can't recall the exact number, please give a best guess]</p>	<p>Variable coded to zero if participant reported prior sexual activity (lifetime), but no activity in the past 12 months.</p> <p>Variable coded to missing if participant reported no prior sexual activity (lifetime).</p> <p>“Don't want to answer” responses coded to missing.</p>			
sex_12mos	<p>Yes/No indicator variable for any sexual activity during the past 12 months.</p>	<p>Variable coded ‘Yes’ if participant reported prior sexual activity (lifetime) and <math>\geq 1</math> partner in the past 12 months.</p> <p>Variable coded ‘No’ if participant reported prior sexual activity (lifetime), but no activity in the past 12 months.</p> <p>Variable coded to missing if participant reported no prior sexual activity (lifetime).</p>	✓	✓	✓
sex_outside_12mos	<p>Q.5. In the past 12 months, did you have sex with somebody living outside of the community?</p>	<p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.</p> <p>“Don't want to answer” responses coded to missing.</p>	✓	✓	✓
age_first_sex	<p>6. How old were you when you had sex for the first time? [If you can't recall the exact age, please give a best guess]</p>	<p>Variable coded to missing if participant reported no prior sexual activity (lifetime).</p> <p>“Don't want to answer” responses coded to missing.</p>	✓	✓	✓

<p>condom_lastsex</p>	<p>Q.7. During the last [most recent] time you had sex, did you or your partner use a condom?</p>	<p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.</p> <p>“Don't want to answer” responses coded to missing.</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>
<p>alcohol_lastsex</p>	<p>You or partner used alcohol during most recent sex</p>	<p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.</p> <p>“Don't want to answer” responses coded to missing.</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>
<p>intergenerational_sex_12mos</p>	<p>Derived variable summarizing age difference between participant and participant’s three most recent sexual partners (past 12 months) as reported in the following partner-specific survey question:</p> <p>Q.9. To the best of your knowledge, how old is this person?</p> <ul style="list-style-type: none"> <li>• less or equal to 18 years old</li> <li>• 19-29 years old</li> <li>• 30-39 years old</li> <li>• 40-49 years old</li> <li>• 50-59 years old</li> </ul>	<p>“Same age category” = participant’s age within the same 10-year age category as all of their three most recent partners.</p> <p>“Partner was older” = one or more of the participant’s three most recent partners was reported to be in a different and older 10-year age category as compared to the participant.</p> <p>“Partner was older” = one or more of the participant’s three most recent partners was reported to be in a different and younger 10-year age category as compared to the participant.</p> <p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>
<p>condom_12mos</p>	<p>Yes/No indicator variable summarizing any reported</p>	<p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>

	<p>condom non-use during the past 12 months as reported in the following partner-specific survey question:</p> <p>Q.17. When you have [had] sex with this partner, how often do you or your partner use a condom?</p>	<p>sexual activity (lifetime), but no activity in the past 12 months.</p> <p>“Don't want to answer” responses coded to missing.</p>			
<p>concurrency_12mos</p>	<p>Yes/No indicator variable summarizing any participation in multiple, concurrent (overlapping) partners during past 12 months as reported in the following partner-specific survey questions:</p> <p>Q.3. When was the last [most recent] time you had sex with this person (how long ago)?</p> <p>Q.5. When was the first time you had sex with this person [how long ago]?</p> <p>Q.7. Do you expect to have sex with this person again?</p> <p>Q.10. Over the past 12 months, during the time you were having a sexual relationship with this person, did YOU</p>	<p>Variable coded to “No” if participant reported (1) no partners, (2) one partner, or (3) <math>\geq 2</math> non-overlapping partners during the past 12 months</p> <p>Variable coded to “Yes” is participant reported <math>\geq 2</math> overlapping partners during the past 12 months</p> <p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or responded “Don’t want to answer” on all supporting variables.</p>	<p>✓</p>	<p>✓</p>	<p>✓</p>

	have sex with other people (including husband/wife)?				
exchange_12mos	<p>Yes/No indicator variable summarizing any participation in transactional sex during past 12 months as reported in the following partner-specific survey questions:</p> <p>Q.8. What type of relationship do you have with this person? [Response = 'Commercial sex worker']</p> <p>Q.11. Have you received money, transport, food/drink, or other goods in exchange for sex from this partner?</p>	<p>Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months, or responded "Don't want to answer" on all supporting variables.</p>	✓	✓	✓

Sexual Behavior: Three most recent partners, past 12 months

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
partner1_inside_comm partner1_outside_comm  partner2_inside_comm partner2_outside_comm  partner3_inside_comm partner3_outside_comm	Q.2. Over the past 12 months, where has this sexual partner lived to the best of your knowledge?	Values are either 'Yes' or missing: partnerX_inside_comm='Yes' indicates partner X lived within the current study community partnerX_outside_comm='Yes' indicates partner X lived outside the current study community	✓	✓	✓
firstpartner_sex_current secondpartner_sex_current thirdpartner_sex_current	Expect to have sex with this partner again	Possible responses: 'Yes', 'No', or missing	✓	✓	✓
firstpartner_lastsex firstpartner_firstsex firstpartner_futuresex  secondpartner_lastsex secondpartner_firstsex secondpartner_futuresex  thirdpartner_lastsex thirdpartner_firstsex thirdpartner_futuresex	Q.3. When was the last [most recent] time you had sex with this person (how long ago)?  Q.5. When was the first time you had sex with this person [how long ago]?  Q.7. Do you expect to have sex with this person again?	Note: These variables are not included in the provided data. However, they were part of the set of variables used to construct derived variable "concurrency_12mos" (described above).  Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  "Don't want to answer" responses coded to missing.	x	x	x
firstpartner_relationship secondpartner_relationship thirdpartner_relationship	Q.8. What type of relationship do you have with this person?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.	✓	✓	✓

		“Don't want to answer” responses coded to missing.			
firstpartner_age secondpartner_age thirdpartner_age	Q.9. To the best of your knowledge, how old is this person? <ul style="list-style-type: none"> <li>• less or equal to 18 years old</li> <li>• 19-29 years old</li> <li>• 30-39 years old</li> <li>• 40-49 years old</li> <li>• 50-59 years old</li> </ul>	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  “Don't want to answer” and “Not sure” responses coded to missing.	✓	✓	✓
firstpartner_goods_exchange secondpartner_goods_exchange thirdpartner_goods_exchange	Q.11. Have you received money, transport, food/drink, or other goods in exchange for sex from this partner?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  “Don't want to answer” responses coded to missing.	✓	✓	✓
firstpartner_sex_freq secondpartner_freq thirdpartner_sex_freq	Q.12. During the last 3 months [of your relationship, if it has ended] how many times did you have sex with this partner?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  “Don't want to answer” responses coded to missing.	✓	✓	✓
firstpartner_hivstatus secondpartner_hivstatus thirdpartner_hivstatus	Q.13. What is this partner's HIV status?	Note: N=79 participants reported that their partner was HIV-negative, but taking ART  Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months	✓	✓	✓



		“Don't want to answer” and “Not sure” responses coded to missing.			
firstpartner_onarv secondpartner_onarv thirdpartner_onarv	Q.15. Is this partner taking antiretroviral treatment?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months  “Don't want to answer” and “Not sure” responses coded to missing.	✓	✓	✓
firstpartner_disclose_hiv secondpartner_disclose_hiv thirdpartner_disclose_hiv	Q.16. Have you told this partner your HIV status?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  “Don't want to answer” responses coded to missing.	✓	✓	✓
firstpartner_condom_freq secondpartner_condom_freq thirdpartner_condom_freq	Q.17. When you have [had] sex with this partner, how often do you or your partner use a condom?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  “Don't want to answer” responses coded to missing.	✓	✓	✓
firstpartner_concurrent secondpartner_concurrent thirdpartner_concurrent	Q.18. To the best of your knowledge, did he/she ever have other sex partners while you two were having a sexual relationship?	Variable coded to missing if participant reported (1) no prior sexual activity (lifetime) or (2) prior sexual activity (lifetime), but no activity in the past 12 months.  “Don't want to answer” responses coded to missing.	✓	✓	✓



## Reproductive Health and Pregnancy

Reproductive Health and Pregnancy questions administered to sexually active (i.e. ever had sex) female participants only.

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
number_children	Q.2. How many children have you given birth to? Please include any children that may have died at (stillbirth) or after birth. Do not include any current pregnancies or miscarriages that occur early in pregnancy (prior to 20 weeks).		✓	✓	✓
menopause	Q.3. Have you reached menopause (more than 12 months without a period)?		✓	✓	✓
familyplan_abstain familyplan_condoms_consist familyplan_condoms_inconsist familyplan_diaphram familyplan_iud familyplan_contra_inject familyplan_contra_oral familyplan_rhythm familyplan_sterilization familyplan_withdrawal familyplan_other	Q.4. In the past 12 months, have you used any methods to prevent pregnancy?	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is 'Yes' for each response selected and missing otherwise  Derived variable incorporating responses provided in the free text field option, "... if 'Other', specify" which matched a pre-specified response provided in the survey.  "Don't want to answer" responses coded to missing.	✓	✓	✓
currently_pregnant	Q.6. Are you currently pregnant?	"Not sure" responses coded to missing.	✓	✓	✓

anc_last_pregnancy	Q.5. During your last pregnancy (not current pregnancy) did you go for antenatal care?	“Don't want to answer” responses coded to missing.	✓	✓	✓
hiv_last_pregnancy	Q.6. During your last pregnancy (not current pregnancy) were you tested for HIV?	“Not sure” responses coded to missing.	✓	✓	✓
more_children	Q.2. Do you wish to have a child now or in the future?	Variable only asked of female participants reporting no current pregnancy.	✓	✓	✓

Medical Diagnoses

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
dx_cancer_ever dx_tb_ever dx_heartdisease_ever dx_hivillness_ever dx_wasting_ever dx_diarrhea_ever dx_yeastinf_ever dx_pneumonia_ever dx_pcp_ever	Q.2. Do you recall or is there a record of having any of the following serious illnesses?	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is ‘Yes’ for each response selected and missing otherwise.	✓	✓	✓

dx_herpes_ever					
dx_cancer_lastyr dx_tb_lastyr dx_heartdisease_lastyr dx_hivillness_lastyr dx_wasting_lastyr dx_diarrhea_lastyr dx_yeastinf_lastyr dx_pneumonia_lastyr dx_pcp_lastyr dx_herpes_lastyr	Q.2. Do you recall or is there a record of having any of the following serious illnesses?	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is 'Yes' for each response selected and date of diagnosis was within past 12 months and missing otherwise.	✓	✓	✓
cancer_type	Q.3. [Interviewer:] What is the cancer diagnosis as recorded?	"Don't want to answer" responses coded to missing.	✓	✓	✓
dx_heartdisease_mi dx_heartdisease_failure dx_heartdisease_stroke	Q.3. [Interviewer:] What is the heart disease or stroke diagnosis as recorded?	Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is 'Yes' for each record identified and missing otherwise.  Derived variable incorporating responses provided in the free text field option, "... if 'Other', specify" which matched a pre-specified response provided in the survey.  "Don't want to answer" responses coded to missing.	✓	✓	✓
tb_type	Q.3. [Interviewer:] What is the tuberculosis diagnosis as recorded?	"Don't want to answer" responses coded to missing.	✓	✓	✓
tbsx_cough tbsx_fever tbsx_lymphnodes tbsx_coughblood tbsx_nightswets tbsx_weightloss	Q.2. – Q.7. Does the participant currently have a [SYMPTOM] that has lasted for more than 2 weeks?	Series of indicator variables derived from a series of survey questions which allowed participant to select multiple responses. Value is 'Yes' for each response and 'No' otherwise.	✓	✓	✓



### Substance Use

Note: Substance use data are not available for ESS participants in community pair 5.

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
alcohol	Q.2. In the past month, how often did you consume alcohol?	“Don't want to answer” responses coded to missing.	✓	✗	✓
smoke	Q.3. Do you currently smoke any tobacco products, such as cigarettes, cigars, or pipes?	“Don't want to answer” responses coded to missing.	✓	✗	✓

## Stigma and Stigma Opinion

All variables in these sections have the same set of responses:

- strongly disagree
- disagree
- uncertain
- agree
- strongly agree

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
anticipate_stigma	Q.2. Would you/have you been hesitant to take HIV test due to fear of people's reactions if you tested HIV positive?	Note: Question administered to HIV-negative participants only.  "Don't want to answer" responses coded to missing.	✓	✗	✗
enacted_shame_stigma	Q.3. I would be ashamed if someone in my family had HIV.	Note: Question administered to HIV-negative participants only.  "Don't want to answer" responses coded to missing.	✓	✗	✗
saliva_stigma	Q.4. I fear I could contract HIV if I come into contact with saliva of person living with HIV.	Note: Question administered to HIV-negative participants only.  "Don't want to answer" responses coded to missing.	✓	✗	✗
teacher_stigma	Q.5. I think that if a teacher is living with HIV but not sick, he/she should be allowed to continue teaching.	Note: Question administered to HIV-negative participants only.  "Don't want to answer" responses coded to missing.	✓	✗	✗



children_stigma	Q.6. Children living with HIV should be able to attend school with HIV-negative children.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗
test_community_stigma	Q.2. People are hesitant to take an HIV test due to fear of others reactions if they test HIV-positive.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗
gossip_community_stigma	Q.3. People talk badly about people living with or thought to be living with HIV.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗
respect_community_stigma	Q.4. People living with or thought to be living with HIV lose respect or standing.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗
enacted_verbal_stigma	Q.5. People living with or thought to be living with HIV are verbally insulted, harassed, and/or threatened.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗
enacted_physical_stigma	Q.6. People living with or thought to be living with HIV are sometimes physically assaulted.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗
enacted_family_stigma	Q.7. People living with or thought to be living with HIV are now more accepted by others.	Note: Question administered to HIV-negative participants only.	✓	✗	✗

		“Don't want to answer” responses coded to missing.			
fear_stigma	Q.8. People living with HIV are less able to financially support themselves and their families.	Note: Question administered to HIV-negative participants only.  “Don't want to answer” responses coded to missing.	✓	x	x
internalize_stigma	Q.2. I think less of myself	Note: Question administered to HIV-positive participants only.  “Don't want to answer” responses coded to missing.	✓	x	x
ashamed_stigma	Q.3. I felt ashamed because of having HIV.	Note: Question administered to HIV-positive participants only.  “Don't want to answer” responses coded to missing.	✓	x	x
friend_stigma	Q.4. I feat that if I disclosed my HIV status to my friends they would lose respect for me.	Note: Question administered to HIV-positive participants only.  “Don't want to answer” responses coded to missing.	✓	x	x
family_stigma	Q.5. I feat that if I disclosed my HIV status to my family they would exclude me from usual family activities.	Note: Question administered to HIV-positive participants only.  “Don't want to answer” responses coded to missing.	✓	x	x
enacted_talk_stigma	Q.6. People have talked bad about me.	Note: Question administered to HIV-positive participants only.  “Don't want to answer” responses coded to missing.	✓	x	x
enacted_respect_stigma	Q.7. I have lost respect or standing in the community.	Note: Question administered to HIV-positive participants only.	✓	x	x

		“Don't want to answer” responses coded to missing.			
enacted_job_stigma	Q.8. I have lost a job because of having HIV.	Note: Question administered to HIV-positive participants only.  “Don't want to answer” responses coded to missing.	✓	✗	✗

Access to Care

Q.7-Q.12 have the same set of responses:

- Strongly agree
- Agree
- Uncertain
- Disagree
- Strongly disagree

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
access_care	<p>Q.3. In the past year, where do you most often get care when you or someone in your family is sick or hurt?</p> <p>Responses:</p> <ul style="list-style-type: none"> <li>- Public or government</li> <li>- Private health facility</li> <li>- Traditional, faith or religious healer/doctor</li> <li>- Pharmacy</li> <li>- Community health worker</li> <li>- OTHER</li> </ul>		✓	✗	✗
medicare_private medicare_public medicare_traditional medicare_communityhealth	<p>Q.5. In the past year where else have you received medical care or treatment.</p>	<p>Series of indicator variables derived from survey question which allowed participant to select multiple responses. Value is 'Yes' for each response selected.</p> <p>"Don't want to answer" responses coded to missing.</p>	✓	✗	✗

overall_access	Q.7. If I need medical care, I can get seen by an appropriate health professional without any trouble.		✓	✗	✗
emergency_access	Q.8. It is hard for me to get medical care in an emergency		✓	✗	✗
expensive_access	Q.9. Sometimes I go without the medical care I need because it is too expensive.		✓	✗	✗
convenient_access	Q.10. Places where I can get medical care are very conveniently located		✓	✗	✗
whenever_access	Q.11. I am able to get medical care whenever I need to		✓	✗	✗
local_hiv_care	Q.12. Would you be willing to come to the clinic within your community to receive HIV care and treatment available locally?		✓	✗	✗

## Primary Endpoint Data

Variable Name	Description	Details	Availability of data		
			BHS (years 1 and 2)	AHS (years 2 and 3)	ESS (year 3)
incidence_cohort	Derived Yes/No indicator variable identifying sample of N=8,551 HIV-negative participants who enrolled in the Evaluation Protocol in year 1, completed $\geq 1$ follow-up visit(s) in years 2 and/or 3, and therefore, form the HIV incidence cohort.	Note: HIV-negative participants who enrolled in year 2 (with or without a follow-up visit in year 3) were not included in the HIV incidence cohort.	✓	✓	✗
endpoint_seroconvert	Yes/No indicator variable for acquisition of HIV infection at study visit.	Variable coded to missing for participants not included in the HIV incidence cohort.	✓	✓	✗
endpoint_persontime	Derived continuous variable summarizing the total estimated person-time (in days) spent at risk for HIV infection over the study period.	<p>Person-time at risk for participants who did not acquire HIV infection during study follow-up was calculated as the difference (in days) between date of enrollment and date of last follow-up visit.</p> <p>Person-time at risk for participants who acquired HIV infection was calculated as the difference (in days) between date of enrollment and the mid-point of the date of last HIV-negative test and first HIV-positive test.</p> <p>Variable coded to missing for participants not included in the HIV incidence cohort.</p>	✓	✓	✗

coverage_baseline_20sample	Derived Yes/No indicator variable to identify the sample used for population coverage levels in community pairs 5, 7, and 13 at <i>baseline</i> .	Note: This variable should be used to subset the relevant data file ("LONGCOHORT_YR1_20SAMP") and analyze change in uptake of HIV prevention interventions as pre-specified by the Evaluation Protocol.	✓	✗	✗
endpoint_coverage_htc	Yes/No indicator variable for uptake of HIV testing and counseling among all HIV-positive participants at study visit.	Calculated among all enrolled participants as the % HIV-tested or diagnosed as HIV-positive = $\frac{\text{(HIV-negative participants with documented negative test result within past 12 months or HIV-positive participants with prior knowledge of positive status)}}{\text{(Enrolled participants)}}$ Note: Per protocol, >99% of participants enrolled in the Evaluation Protocol who completed AHS in years 2 and 3 were re-tested for HIV infection in both arms; thus, coverage of HIV testing and counseling, including achievement of UNAIDS 90-90-90 targets cannot be analyzed.	✓	✗	✓
endpoint_coverage_hivpos	Yes/No indicator variable for prior knowledge of HIV-positive status among all HIV-positive participants at study visit.	Calculated among all HIV-positive participants as the % diagnosed HIV-positive = $\frac{\text{(HIV-positive participants with prior knowledge of positive status)}}{\text{(HIV-positive participants)}}$ Note: Per protocol, >99% of participants enrolled in the Evaluation Protocol who completed AHS in years 2 and 3 were re-tested for HIV infection in both arms; thus, coverage of HIV testing and counseling and achievement of UNAIDS 90-90-90 targets cannot be analyzed.	✓	✗	✓

endpoint_coverage_onart	Yes/No indicator variable for uptake of ART among all HIV-positive participants at study visit.	Calculated among all HIV-positive participants as the % on treatment = (HIV-positive participants with prior knowledge of positive status currently receiving ART) / (HIV-positive participants)	✓	✓	✓
endpoint_coverage_vlsupp	Yes/No indicator variable for viral suppression among all HIV-positive participants at study visit.	Calculated among all HIV-positive participants as the % virally suppressed (combined overall UNAIDS target) = (HIV-positive participants with prior knowledge of their HIV-positive status receiving ART who have HIV-1 RNA <400 copies per mL) / (HIV-positive participants)	✓	✓	✓
mc_cohort_20samp	Yes/No indicator variable for inclusion in circumcision analysis				
endpoint_coverage_mc	Yes/No indicator variable for uptake of male circumcision at study visit.	Calculated among all HIV-negative participants aged 16-49 years as the % circumcised = (HIV-negative male participants aged 16–49 years self-reported to be circumcised) / (HIV-negative male participants aged 16–49 years)	✓	✓	✓
endpoint_first_90	Yes/No indicator variable for prior knowledge of HIV-positive status (i.e. diagnosed) among all HIV-positive participants at study visit.	Calculated among all HIV-positive participants as the % diagnosed HIV-positive = (HIV-positive participants with prior knowledge of positive status) / (HIV-positive participants)  Note: Per protocol, >99% of participants enrolled in the Evaluation Protocol who completed AHS in years 2 and 3 were re-tested for HIV infection in both arms; thus, coverage of HIV testing and counseling, including achievement of UNAIDS 90-90-90 targets cannot be analyzed.	✓	x	✓



endpoint_second_90	Yes/No indicator variable for uptake of ART among diagnosed persons at study visit.	Calculated among diagnosed HIV-positive participants as the % on treatment = (HIV-positive participants with prior knowledge of positive status currently receiving ART) / (diagnosed HIV-positive participants)	✓	✓	✓
endpoint_third_90	Yes/No indicator variable for viral suppression among those diagnosed and on ART at study visit.	Calculated among diagnosed HIV-positive participants on treatment as the % virally suppressed = (HIV-positive participants with prior knowledge of their HIV-positive status receiving ART who have HIV-1 RNA <400 copies per mL) / (diagnosed HIV-positive participants on ART)	✓	✓	✓
endpoint_overall_90	Yes/No indicator variable for viral suppression among all HIV-positive persons (irrespective of diagnosis or ART status) at study visit.	Calculated among all HIV-positive participants as the % virally suppressed (combined overall UNAIDS target) = (HIV-positive participants with prior knowledge of their HIV-positive status receiving ART who have HIV-1 RNA <400 copies per mL) / (HIV-positive participants)	✓	✓	✓
endpoint_death	Yes/No indicator variable for death during study follow-up.	Note: Death data available for participants enrolled in Evaluation Protocol in years 1 and 2 only.	✓	✓	✗
death_days	Days from enrollment to documented date of death during study follow-up.	Note: Death data available for participants enrolled in Evaluation Protocol in years 1 and 2 only.	✓	✓	✗
death_primary_cause	Derived variable of the primary cause of death as reported by family member(s) and/or recorded on the death certificate.	Note: Death data available for participants enrolled in Evaluation Protocol in years 1 and 2 only.	✓	✓	✗

death_group_cause	Derived variable summarizing the cause of death according to the following categories:  Any Infectious Cause Any Non-infectious Cause Trauma Unknown	Note: Death data available for participants enrolled in Evaluation Protocol in years 1 and 2 only.	✓	✓	✗
-------------------	---	--	---	---	---