

Dataset Property (2 of 2)	
Dataset	PRECOHORT
Dataset Label	BASELINE DATASET (PRE-COHORT)
Purpose	ANALYSIS DATASET FOR PRE-COHORT SURVEY AT T1, INCLUDING KEY DEMOGRAPHIC VARIABLES AT T1.
Dataset Location	/trials/epi/shims001/analysis/manuscripts/share/live/adata
Source Code	/TRIALS/EPI/SHIMS001/ANALYSIS/MANUSCRIPTS/SHARE/LIVE/CODE/CREATE_ADATA/D_PRECOHORT.SAS
Source Dataset(s)	HC/AHA/REQ/RQ/EQ/HTR/NIC/NRL/ENR FORMS. PLEASE REFER TO SOURCE CODE AND /CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.
Source Dataset Type	SAS
Key Variable(s)	PTID
Structure	BY PTID
Number of Variables	56

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
1	ART	ART use	ART USE AT T1	COALESCE(RQCTART, EQCTART)	Numeric	YNDKR	1 = Yes; 2 = NO; 3 = DK; 4 = REF	
2	EA	EA	EA NUMBER	SWAZILAND EA USED FOR SAMPLING	Numeric	BEST12.		
3	HIVdetail	HIV status (detailed)	DETAILED CATEGORIZATION OF HIV STATUS AT T1	IF HIVPOS = 0 THEN HIVDETAIL = 0; IF HIVPOS = 1 THEN DO; HIVDETAIL = HIVPOS; /* PREVIOUSLY DIAGNOSED*/ IF SELFHIV IN (0, 2) THEN HIVDETAIL = 2; /*NEWLY DIAGNOSED*/ END;	Numeric	HIVDETAIL	0 = HIV-; 1 = previously diagnosed HIV+; 2 = newly diagnosed HIV+	
4	HIVpos	HIV status	HIV STATUS AT T1	COMPOSITE VARIABLE OF HTRT1R, HTRT2R, NICAFFS, AND NRLAFFS. PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS FOR DETAILED DEFINITION.	Numeric	Y10N	0 = No; 1 = Yes	
5	REQerdt	RQ/EQ date	A1 ENROLLMENT DATE	DATE RECORDED ON REQ FORM	Numeric	DDMMYY10.		

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
6	age_circum	circumcision age	CIRCUMCISION AGE AT T1	AGE_CIRCUM = AGE_FINAL + (CIRCUM_T1DT-REQERDT)/365.25 +1	Numeric	BEST12.	age in years.	
7	age_final	age	AGE AT T1	A COMPOSITE VARIABLE OF REQYEARS, REQDOB_NUM, AND AGE RECORDED ON HC/AHA. PLEASE REFER TO /CODE/CREATE_ADATA/D_RREQ.SAS AND /CODE/OPEN/T_DISP_FIG.SAS FOR DETAILS.	Numeric	BEST12.	age in years.	
8	agegrp_final18	age group	AGE GROUP AT T1 (18-49)	AGE GROUP AT T1 RANGING FROM 18-49	Numeric	AGE18GRP.	1='18-19'; 2='20-24'; 3='25-29'; 4='30-34'; 5='35-39'; 6='40-44'; 7='45-49'	
9	circum	circumcision status	CIRCUMCISION STATUS AT T1	A COMPOSITE VARIABLE OF RQCIRCUM, EQCIRCUM, RQWCYyyy, AND EQWCYyyy. PLEASE REFER TO /CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	CIRCUM	1 = Circumcised; 2 = Uncircumcised; 3 = DK/REF; 4 = Missing item; 5 = Missing plate; 6 = Q	
10	circum_T1dt	circumcision date	CIRCUMCISION DATE AT T1	COALESCEC(RQWCYyyy, EQWCYyyy)	Numeric	DDMMYY10.	date format.	
11	circumdur	circumcision duration	CIRCUMCISION DURATION AT T1	(REQERDT-CIRCUM_T1DT)/365.25	Numeric	BEST12.	duration in years.	
12	condom	condom use	CONDOM USE AT T1	COALESCE(RQHOCU, EQSEXTC). PLEASE REFER TO /CODE/CREATE_ADATA/D_PRECOHORT.SAS FOR DETAILS.	Numeric	RQHOCU	1='sometimes'; 2='always'; 3='never'; 4='DK'; 5='REF'	
13	edu	education	EDUCATION AT T1	DERIVED BASED ON REQHLA.	Numeric	EDU	1 = Primary; 2 = Secondary; 3 = Tertiary; 4 = Did not attend/DK/REF; 5 = Missing item	
14	eligible_a1	Eligible in A1	ELIGIBILITY OF A1 PARTICIPATION	A COMPOSITE VARIABLE BASED ON HIVPOS AND ENRPEA1. PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS.	Numeric	Y10N	0 = No; 1 = Yes	
15	eligible_enrolled_a1	Eligible and enrolled in A1	ELIGIBILITY AND ENROLLMENT STATUS OF A1 PARTICIPATION	PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS.	Numeric	Y10N	0 = No; 1 = Yes	

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
16	emply	employment	EMPLOYMENT AT T1	COALESCE(RQEMPL, EQEMPL)	Numeric	RQEMPL	1='regularly employed full time'; 2='employed part-time'; 3='employed seasonally'; 4='self-employed'; 5='unemployed/looking for work'; 6='unemployed/not looking for work'; 7='retired or disabled'; 8='other'; 9='REF'	
17	enrolled_a1	Enrolled in A1	ENROLLMENT STATUS OF A1 PARTICIPATION	A COMPOSITE VARIABLE OF ENRPEA1 AND ENRENA1. PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS.	Numeric	Y10N	0 = No; 1 = Yes	
18	gender_final	gender	GENDER AT T1	DERIVED BASED ON REQSEX AND SEX ON HC/AHA FORMS.	Numeric	SEX	1 = Male; 2 = Female	
19	has_rreq	has RQ/EQ	WHETHER THE PARTICIPANT HAD A RQ/EQ FORM	HAS_RREQ=1 IF THE PARTICIPANT HAD A REQ, RQ OR EQ FORM.	Numeric	Y10N	0 = No; 1 = Yes	
20	hhid	Household ID	HOUSEHOLD ID	HOUSEHOLD ID DERIVED FROM PTID DIGITS #5-6	Numeric	BEST12.	1-27	
21	lasthiv	result of last HIV test	RESULT OF LAST HIV TEST AT T1	PLEASE REFER TO /CODE/OPEN/T_SELFHIV.SAS FOR DETAILS.	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
22	marital	marital status	MARITAL STATUS AT T1	PLEASE REFER TO CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	MARITAL	1 = Not married, ever had sex; 2 = Not married, never had sex; 3 = Not married, unknown sex history; 4 = Married, living with partner; 5 = Married, partner stays elsewhere ² ; 6 = Married, unknown living situation; 7 = REF; 8 = Missing marriage item; 9 = Missing marriage plate	

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
23	marry	marriage	MARRIAGE STATUS AT T1	DERIVED BASED ON RQMARRIE AND EQMARRIE. PLEASE REFER TO CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	RQMARRIE	1 = Currently married; 2 = Married and living with a man/woman; 3 = Not married or not living together with a man/woman as married; 4 = REF; 5 = Missing item; 6 = Missing plate	
24	mcontact	pre-cohort enrollment status	PRE-COHORT ENROLLMENT STATUS	PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS FOR DETAILS.	Numeric	MCONTACT	0 = member is illegible; 1 = member is eligible but refused; 2 = member is eligible and participated; 3 = member is eligible but unable to contact; 4 = member is eligible but unknown enrollment status	
25	never_sex	never had sex	WHETHER HAD SEX PRIOR TO T1	DERIVED BASED ON RQHNREF AND EQSEXFTN. PLEASE REFER TO CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	RQHNREF	1 = have never had sex; 2 = REF; 3 = Missing item; 4 = Missing plate	
26	no2p	Has P2	WHETHER HAD A 2ND PARTNER	COALESCE(RQNO2P, EQNO2P)	Numeric	Y10N	0 = No; 1 = Yes	
27	no3p	Has P3	WHETHER HAD A 3RD PARTNER	COALESCE(RQNO3P, EQNO3P)	Numeric	Y10N	0 = No; 1 = Yes	
28	p1art	P2 ART status	1ST PARTNER ART STATUS AT T1	COALESCE(RQP1TART, EQP1TART)	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
29	p1gender	P1gender	1ST PARTNER GENDER AT T1	COALESCE(RQP1PS, EQP1PS)	Numeric	SEX	1 = Male; 2 = Female	
30	p1hiv	P1 HIV status	1ST PARTNER HIV STATUS AT T1	COALESCE(RQP1THIV, EQP1TPHI)	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
31	p2art	P2 ART status	2ND PARTNER ART STATUS AT T1	COALESCE(RQP2TART, EQP2TART)	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
32	p2gender	P2gender	2ND PARTNER GENDER AT T1	COALESCE(RQP2PS, EQP2PS)	Numeric	SEX	1 = Male; 2 = Female	
33	p2hiv	P2 HIV status	2ND PARTNER HIV STATUS AT T1	COALESCE(RQP2THIV, EQP2TPHI)	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
34	p3art	P3 ART status	3RD PARTNER ART STATUS AT T1	COALESCE(RQP3TART, EQP3TART)	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
35	p3gender	P3gender	3RD PARTNER GENDER AT T1	COALESCE(RQP3PS, EQP3PS)	Numeric	SEX	1 = Male; 2 = Female	
36	p3hiv	P3 HIV status	3RD PARTNER HIV STATUS	COALESCE(RQP3THIV, EQP3TPHI)	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
37	partners	number of partners surveyed	NUMBER OF PARTNERS SURVEYED AT T1	DERIVED BASED ON P#GENDER	Numeric	SEXPRTF	1 = No partners; 2 = 1 partner; 3 =>1 partner	
38	preg	pregnancy status	FEMALE PARTICIPANT PREGNANCY STATUS AT T1	DERIVED BASED ON RQPREG AND EQPREG. PLEASE REFER TO /CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	PREG	1 = Yes; 2 = No; 3 = DK/REF; 4 = Missing item; 5 = Missing plate; 6 = Q	
39	probselect	probability of selection	PROBABILITY OF SELECTION OF EAS	PLEASE REFER TO /CODE/CREATE_ADATA/D_WEIGHT.SAS FOR DETAILS.	Numeric		continuous.	
40	prtnloc	partner location	PARTNER LOCATION AT T1	COALESCE(RQPLWY, EQPSWY). PLEASE REFER TO /CODE/CREATE_ADATA/D_PRECOHORT.SAS FOR MORE DETAILS.	Numeric	PRTNLOC	-1 = Missing/REF; 0 = Not married or living with a partner; 1 = Living with me; 2 = Staying elsewhere	
41	uid	uid	PATIENT ID NUMBER	A-BBB-CC-D-EE; A=REGION; BBB=AREA; CC=HHID; D=CHECK; EE=MEMBER #	Numeric	SHIMPTID	LOW - HIGH = 9-999-99-9-99	
42	rakingwgt	RakedWgt	RAKED WEIGHT ACCORDING TO POPULATION	PLEASE REFER TO /CODE/CREATE_ADATA/D_RAKING.SAS	Numeric		continuous	
43	refused	Refused to A1	REFUSAL STATUS TO A1 AT T1	PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS FOR DETAILS.	Numeric	Y10N	0 = No; 1 = Yes	
44	region	Region Number	REGION NUMBER	FIRST DIGIT OF PTID	Character		1-4	
45	resid	residence	RESIDENCE	BASED ON EA NUMBER, PLEASE REFER TO /CODE/CREATE_ADATA/D_HHCOUNT.SAS FOR DETAILS.	Numeric	RESID	1 = Urban; 2 = Rural	

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
46	selfHIV	self HIV status	SELF-REPORTED HIV STATUS AT T1	PLEASE REFER TO /CODE/OPEN/T_SELFHIV.SAS FOR DETAILS.	Numeric	SELFHIV	0 = Never had HIV test/DK/REF/Missing; 1 = Ever had HIV test - Postive; 2 = Ever had HIV test - Negative; 3 = Ever had HIV test - Indeterminate; 4 = Ever had HIV test - DK; 5 = Ever had HIV test - REF; 6 = Ever had HIV test - Did not receive results/DK/REF/Missing; 7 = Missing Plate	
47	serial	Serial	SERIAL NUMBER	SERIAL NUMBER CORRESPONDING TO EA NUMBER	Numeric	BEST9.	1-575	
48	sex_age	sex debut age	SEX DEBUT AGE AT T1	DERIVED BASED ON RQFTAGE AND EQSEXFT. PLEASE REFER /CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	SEXAGE	-2 = Missing item; -1 = Missing plate; age in years	
49	sexprtn	sex partners in the past 6 months (cat)	CATEGORICAL VARIABLE OF SEXPRTNUM AT T1		Numeric	SEXPRT	1 = 0; 2 = 1; 3 = 2 or more; 4 = REF; 5 = Missing item; 6 = Missing plate	
50	sexprtnum	sex partners in the past 6 months (cont)	NUMBER OF SEX PARTNERS IN THE PAST 6 MONTHS AT T1	DERIVED BASED ON RQTNP6M AND EQSEXT. PLEASE REFER TO /CODE/CREATE_ADATA/D_DEMO1.SAS FOR DETAILS.	Numeric	SEXPRTNUM	-3 = REF; -2 = Missing item; -1 = Missing plate	
51	sexyn	SexYN	WHETHER HAD SEX AT T1	INVERSE OF NEVER_SEX	Numeric	SEXYN	0 = No; 1 = Yes; 2 = REF; 3 = Missing both items; 4 = Missing plate; 5 = Q ⁴	
52	testdate	last HIV test date	LAST HIV TEST DATE AT T1	DERIVED BASED ON RQLTMY AND EQWTHIVY. PLEASE REFER TO /CODE/OPEN/T_SELFHIV.SAS	Numeric	TESTDATE	0 = No date of last test; 1 = Test prior to 2010; 2 - 3 = Test 2010/11	

Variable Property of PRECOHORT								
N	Variable	Label	Description	Definition	Type	Format	Format Definition	Note
53	tested	prior HIV test	WHETHER TESTED FOR HIV PRIOR TO T1	DERIVED BASED ON RQTEST AND EQTFAID. PLEASE REFER TO /CODE/OPEN/T_SELFHIV.SAS	Numeric	YNDKREF	1 = Yes; 2 = No; 3 = DK; 4 = REF	
54	unresolved	Unresolved status	WHETHER QC WAS RESOLVED	PLEASE REFER TO /CODE/OPEN/T_DISP_FIG.SAS FOR DETAILS.	Numeric	Y10N	0 = No; 1 = Yes	
55	wgt	UnrakedWgt	UNRAKED WEIGHT	1/PROBSELECT * WT_RESPONSE. PLEASE REFER TO /CODE/CREATE_ADATA/D_RAKING.SAS FOR DETAILS.	Numeric		continuous	
56	wt_response	ResponseWgt	RESPONSE WEIGHT	PLEASE REFER TO /CODE/CREATE_ADATA/D_POSTWEIGHT.SAS FOR DETAILS.	Numeric		continuous	