Prevalence Dataset Documentation

Number of observations: 18172
Number of variables: 23
Purpose: Analysis dataset for prevalence manuscript

Variable	Description	Derivation[1]	SAS Format Definition
uid	De-identified participant ID	De-identified from original ptid	
uhhid	De-identified Household ID	De-identified from original hhid	
SHIMS_EA	Swaziland EA number	Swaziland EA : used for sampling	
Region	Swaziland region	1 of 4 regions	1= “Hhohho”; 2= “Manzini”; 3= “Shiselweni”; 4= “Lubombo”;
urban	residence: rural/urban	Household residence classification	1= “Urban” 2= “Rural”
REQ_erdt	Participant baseline survey enrollment date (T1)	Date recorded on RQ/EQ form: REQerdt	
gender	Participant gender	REQsex	1= “Male” 2= “Female”
age	Participant age at T1	Derived from REQmy when provided, REQyears if REQmy not available.	
age_grp	Participant age group at T1	Categories created based on age variable	1=‘18-19;2=‘20-24’;3=‘25-29’;4=‘30-34’; 5=‘35-39’;6=‘40-44’;7= ‘45-49’;
education	Participant highest level of education at T1	REQhlsa; responses “DK” and “REF” categorized as “Missing”	1=’Primary’; 2=’Secondary’; 3=’Tertiary’; 4=’Did not attend’;

[1] Variable labels referenced in this column are taken from the annotated CRF for the routine questionnaire. Note that the corresponding variables from the extended questionnaire might have slightly different labels recorded on the EQ annotated CRF.

			5='Missing';
marital	Combined variable indicating participant marital status, partner location, and previous sexual activity	If RQmarrie = "No" & sex_age > 0 → "Not married, ever had sex" If RQmarrie = "No" & sex_age = "Have never had sex" → "Not married, never had sex" If RQmarrie = "No" & sex_age = "Missing" → "Not married, unknown sex history" If RQmarrie = "Yes, married" or "Yes, living with…" & RQplwy = "Living with me" → "Married, living with partners" If RQmarrie = "Yes" & RQplwy = "Staying elsewhere" → "Married, partners living elsewhere" If RQmarrie = "Yes" & RQplwy = "REF"/"Missing" → "Married, unknown living situation" If RQmarrie = "no"/"REF" → "Missing"	1= "Not married, ever had sex"; 2= "Not married, never had sex"; 3= "Not married, unknown sex history"; 4= "Married, living with partners"; 5= "Married, partners living elsewhere"; 6= "Married, unknown living situation"; 7-9= "Missing";
sex_age	Participant reported age of first sexual encounter	if RQhnref= "have never had sex" or RQftage=0 then sex_age= "Have never had sex"; else sex_age = RQftage;	-1= "Never had sex" -2= "Missing"
sexprt_num	Participant reported number of sexual partners in the 6 months prior to T1	If sex_age = "have never had sex", then sexprt_num = 0; else sexprt_num = RQtnp6m	-1, -2,-3 = "Missing";
sexprt_cat	Categorical variable for sexprt_num	Categoried created based on sexprt_num: 0, 1, 2+	1= "0"; 2= "1" 3= "2 or more"; 4-6= "Missing"
preg	Female participant pregnancy status at T1	RQpreg	1= "Yes"; 2= "No"; 3-5= "Missing";
circum	Male participant circumcision status at T1	RQcircum	1= "Circumcised"; 2= "Uncircumcised"; 3-5= "Missing";
hiv_testhistory	Participant reported HIV test history with results if reported	Derived based on RQtest, RQresult, RQrltest: If rqtest = "No" → "Never had test before" If rqtest = "Yes" & RQresult = "No"/"DK"/"REF"/"Missing" → "Ever had test – did not receive result" If rqtest = "Yes" & RQresult = "Yes" → hiv_testhistory = "Ever	0= "Never had test before" 1= "Ever had test-Positive" 2='Ever had test -

		had HIV test, [RQrltest value]"	Negative" 3= "Ever had test- result Indeterminate" 4= "Ever had test- result DK" 6= "Ever had test – did not receive result" 7 = "Missing"
hivstatus_t1	Participant HIV test result from SHIMS-administered testing	Derived based on HTR form and lab datasets (NIC, NRL): Status determined by rapid test results (HTRt1r, HTRt2r), with correction/confirmation by NIC/NRL results when available.	1= "Neg" 2= "Pos"
aware_hiv_art	Derived variable indicating HIV+ participants' awareness of HIV status and current ART use if aware	Derived based on hiv_testhistory, RQctart, only defined for hivstatus_t1 = "Pos": If hiv_testhistory = "Yes, positive" & RQctart = "Yes" → aware_hiv_art = "Aware, ART use" If hiv_testhistory = "Yes, positive" & RQctart = "No" → aware_hiv_art = "Aware, no ART use" Other values of hiv_testhistory → aware_hiv_art = "Unaware"	1= "Unaware"; 2= "Aware, no ART use"; 3= "Aware, ART use"; 4= "Aware, unknown ART use"
wt_poststratification	Poststratification survey weights	Poststratification weights calculated using 2007 DHS population census data for region, gender, resid, and age categories.	
wt_design	Probability of selection survey weights	Design weights calculated probability of selection of EAs and households within EAs for each of eight strata defined by region and resid (4 regions, urban/rural).	
wt_population	Final survey weights scaled to Swaziland total population size (N = 386179)	Combined design and poststratification weights, scaled to the size of Swaziland's total population.	
wt_cohort	Final survey weights, with raking applied, that are scaled to the size of the T1 cross-sectional survey participants (18172). Used in prevalence analysis.	Combined design and poststratification weights, scaled to the size of the T1 cross-sectional survey participants. Used in the prevalence analysis/manuscript (Bicego 2013).	