

South Africa - AHRI.Surveillance Episodes HIV Dataset 2023

Sweetness H Dube

Report generated on: September 6, 2023

Visit our data catalog at: <https://data.ahri.org/index.php>

Identification

SURVEY ID NUMBER
AHRI.SEH2023.v1

TITLE
AHRI.Surveillance Episodes HIV Dataset 2023

COUNTRY

Name	Country code
South Africa	ZA

ABSTRACT

Africa Health Research Institute (AHRI) is a multidisciplinary, independent research institute. AHRI's goal is to become a source of fundamental discoveries into the susceptibility, transmission and cure of HIV, TB and related diseases. We also seek ways to improve diagnosis, prevention and treatment. To achieve this, we bring together leading researchers from different fields, use cutting edge science to improve people's health, and help to train the next generation of African scientists.

AHRI has conducted individual and household demographic surveillance since 2000. Demographic surveillance data are processed for longitudinal analyses by organizing the demographic surveillance data into individual and household residency episodes at a geographical physical location, and individual membership episodes within a household. Start events of residency and memberships include enumeration, birth, in-migration and relocating into a household from within the study population. Exit events of residency and memberships include death (by cause), out-migration and relocating to another location in the study population. AHRI annually produces a suite of Individual Demographic surveillance datasets which contains detailed records of all individuals under surveillance since January 2000. The datasets detail individual residency and membership patterns, household mortality (including cause of death), fertility, migration, HIV status, health care utilization, and household socio-economic profile.

Each record in the dataset represents a period of observation for an individual during which all the recorded characteristics of the individual stay constant. For example, on the birthday of the individual a new episode will start, because the age of the individual has changed. An out-migration will result in a new episode, because the location or residential status has changed. Any change in one of the status values, such as education, marital status, social-economic status and HIV will likewise result in a new episode on the date of the change.

There are four versions of the Demographic surveillance datasets:

1. SurveillanceEpisodesBasic
2. SurveillanceEpisodesBasicYrAge
3. SurveillanceEpisodesHIV
4. SurveillanceEpisodesYrAgeDelivery

The definitions of the SurveillanceEpisodes vary slightly between the four versions. All SurveillanceEpisodesBasic are specific to a) individual b) residency type except that a new episode starts if the age changes (SurveillanceEpisodes BasicYrAge), HIV status changes (SurveillanceEpisodesBasicHIV). The SurveillanceEpisodesBasic datasets contain specific variables in addition to the minimum variables that each dataset has. For instance, SurveillanceEpisodesBasicHIV contains all HIV related variables in addition to the basic minimum variables that all surveillance episodes contain. Causes and likelihood of deaths from InterVA 5 are also included. Table 1 below describes the seven Demographic surveillance datasets

KIND OF DATA

Event History Data

UNIT OF ANALYSIS

Episodes of exposure

Version

VERSION DESCRIPTION

v1.0.0

Scope

TOPICS

Topic	Vocabulary	URI
Surveillance Episodes Basic HIV	Africa Health Research Institute	www.ahri.org

KEYWORDS

Keyword	Vocabulary	URI
Fertility, Mortality, Migration, Social-economic status, HIV, ART	Africa Health Research Institute	www.ahri.org

Coverage

GEOGRAPHIC COVERAGE

AHRI is situated in the south-east portion of the uMkhanyakude district of KwaZulu-Natal province near the town of Mtubatuba. It is bounded on the west by the Umfolozi-Hluhluwe nature reserve, on the south by the Umfolozi river, on the east by the N2 highway (except from portions where the KwaMsane township strangles the highway) and in the north by the provincial road R22. The surveillance area is approximately 845km² in size.

UNIVERSE

Households resident in dwellings within the study area are eligible for inclusion in the household component of AHRI Demographic Surveillance. All individuals identified by the household proxy informant as members of the households are enumerated. A resident household member is an individual that intends to sleep majority of time at the dwelling occupied by the household over a four-month period. Households include resident and non-resident members. An individual is a non-resident member if they have close ties to the household, but do not physically reside with the household at the dwelling most of the time. Non-resident members can also be called temporary migrants and they are enumerated within the household list. Because household membership is not tied to physical residency, an individual may be a member of more than one household.

Producers and sponsors

PRIMARY INVESTIGATORS

Name	Affiliation
Kobus Herbst	AHRI, SAPRIN
Guy Harling	AHRI
Mark Siedner	AHRI
Willem Hanekom	AHRI

PRODUCERS

Name
Africa Health Research Institute

FUNDING AGENCY/SPONSOR

Name	Abbreviation	Role
Wellcome Trust	WT	Core funding
SAPRIN	SAPRIN	

OTHER IDENTIFICATIONS/ACKNOWLEDGMENTS

Name	Affiliation	Role
AHRI Research Data Management Team	AHRI	Research Data Management
AHRI Data Collection Team	AHRI	Data Collection
AHRI Community Engagement Team	AHRI	Community Engagement

Sampling

SAMPLING PROCEDURE

This dataset is not based on a sample but contains information from the complete demographic surveillance areas.

Data Collection

DATES OF DATA COLLECTION

Start	End
2000-01-01	2023-07-31

Data Processing

DATA EDITING

Data collected in this study was centrally stored on AHRI's secure database servers. Data was collected through interviews at the call centre on central server using Survey Solutions, and in the field on paper and encrypted tablet computers using REDCap, and then securely uploaded to a central server. Database rules and constraints were enforced on the master database to ensure that only valid data was stored. Personally identifiable data was restricted to separate database tables accessible only to data management staff .

Access policy

ACCESS CONDITIONS

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.

CITATION REQUIREMENTS

Herbst, K., Harling, G., Siedner, M., & Hanekom, W. (2023). AHRI.Surveillance Episodes HIV Dataset 2023 (Version 1) [Data set]. Africa Health Research Institute. DOI:<https://doi.org/10.23664/AHRI.SEH2023>

Metadata production

DDI DOCUMENT ID

DDI.AHRI.SEH2023.v1

PRODUCERS

Name	Abbreviation
------	--------------

Africa Health Research Institute	AHRI
----------------------------------	------

Data Dictionary

Data file	Cases	Variables
SurveillanceEpisodesHIV	786200	36

Data file: SurveillanceEpisodesHIV

Cases: 786200

Variables: 36

Variables

ID	Name	Label	Question
V282	IntId	Unique PIP internal individual identifier	
V283	Resident	Whether individual is resident for duration of episode	
V284	LocationId	Where individual was resident, household residence if non-resident	
V285	HouseholdId	Unique household identifier of the household the individual is a member of	
V286	Sex	Sex of individual	
V287	DoB	Date of birth	
V288	DoD	Date of death	
V289	MotherId	Mother's IndividualId	
V290	FatherId	Father's IndividualId	
V291	StartDate	Start date of episode (inclusive)	
V292	EndDate	End date of episode (inclusive)	
V293	Days	Duration in days of episode	
V294	Born	Episode starts with the birth of the individual	
V295	Enumeration	Episode starts with an enumeration	
V296	InMigration	Episode starts with an in-migration	
V297	LocationEntry	Episode starts with an internal migration	
V298	ExtResStart	Flag to indicate start of external residence	
V299	Participation	Resume participation after refusal	
V300	Died	Episode ends with the death of the individual	
V301	OutMigration	Episode ends with the out-migration	
V302	LocationExit	Episode ends with an internal migration	
V303	ExtResEnd	Flag to indicate end of external residence	
V304	Refusal	Individual refused follow-up	
V305	LostToFollowUp	Individual was lost to follow-up at the end of the episode	
V306	Current	Individual still under surveillance	
V307	MembershipStart	Flag to indicate start of household membership	
V308	MembershipEnd	Flag to indicate end of household membership	
V309	Memberships	Number of concurrent household memberships	
V310	Gap	Gap in individual exposure record	
V311	HIVStatus	HIV Status for the duration of the episode	
V312	ResultDate	Date of HIV Result	
V313	HIVResult	Result of DBS HIV test	
V314	StatusStart	Start of HIV status episodes	
V315	StatusEnd	End of HIV status episodes	
V316	Episode	This episode number (first=1, last=Episodes)	
V317	Episodes	Total number of episodes for individual	

Total: 36

IINTID: Unique PIP internal individual identifier**Data file: SurveillanceEpisodesHIV****Overview**

Valid: 786200 Invalid: 0 Minimum: 11 Maximum: 271715 Mean: 104326.179 Standard deviation: 70975.068

Type: Continuous Decimal: 0 Width: 12 Range: 11 - 271715 Format: Numeric

RESIDENT: Whether individual is resident for duration of episode**Data file: SurveillanceEpisodesHIV****Overview**Valid: 786200 Invalid: 0
Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric**Questions and instructions**

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	214983	27.3%
1	Yes	571217	72.7%

LOCATIONID: Where individual was resident, household residence if non-resident**Data file: SurveillanceEpisodesHIV****Overview**Valid: 786200 Invalid: 0 Minimum: 0 Maximum: 354564 Mean: 13682.215 Standard deviation: 29372.704
Type: Continuous Decimal: 0 Width: 12 Range: 0 - 354564 Format: Numeric**HOUSEHOLDID: Unique household identifier of the household the individual is a member of****Data file: SurveillanceEpisodesHIV****Overview**Valid: 786200 Invalid: 0 Minimum: 11 Maximum: 46662 Mean: 14739.75 Standard deviation: 12696.193
Type: Continuous Decimal: 0 Width: 12 Range: 11 - 46662 Format: Numeric**SEX: Sex of individual****Data file: SurveillanceEpisodesHIV**

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 12 Range: 1 - 3 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
1	Male	336384	42.8%
2	Female	449814	57.2%
3	Unknown	2	0%

DOB: Date of birth

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Minimum: 1890-01-01 Maximum: 2023-07-19
 Type: Discrete Width: 11 Range: - Format: character

DOD: Date of death

Data file: SurveillanceEpisodesHIV

Overview

Valid: 71669 Minimum: 2000-01-02 Maximum: 2023-07-27
 Type: Discrete Width: 11 Range: - Format: character

MOTHERID: Mother's IndividualId

Data file: SurveillanceEpisodesHIV

Overview

Valid: 506299 Invalid: 279901 Minimum: 12 Maximum: 271612 Mean: 79008.562 Standard deviation: 62331.455
 Type: Continuous Decimal: 0 Width: 12 Range: 12 - 271612 Format: Numeric

FATHERID: Father's IndividualId

Data file: SurveillanceEpisodesHIV

Overview

Valid: 270274 Invalid: 515926 Minimum: 11 Maximum: 271324 Mean: 68677.569 Standard deviation: 56985.212
 Type: Continuous Decimal: 0 Width: 12 Range: 11 - 271324 Format: Numeric

STARTDATE: Start date of episode (inclusive)**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Minimum: 2000-01-01 Maximum: 2023-07-30
 Type: Discrete Width: 11 Range: - Format: character

ENDDATE: End date of episode (inclusive)**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Minimum: 2000-01-01 Maximum: 2023-07-31
 Type: Discrete Width: 11 Range: - Format: character

DAYS: Duration in days of episode**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Invalid: 0 Minimum: 1 Maximum: 8613 Mean: 1186.06 Standard deviation: 1316.787
 Type: Continuous Decimal: 0 Width: 12 Range: 1 - 8613 Format: Numeric

BORN: Episode starts with the birth of the individual**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	740073	94.1%
1	Yes	46127	5.9%

ENUMERATION: Episode starts with an enumeration**Data file:** SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	647914	82.4%
1	Yes	138286	17.6%

INMIGRATION: Episode starts with an in-migration

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	638543	81.2%
1	Yes	147657	18.8%

LOCATIONENTRY: Episode starts with an internal migration

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	728571	92.7%
1	Yes	57629	7.3%

EXTRESSTART: Flag to indicate start of external residence**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	623353	79.3%
1	Yes	162847	20.7%

PARTICIPATION: Resume participation after refusal**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	786196	100%
1	Yes	4	0%

DIED: Episode ends with the death of the individual**Data file:** SurveillanceEpisodesHIV**Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	760339	96.7%
1	Yes	25861	3.3%

OUTMIGRATION: Episode ends with the out-migration

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	615245	78.3%
1	Yes	170955	21.7%

LOCATIONEXIT: Episode ends with an internal migration

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	728559	92.7%
1	Yes	57641	7.3%

EXTRESEND: Flag to indicate end of external residence

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0

Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	657677	83.7%
1	Yes	128523	16.3%

REFUSAL: Individual refused follow-up

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	782803	99.6%
1	Yes	3397	0.4%

LOSTTOFOLLOWUP: Individual was lost to follow-up at the end of the episode

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	769708	97.9%
1	Yes	16492	2.1%

CURRENT: Individual still under surveillance**Data file: SurveillanceEpisodesHIV****Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	632868	80.5%
1	Yes	153332	19.5%

MEMBERSHIPSTART: Flag to indicate start of household membership**Data file: SurveillanceEpisodesHIV****Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	452000	57.5%
1	Yes	334200	42.5%

MEMBERSHIPEND: Flag to indicate end of household membership**Data file: SurveillanceEpisodesHIV****Overview**

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%

0	No	604720	76.9%
1	Yes	181480	23.1%

MEMBERSHIPS: Number of concurrent household memberships

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
Type: Discrete Decimal: 0 Width: 12 Range: 1 - 10 Format: Numeric

GAP: Gap in individual exposure record

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	771559	98.1%
1	Yes	14641	1.9%

HIVSTATUS: HIV Status for the duration of the episode

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
Type: Discrete Decimal: 0 Width: 13 Range: 0 - 9 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
0	Unknown	345392	43.9%
1	Pre-negative	106180	13.5%
2	Negative	127835	16.3%
3	Post-negative	89778	11.4%

4	Pre-positive	36168	4.6%
5	Positive	79995	10.2%
9	Error	852	0.1%

RESULTDATE: Date of HIV Result

Data file: SurveillanceEpisodesHIV

Overview

Valid: 440808 Minimum: 2001-02-21 Maximum: 2023-05-05
 Type: Discrete Width: 11 Range: - Format: character

HIVRESULT: Result of DBS HIV test

Data file: SurveillanceEpisodesHIV

Overview

Valid: 440808 Invalid: 0
 Type: Discrete Width: 1 Range: - Format: character

STATUSSTART: Start of HIV status episodes

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	424615	54%
1	Yes	361585	46%

STATUSEND: End of HIV status episodes

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0
 Type: Discrete Decimal: 0 Width: 15 Range: -1 - 1 Format: Numeric

Questions and instructions

CATEGORIES

Value	Category	Cases	
-1	Unknown/Missing	0	0%
0	No	678774	86.3%
1	Yes	107426	13.7%

EPISODE: This episode number (first=1, last=Episodes)

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0 Minimum: 1 Maximum: 26 Mean: 3.39 Standard deviation: 2.888
 Type: Continuous Decimal: 0 Width: 12 Range: 1 - 26 Format: Numeric

EPISODES: Total number of episodes for individual

Data file: SurveillanceEpisodesHIV

Overview

Valid: 786200 Invalid: 0 Minimum: 1 Maximum: 26 Mean: 5.781 Standard deviation: 4.114
 Type: Continuous Decimal: 0 Width: 12 Range: 1 - 26 Format: Numeric

Download related resources

Technical documents

DDI:SEH2023

Title	DDI:SEH2023
Author(s)	Sweetness H Dube
Date	15/08/2023
Country	South Africa
Language	English
Publisher(s)	Sweetness H Dube
Filename	AHRI.SEH2023.v1.pdf
