

**South Africa**

**Dr Kobus Herbst, AHRI, SAPRIN**

**Dr Guy Harling, AHRI**

**Dr Mark Siedner, AHRI**

**Prof Willem Hanekom, AHRI**

**AHRI.Surveillance Episodes Year Age Coresidency**

**Study Documentation**

February 19, 2026

# Metadata Production

<b>Metadata Producer(s)</b>	Africa Health Research Institute (AHRI)
<b>Identification</b>	DDI.AHRI.SurveillanceEpisodesYrAge.Coresidency

# Table of Contents

<a href="#">Overview</a> .....	4
<a href="#">Scope &amp; Coverage</a> .....	5
<a href="#">Producers &amp; Sponsors</a> .....	5
<a href="#">Sampling</a> .....	5
<a href="#">Data Collection</a> .....	6
<a href="#">Accessibility</a> .....	6
<a href="#">Files Description</a> .....	7
<a href="#">SurveillanceEpisodesYrAge_Coresidency</a> .....	7
<a href="#">Variables List</a> .....	8
<a href="#">SurveillanceEpisodesYrAge_Coresidency</a> .....	8
<a href="#">Variables Description</a> .....	10
<a href="#">SurveillanceEpisodesYrAge_Coresidency</a> .....	11

## AHRI.Surveillance Episodes Year Age Coresidency

### Overview

<b>Identification</b>	AHRI.SurveillanceEpisodesYrAge.Coresidency
<b>Version</b>	v2.0.0

#### 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.

The following are the versions of the Demographic surveillance datasets:

- CoresidencyEpisodes
- EducationCleaned
- EducationStatuses
- HIVStatusEpisodes
- HouseholdAssets
- HouseholdMap
- HouseholdObservations
- HouseholdSocioEconomicStatus
- HouseholdStatusEpisodes
- IndividualHSEEpisodes
- IndividualMap
- LabourEpisodes
- LabourStatuses
- LocationMap
- MaritalStatuses
- PartnerEpisodes
- SurveillanceEpisodesBasic
- SurveillanceEpisodesYTAge
- SurveillanceEpisodes YrAge\_Coresidency
- SurveillanceEpisodes YrAge\_HIV
- SurveillanceEpisodes VrAge\_HSE
- SurveillanceEpisodes YrAge\_Labour
- SurveillanceEpisodes YrAge\_Partners
- SurveillanceEpisodes YrAgeDel

SurveillanceEpisodes VrAgeDel_HSE SurveillanceEpisodesYrAgeDeL_HSE_Labour SurveillanceEpisodesYrAgeDeL_HSE_Labour_Partner_Parents SurveillanceEpisodes VrAgeDel_HSE_Labour_Partner_Parents_HIV SurveillanceEpisodesYrAgeDeL_HSE_Labour_Partners	
<b>Kind of Data</b>	Event History Data
<b>Unit of Analysis</b>	Episodes of exposure

<b>Scope &amp; Coverage</b>	
<b>Keywords</b>	Fertility, Mortality, Migration, Social-economic status, HIV, ART
<b>Topics</b>	Surveillance Episodes Year Age_Coresidency
<b>Time Period(s)</b>	2000-2024
<b>Countries</b>	South Africa
<b>Geographic Coverage</b> 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 form portions where the KwaMsane township strangles the highway) and in the north by the provincial road R22. The surveillance area is approximately 850km2 in size.	
<b>Universe</b> 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.	

<b>Producers &amp; Sponsors</b>	
<b>Primary Investigator(s)</b>	Dr Kobus Herbst, AHRI, SAPRIN Dr Guy Harling, AHRI Dr Mark Siedner, AHRI Prof Willem Hanekom, AHRI
<b>Other Producer(s)</b>	Africa Health Research Institute (AHRI)
<b>Funding Agency/ies</b>	Wellcome Trust (WT) , Core funding SAPRIN (SAPRIN)
<b>Other Acknowledgment(s)</b>	AHRI Research Data Management Team , Research Data Management , AHRI AHRI Data Collection Team , Data Collection , AHRI AHRI Community Engagement Team , Community Engagement , AHRI

<b>Sampling</b>	
<b>Sampling Procedure</b> This dataset is not based on a sample but contains information from the complete demographic surveillance areas.	

## Data Collection

<b>Data Collection Dates</b>	start 2000-01-01 end 2024-12-31
------------------------------	------------------------------------

## Accessibility

### 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.

<https://github.com/SAPRIN-MRC/SaprinEpisodes.jl.git>

### Citation Requirements

Herbst, K., Harling, G., Siedner, M., & Hanekom, W. (2026). AHRI.Surveillance Episodes Year Age Coresidency [Data set]. Africa Health Research Institute.

DOI:<https://doi.org/10.23664/AHRI.SURVEILLANCEEPISODESYRAGE.CORESIDENTY>

# Files Description

Dataset contains 1 file(s)

<b>SurveillanceEpisodesYrAge_Coresidency</b>	
<b># Cases</b>	6234957
<b># Variable(s)</b>	45

# Variables List

Dataset contains 45 variable(s)

File SurveillanceEpisodesYrAge_Coresidency							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	<a href="#">Individu..</a>	Unique individual identifier	continuous	numeric-9.0	6234957	0	-
2	<a href="#">Sex</a>	Sex of individual	discrete	numeric-9.0	6234957	0	-
3	<a href="#">DoB</a>	Date of birth	discrete	character-11	6234940	-	-
4	<a href="#">DoD</a>	Date of death	discrete	character-11	511733	-	-
5	<a href="#">MotherId</a>	Mother's IndividualId	continuous	numeric-9.0	4383709	1851248	-
6	<a href="#">FatherId</a>	Father's IndividualId	continuous	numeric-9.0	2299078	3935879	-
7	<a href="#">LocationId</a>	Where individual was resident, household residence if non-resident	continuous	numeric-9.0	6234957	0	-
8	<a href="#">Househol..</a>	Unique household identifier of the household the individual is a member of	continuous	numeric-9.0	6234957	0	-
9	<a href="#">HHRelati..</a>	Relationship to head of household	discrete	numeric-22.0	6234957	0	-
10	<a href="#">Resident</a>	Whether individual is resident for duration of episode	discrete	numeric-15.0	6234957	0	-
11	<a href="#">Episode</a>	This episode number (first=1, last=Episodes)	continuous	numeric-9.0	6234957	0	-
12	<a href="#">Episodes</a>	Total number of episodes for individual	continuous	numeric-9.0	6234957	0	-
13	<a href="#">StartDate</a>	Start date of episode (inclusive)	discrete	character-11	6234957	-	-
14	<a href="#">EndDate</a>	End date of episode (inclusive)	discrete	character-11	6234957	-	-
15	<a href="#">Days</a>	Duration in days of episode	continuous	numeric-9.0	6234957	0	-
16	<a href="#">Calendar..</a>	Calendar year during episode	continuous	numeric-9.0	6234957	0	-
17	<a href="#">Age</a>	Age in completed years during episode	continuous	numeric-9.0	6234957	0	-
18	<a href="#">AgeGroup</a>	Age group during episode	discrete	numeric-9.0	6234957	0	-
19	<a href="#">Born</a>	Episode starts with the birth of the individual	discrete	numeric-15.0	6234957	0	-
20	<a href="#">Enumerat..</a>	Episode starts with an enumeration	discrete	numeric-15.0	6234957	0	-
21	<a href="#">InMigrat..</a>	Episode starts with an in-migration	discrete	numeric-15.0	6234957	0	-
22	<a href="#">Location..</a>	Episode starts with an internal migration	discrete	numeric-15.0	6234957	0	-
23	<a href="#">ExtResSt..</a>	Flag to indicate start of external residence	discrete	numeric-15.0	6234957	0	-
24	<a href="#">Particip..</a>	Resumed participation after refusal	discrete	numeric-15.0	6234957	0	-

<b>File SurveillanceEpisodesYrAge_Coresidency</b>							
#	Name	Label	Type	Format	Valid	Invalid	Question
25	<a href="#">Died</a>	Episode ends with the death of the individual	discrete	numeric-15.0	6234957	0	-
26	<a href="#">OutMigra..</a>	Episode ends with the out-migration	discrete	numeric-15.0	6234957	0	-
27	<a href="#">Location..</a>	Episode ends with an internal migration	discrete	numeric-15.0	6234957	0	-
28	<a href="#">ExtResEnd</a>	Flag to indicate end of external residence	discrete	numeric-15.0	6234957	0	-
29	<a href="#">Refusal</a>	Individual refused follow-up	discrete	numeric-15.0	6234957	0	-
30	<a href="#">LostToFo..</a>	Individual was lost to follow-up at the end of the episode	discrete	numeric-15.0	6234957	0	-
31	<a href="#">Current</a>	Individual still under surveillance	discrete	numeric-15.0	6234957	0	-
32	<a href="#">HHRelati..</a>	Relationship to head of household changed during this episode	discrete	numeric-15.0	6234957	0	-
33	<a href="#">Membersh..</a>	Flag to indicate start of household membership	discrete	numeric-15.0	6234957	0	-
34	<a href="#">Membersh..</a>	Flag to indicate end of household membership	discrete	numeric-15.0	6234957	0	-
35	<a href="#">GapStart</a>	Start of a gap in individual exposure record	discrete	numeric-15.0	6234957	0	-
36	<a href="#">GapEnd</a>	End of gap in individual exposure record	discrete	numeric-15.0	6234957	0	-
37	<a href="#">AgeChange</a>	Flag age changed	discrete	numeric-15.0	6234957	0	-
38	<a href="#">YrChange</a>	Flag calendar year changed	discrete	numeric-15.0	6234957	0	-
39	<a href="#">Membersh..</a>	Number of concurrent household memberships	continuous	numeric-9.0	6234957	0	-
40	<a href="#">Educatio..</a>	Education level achieved during this calendar year	discrete	numeric-18.0	4570696	1664261	-
41	<a href="#">MotherDead</a>	Mother is dead	discrete	numeric-9.0	6234957	0	-
42	<a href="#">MotherCo..</a>	Mother is co-resident	discrete	numeric-19.0	6234957	0	-
43	<a href="#">FatherDead</a>	Father is dead	discrete	numeric-9.0	6234957	0	-
44	<a href="#">FatherCo..</a>	Father is co-resident	discrete	numeric-19.0	6234957	0	-
45	<a href="#">Parental..</a>	Change in parental status	discrete	numeric-15.0	6234957	0	-

# Variables Description

Dataset contains 45 variable(s)

## File : SurveillanceEpisodesYrAge\_Coresidency

### # IndividualId: Unique individual identifier

**Information** [Type= continuous] [Format=numeric] [Range= 1-263153] [Missing=\*]

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-] [Mean=117323.627 /-] [StdDev=77591.694 /-]

### # Sex: Sex of individual

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0		59	0.0%
1	Male	2955054	47.4%
2	Female	3279844	52.6%
3	Unknown	0	

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

### # DoB: Date of birth

**Information** [Type= discrete] [Format=character] [Missing=\*]

**Statistics [NW/ W]** [Valid=6234940 /-]

### # DoD: Date of death

**Information** [Type= discrete] [Format=character] [Missing=\*]

**Statistics [NW/ W]** [Valid=511733 /-]

### # MotherId: Mother's IndividualId

**Information** [Type= continuous] [Format=numeric] [Range= 8-263150] [Missing=\*]

**Statistics [NW/ W]** [Valid=4383709 /-] [Invalid=1851248 /-] [Mean=154798.167 /-] [StdDev=66868.327 /-]

### # FatherId: Father's IndividualId

**Information** [Type= continuous] [Format=numeric] [Range= 49-263146] [Missing=\*]

**Statistics [NW/ W]** [Valid=2299078 /-] [Invalid=3935879 /-] [Mean=164148.964 /-] [StdDev=59796.413 /-]

### # LocationId: Where individual was resident, household residence if non-resident

**Information** [Type= continuous] [Format=numeric] [Range= 0-37292] [Missing=\*]

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-] [Mean=18484.325 /-] [StdDev=10754.421 /-]

### # HouseholdId: Unique household identifier of the household the individual is a member of

**Information** [Type= continuous] [Format=numeric] [Range= 1-41014] [Missing=\*]

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-] [Mean=20450.464 /-] [StdDev=11782.474 /-]

### # HHRelationshipTypeId: Relationship to head of household

**Information** [Type= discrete] [Format=numeric] [Range= 1-12] [Missing=\*]

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Head	784037	12.6%
2	Spouse	303467	4.9%
3	Child	2097566	33.6%
4	Son/Daughter-inlaw	26769	0.4%
5	Grandchild	1441906	23.1%
6	Parent	43228	0.7%

## File : SurveillanceEpisodesYrAge\_Coresidency

### # HHRelationshipTypeId: Relationship to head of household

Value	Label	Cases	Percentage
7	Parent-inlaw	1584	0.0%
8	Grandparent	7868	0.1%
9	Sibling	1040701	16.7%
10	Other relative	13094	0.2%
11	Domestic worker/tenant	330683	5.3%
12	Unrelated/Other	144054	2.3%

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

### # Resident: Whether individual is resident for duration of episode

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	1787224	28.7%
1	Yes	4447733	71.3%

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

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

<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 1-73] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-] [Mean=18.678 /-] [StdDev=14.347 /-]

### # Episodes: Total number of episodes for individual

<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 1-73] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-] [Mean=36.357 /-] [StdDev=16.951 /-]

### # StartDate: Start date of episode (inclusive)

<b>Information</b>	[Type= discrete] [Format=character] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-]

### # EndDate: End date of episode (inclusive)

<b>Information</b>	[Type= discrete] [Format=character] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-]

### # Days: Duration in days of episode

<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 1-366] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-] [Mean=166.124 /-] [StdDev=105.86 /-]

### # CalendarYear: Calendar year during episode

<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 2000-2024] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-] [Mean=2013.455 /-] [StdDev=7.306 /-]

### # Age: Age in completed years during episode

<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 0-1499] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-] [Mean=24.738 /-] [StdDev=18.448 /-]

### # AgeGroup: Age group during episode

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 0-18] [Missing=*]
--------------------	---

## File : SurveillanceEpisodesYrAge\_Coresidency

### # AgeGroup: Age group during episode

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	<1	122054	2.0%
1	01-04	555057	8.9%
2	05-09	731042	11.7%
3	10-14	727877	11.7%
4	15-19	721849	11.6%
5	20-24	697613	11.2%
6	25-29	615905	9.9%
7	30-34	497336	8.0%
8	35-39	382668	6.1%
9	40-44	284856	4.6%
10	45-49	220576	3.5%
11	50-54	175562	2.8%
12	55-59	142958	2.3%
13	60-64	115450	1.9%
14	65-69	84686	1.4%
15	70-74	62138	1.0%
16	75-79	42241	0.7%
17	80-84	28111	0.5%
18	85+	26978	0.4%

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

### # Born: Episode starts with the birth of the individual

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

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6185501	99.2%
1	Yes	49456	0.8%

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

### # Enumeration: Episode starts with an enumeration

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

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6096170	97.8%
1	Yes	138787	2.2%

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

### # InMigration: Episode starts with an in-migration

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

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-]

## File : SurveillanceEpisodesYrAge\_Coresidency

### # InMigration: Episode starts with an in-migration

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6077271	97.5%
1	Yes	157686	2.5%

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

### # LocationEntry: Episode starts with an internal migration

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6182043	99.2%
1	Yes	52914	0.8%

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

### # ExtResStart: Flag to indicate start of external residence

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6058873	97.2%
1	Yes	176084	2.8%

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

### # Participation: Resumed participation after refusal

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6234957	100.0%
1	Yes	0	

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

### # Died: Episode ends with the death of the individual

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6207253	99.6%
1	Yes	27704	0.4%

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

### # OutMigration: Episode ends with the out-migration

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

## File : SurveillanceEpisodesYrAge\_Coresidency

### # OutMigration: Episode ends with the out-migration

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6011135	96.4%
1	Yes	223822	3.6%

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

### # LocationExit: Episode ends with an internal migration

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6182043	99.2%
1	Yes	52914	0.8%

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

### # ExtResEnd: Flag to indicate end of external residence

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6141656	98.5%
1	Yes	93301	1.5%

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

### # Refusal: Individual refused follow-up

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6232241	100.0%
1	Yes	2716	0.0%

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

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

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

### # Current: Individual still under surveillance

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6075753	97.4%
1	Yes	159204	2.6%

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

## File : SurveillanceEpisodesYrAge\_Coresidency

### # HHRelationship: Relationship to head of household changed during this episode

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6165161	98.9%
1	Yes	69796	1.1%

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

### # MembershipStart: Flag to indicate start of household membership

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	5889718	94.5%
1	Yes	345239	5.5%

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

### # MembershipEnd: Flag to indicate end of household membership

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	5889718	94.5%
1	Yes	345239	5.5%

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

### # GapStart: Start of a gap in individual exposure record

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6220125	99.8%
1	Yes	14832	0.2%

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

### # GapEnd: End of gap in individual exposure record

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	6220125	99.8%
1	Yes	14832	0.2%

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

### # AgeChange: Flag age changed

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

## File : SurveillanceEpisodesYrAge\_Coresidency

### # AgeChange: Flag age changed

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	3415035	54.8%
1	Yes	2819922	45.2%

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

### # YrChange: Flag calendar year changed

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

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	3541184	56.8%
1	Yes	2693773	43.2%

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

### # Memberships: Number of concurrent household memberships

Information [Type= continuous] [Format=numeric] [Range= 1-28] [Missing=\*]

Statistics [NW/ W] [Valid=6234957 /-] [Invalid=0 /-] [Mean=1.07 /-] [StdDev=0.273 /-]

### # EducationStatus: Education level achieved during this calendar year

Information [Type= discrete] [Format=numeric] [Range= -1-99] [Missing=\*]

Statistics [NW/ W] [Valid=4570696 /-] [Invalid=1664261 /-]

Value	Label	Cases	Percentage
-1	Unknown	76725	1.7%
0	Grade 0	70207	1.5%
1	Grade 1	144873	3.2%
2	Grade 2	187306	4.1%
3	Grade 3	188089	4.1%
4	Grade 4	238191	5.2%
5	Grade 5	178280	3.9%
6	Grade 6	192193	4.2%
7	Grade 7	252758	5.5%
8	Grade 8	249631	5.5%
9	Grade 9	237799	5.2%
10	Grade 10	375622	8.2%
11	Grade 11	385322	8.4%
12	Grade 12	1454762	31.8%
13	Special School	0	
21	Higher Certificate	0	
22	National Diploma	0	
23	Degree	0	
24	Honours	0	
25	Masters	0	
26	Doctors	0	

## File : SurveillanceEpisodesYrAge\_Coresidency

### # EducationStatus: Education level achieved during this calendar year

Value	Label	Cases	Percentage
31	ABET 1	0	
32	ABET 2	0	
33	ABET 3	0	
34	ABET 4	0	
41	FET 1	0	
42	FET 2	0	
43	FET 3	0	
50	LT 1 year	0	
51	Creche	0	
96	Refused	0	
97	NAD	0	
98	No education	338938	7.4%
99	Not attending	0	
Sysmiss		1664261	

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

### # MotherDead: Mother is dead

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown	1049647	16.8%
0	No	3696148	59.3%
1	Yes	1489162	23.9%

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

### # MotherCoresident: Mother is co-resident

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-2] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown	1062187	17.0%
0	No	1769495	28.4%
1	Non-resident member	622492	10.0%
2	Resident member	2780783	44.6%

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

### # FatherDead: Father is dead

<b>Information</b>	[Type= discrete] [Format=numeric] [Range= -1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown	1857975	29.8%
0	No	1703465	27.3%
1	Yes	2673517	42.9%

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

## File : SurveillanceEpisodesYrAge\_Coresidency

### # FatherCoresident: Father is co-resident

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown	1897418	30.4%
0	No	2881866	46.2%
1	Non-resident member	533584	8.6%
2	Resident member	922089	14.8%

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

### # ParentalStatusChange: Change in parental status

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

**Statistics [NW/ W]** [Valid=6234957 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
-1	Unknown/Missing	0	
0	No	5704284	91.5%
1	Yes	530673	8.5%

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