

## **South Africa**

**Kobus Herbst, Africa Health Research Institute**  
**Willem Hanekom, Africa Health Research Institute**  
**Mark Siedner, Africa Health Research Institute**  
**Guy Harling, Africa Health Research Institute**

## **Network for Analysing Longitudinal Population- based HIV/AIDS data on Africa (ALPHA Network)**

### **Study Documentation**

July 31, 2024

# Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Identification	DDI.AHRI.AN01-03.Dataset.6.1

# Table of Contents

<a href="#">Overview.....</a>	<a href="#">4</a>
<a href="#">Scope &amp; Coverage.....</a>	<a href="#">4</a>
<a href="#">Producers &amp; Sponsors.....</a>	<a href="#">4</a>
<a href="#">Sampling.....</a>	<a href="#">5</a>
<a href="#">Data Collection.....</a>	<a href="#">5</a>
<a href="#">Accessibility.....</a>	<a href="#">5</a>
<a href="#">Files Description.....</a>	<a href="#">6</a>
<a href="#">AN01-03 Dataset 6.1.....</a>	<a href="#">6</a>
<a href="#">Variables List.....</a>	<a href="#">7</a>
<a href="#">AN01-03 Dataset 6.1.....</a>	<a href="#">7</a>
<a href="#">Variables Description.....</a>	<a href="#">8</a>
<a href="#">AN01-03 Dataset 6.1.....</a>	<a href="#">9</a>

## Network for Analysing Longitudinal Population-based HIV/AIDS data on Africa (ALPHA Network)

Overview	
<b>Identification</b>	AHRI.AN01-03.Dataset.6.1
<b>Abstract</b> The ALPHA Network is a collaboration between 10 longitudinal studies in sub-Saharan Africa. These studies collect data on HIV infection alongside demographic, behavioural, socio-economic and clinical data from residents of the study areas. The Network harmonises these data and conducts comparable and pooled analyses on HIV-related research questions. AHRI contributes data from ACDIS and PIP to the Network and this is referred to within the Network as uMkhanyakude. This dataset contains the sex and date of birth for the population resident in the study area and dates of migrations and deaths. It is used to construct the population denominator for all analyses.	
<b>Kind of Data</b>	Surveillance data
<b>Unit of Analysis</b>	Each observation describes an individual's period of residence in a household

Scope & Coverage	
<b>Keywords</b>	ALPHA; ALPHA Network; HIV incidence; HIV mortality; ANDLA Network; residency
<b>Topics</b>	Demography Epidemiology HIV Mortality Incidence
<b>Time Period(s)</b>	2000-2024
<b>Countries</b>	South Africa
<b>Geographic Coverage</b> Africa Health Research Institute Population Intervention Study Area (PIPSA)	
<b>Universe</b> Household members of the Africa Health Research Institute Population Intervention Programme Study Area	

Producers & Sponsors	
<b>Primary Investigator(s)</b>	Kobus Herbst, Africa Health Research Institute Willem Hanekom, Africa Health Research Institute Mark Siedner, Africa Health Research Institute Guy Harling, Africa Health Research Institute
<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>	Thobeka Mngomezulu , Data collection , AHRI Phumzile Dlamini , Data collection , AHRI Nomathamsanqa Majozi , Public Engagement , AHRI Siyabonga Nxumalo , Data extraction , AHRI Sweetness Dube , Data Documentation , AHRI Dickman Gareta , Data Management oversight , AHRI Eugene Ehlers , Software Development , AHRI Theresa Smit , Laboratory oversight , AHRI Gregory Ording-Jesperse , Laboratory data management , AHRI

## Sampling

### Sampling Procedure

Household demographic surveillance

## Data Collection

<b>Data Collection Dates</b>	start 2000-01-01 end 2024-06-04
------------------------------	------------------------------------

## 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.

### Citation Requirements

Herbst, K., Hanekom, W., Siedner, M., & Harling, G. (2024). Network for Analysing Longitudinal Population-based HIV/AIDS data on Africa (ALPHA Network) [Data set]. Africa Health Research Institute.  
DOI:<https://doi.org/10.23664/AHRI.AN01-03.DATASET.6.1>

# Files Description

Dataset contains 1 file(s)

AN01-03 Dataset 6.1	
# Cases	603551
# Variable(s)	12

# Variables List

Dataset contains 12 variable(s)

File AN01-03 Dataset 6.1							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	<a href="#">study_name</a>	Name of study field site	discrete	character-5	603551	0	-
2	<a href="#">idno</a>	Person ID number	continuous	numeric-12.0	603551	0	-
3	<a href="#">hhold_id</a>	Household ID number (AC = BSIIntId)	continuous	numeric-12.0	603551	0	-
4	<a href="#">hhold_id_.</a>	Household ID number (AC=HHIntId)	continuous	numeric-12.0	603523	28	-
5	<a href="#">sex</a>	Male or female	discrete	numeric-12.0	603551	0	-
6	<a href="#">dob</a>	Date of birth- best estimate	discrete	character-11	603551	-	-
7	<a href="#">residence</a>	Type of area within DSS	discrete	numeric-12.0	603551	0	-
8	<a href="#">event</a>	A code identifying the event that has occurred	discrete	numeric-12.0	603551	0	-
9	<a href="#">event_date</a>	Date on which the event occurred	discrete	character-11	603551	-	-
10	<a href="#">type_of_.</a>	Type of entry	discrete	numeric-12.0	603551	0	-
11	<a href="#">obs_date</a>	Observation date	discrete	character-11	603551	-	-
12	<a href="#">obs_round</a>	Observation round	continuous	numeric-12.0	603551	0	-

# Variables Description

**Dataset contains 12 variable(s)**



## File : AN01-03 Dataset 6.1

### # study\_name: Name of study field site

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-]

### # idno: Person ID number

Information	[Type= continuous] [Format=numeric] [Range= 11-275968] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-] [Mean=113488.146 /-] [StdDev=72318.751 /-]

### # hhold\_id: Household ID number (AC = BSIntId)

Information	[Type= continuous] [Format=numeric] [Range= 11-355646] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-] [Mean=15662.284 /-] [StdDev=37510.845 /-]

### # hhold\_id\_extra: Household ID number (AC=HHIntId)

Information	[Type= continuous] [Format=numeric] [Range= 11-47423] [Missing=*]
Statistics [NW/ W]	[Valid=603523 /-] [Invalid=28 /-] [Mean=14805.605 /-] [StdDev=12943.274 /-]

### # sex: Male or female

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Male	276054	45.7%
2	Female	327497	54.3%

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.

### # dob: Date of birth- best estimate

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-]

### # residence: Type of area within DSS

Information	[Type= discrete] [Format=numeric] [Range= 0-9] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Urban	48795	8.1%
1	Peri-urban	176336	29.2%
2	Rural	377428	62.5%
9		992	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.

### # event: A code identifying the event that has occurred

Information	[Type= discrete] [Format=numeric] [Range= 1-99] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	baseline recruitment	113318	18.8%
2	birth	39556	6.6%
3	external in-migration	136910	22.7%
4	Internal in-migration	62758	10.4%
5	Found after lost to follow up	0	

## File : AN01-03 Dataset 6.1

### # event: A code identifying the event that has occurred

Value	Label	Cases	Percentage
6	became eligible for study	0	
11	present in study site	0	
12	death	18485	3.1%
13	out-migration	164659	27.3%
14	Internal out-migration	54725	9.1%
15	lost to follow-up	7535	1.2%
16	became ineligible for study	2369	0.4%
99		3236	0.5%

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.

### # event\_date: Date on which the event occurred

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-]

### # type\_of\_date: Type of entry

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Reported by HH informant at interview	507284	84.0%
2	Reported by key informant	9735	1.6%
3	Imputed	86532	14.3%

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.

### # obs\_date: Observation date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-]

### # obs\_round: Observation round

Information	[Type= continuous] [Format=numeric] [Range= 1-47] [Missing=*]
Statistics [NW/ W]	[Valid=603551 /-] [Invalid=0 /-] [Mean=24.553 /-] [StdDev=15.75 /-]