

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

August 13, 2026

Metadata Production

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

Table of Contents

Overview.....	4
Scope & Coverage.....	4
Producers & Sponsors.....	4
Sampling.....	5
Data Collection.....	5
Accessibility.....	5
Files Description.....	6
AHRI.AN01-03.Dataset.6.1.....	6
Variables List.....	7
AHRI.AN01-03.Dataset.6.1.....	7
Variables Description.....	8
AHRI.AN01-03.Dataset.6.1.....	9

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

Overview	
Identification	AHRI.AN01-03.Dataset.6.1
Version	v1.1.0
Abstract 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.	
Kind of Data	Surveillance data
Unit of Analysis	Each observation describes an individual's period of residence in a household

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

Producers & Sponsors	
Primary Investigator(s)	Kobus Herbst, Africa Health Research Institute Willem Hanekom, Africa Health Research Institute Mark Siedner, Africa Health Research Institute Guy Harling, Africa Health Research Institute
Other Producer(s)	Africa Health Research Institute (AHRI)
Funding Agency/ies	Wellcome Trust (WT) , Core funding SAPRIN (SAPRIN)
Other Acknowledgment(s)	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

Data Collection Dates	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 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. 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 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)

AHRI.AN01-03.Dataset.6.1	
# Cases	592430
# Variable(s)	11

Variables List

Dataset contains 11 variable(s)

File AHRI.AN01-03.Dataset.6.1							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	idno	Person ID number	discrete	character-36	592430	0	-
2	study_name	Name of your study field site	discrete	character-5	592430	0	-
3	sex	Male or female	discrete	numeric-9.0	592430	0	-
4	dob	Date of birth- best estimate	discrete	character-11	592429	-	-
5	residence	Type of area within DSS	discrete	numeric-9.0	0	592430	-
6	entry_date	Date of start of residence episode	discrete	character-11	592430	-	-
7	entry_type	Type of entry	discrete	numeric-25.0	592430	0	-
8	exit_date	Date of end of residence episode	discrete	character-11	592430	-	-
9	exit_type	Type of exit	discrete	numeric-27.0	592430	0	-
10	hhold_id	Household ID number	discrete	character-36	592430	0	-
11	hhold_id_.	Household ID number	discrete	character-36	592059	0	-

Variables Description

Dataset contains 11 variable(s)

File : AHRI.AN01-03.Dataset.6.1

idno: Person ID number

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

Statistics [NW/ W] [Valid=592430 /-] [Invalid=0 /-]

study_name: Name of your study field site

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

Statistics [NW/ W] [Valid=592430 /-] [Invalid=0 /-]

sex: Male or female

Information [Type= discrete] [Format=numeric] [Range= 0-2] [Missing=*]

Statistics [NW/ W] [Valid=592430 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0		9	0.0%
1	Male	277273	46.8%
2	Female	315148	53.2%

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.

dob: Date of birth- best estimate

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

Statistics [NW/ W] [Valid=592429 /-]

residence: Type of area within DSS

Information [Type= discrete] [Format=numeric] [Missing=*]

Statistics [NW/ W] [Valid=0 /-] [Invalid=592430 /-]

entry_date: Date of start of residence episode

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

Statistics [NW/ W] [Valid=592430 /-]

entry_type: Type of entry

Information [Type= discrete] [Format=numeric] [Range= 1-4] [Missing=*]

Statistics [NW/ W] [Valid=592430 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Baseline recruitment	138787	23.4%
2	Birth	49456	8.3%
3	In-migration	210600	35.5%
4	Became eligible for study	193587	32.7%

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.

exit_date: Date of end of residence episode

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

Statistics [NW/ W] [Valid=592430 /-]

exit_type: Type of exit

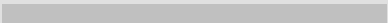
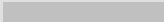
Information [Type= discrete] [Format=numeric] [Range= 1-5] [Missing=*]

Statistics [NW/ W] [Valid=592430 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Present in study site	159204	26.9%

File : AHRI.AN01-03.Dataset.6.1

exit_type: Type of exit

Value	Label	Cases	Percentage
2	Death	27704	 4.7%
3	Out-migration	276736	 46.7%
4	Lost to follow-up	12861	 2.2%
5	Became ineligible for study	115925	 19.6%

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.

hhold_id: Household ID number

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=592430 /-] [Invalid=0 /-]

hhold_id_extra: Household ID number

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=592059 /-] [Invalid=0 /-]