

**South Africa**

**Africa Health Research Institute  
Network for Analysing Longitudinal Population HIV/ AIDS data  
on Africa, London School of Hygiene and Tropical Medicine**

**ALPHA HIV Incidence and Mortality  
data, uMkhanyakude, South Africa**

**Study Documentation**

May 13, 2022

# Metadata Production

<b>Metadata Producer(s)</b>	Africa Health Research Institute (AHRI) , Metadata creator Network for Analysing Longitudinal Population HIV/ AIDS data on Africa (ALPHA) , London School of Hygiene and Tropical Medicine , Metadata creator DataFirst , University of Cape Town , Metadata creator
<b>Production Date</b>	May 21, 2022
<b>Version</b>	v01
<b>Identification</b>	zar-umkhanyakude-alpha-him-2002-2017-v1

# 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="#">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="#">incidence uMkhanyakude.....</a>	<a href="#">6</a>
<a href="#">mortality uMkhanyakude.....</a>	<a href="#">6</a>
<a href="#">Variables List.....</a>	<a href="#">8</a>
<a href="#">incidence uMkhanyakude.....</a>	<a href="#">8</a>
<a href="#">mortality uMkhanyakude.....</a>	<a href="#">8</a>
<a href="#">Variables Description.....</a>	<a href="#">10</a>
<a href="#">incidence uMkhanyakude.....</a>	<a href="#">11</a>
<a href="#">mortality uMkhanyakude.....</a>	<a href="#">14</a>

## ALPHA HIV Incidence and Mortality data, uMkhanyakude, South Africa (ALPHA HIM - U 2000-2017)

Overview	
Type	Health and Demographic Surveillance System (HDSS)
Identification	zar-umkhanyakude-alpha-him-2002-2017-v1
Version	Version 1: Edited, anonymised data available with restrictions
Series	The ALPHA (Analysing Longitudinal Population-based HIV/AIDS Data in Africa) Network is a collaboration among 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. The ALPHA HIV Incidence and Mortality data, uMkhanyakude is part of the international data harmonisation and analysis programme of the ALPHA network.
<p><b>Abstract</b></p> <p>This dataset is created from harmonising longitudinal population-based demographic surveillance and repeated serological survey data collected by Umkhanyakude study, South Africa. The data harmonisation is coordinated by the ALPHA Network which curates individual-level data on demographic surveillance, verbal autopsy interviews, serological and sexual behaviour surveys, and individually-linked data from HDSS and medical facilities.</p> <p>The data from the Umkhanyakude site in South Africa include two data files:</p> <p>HIV incidence data data file:</p> <p>This data file includes data on resident study participants aged 15-49 who have had one negative HIV test and at least one subsequent HIV test done within the study for research purposes. Data may be split into multiple observations for each person to describe time spent in different residencies, different age groups and different calendar year periods. All timings are given as age at event, instead of the dates. Dates of birth were rounded to the 15th of the month, and all other dates were offset by the same amount and the time elapsed between the rounded dates is used as the age at event. To enable multiple imputation of seroconversion dates we have provided the times (ages) of last negative test and first positive test.</p> <p>Mortality data</p> <p>This data file was created as described here (<a href="https://gatesopenresearch.org/articles/1-4">https://gatesopenresearch.org/articles/1-4</a>)</p>	
Kind of Data	Event history data
Unit of Analysis	Individuals

### Scope & Coverage

Time Period(s)	2000-2017
Countries	South Africa
<p><b>Geographic Coverage</b></p> <p>The data covers the Umkhanyakude study area in Kwazulu Natal, south Africa</p>	

### Producers & Sponsors

Primary Investigator(s)	Africa Health Research Institute
-------------------------	----------------------------------

	Network for Analysing Longitudinal Population HIV/ AIDS data on Africa, London School of Hygiene and Tropical Medicine
<b>Other Producer(s)</b>	
<b>Funding Agency/ies</b>	Wellcome Trust Bill and Melinda Gates Foundation
<b>Other Acknowledgment(s)</b>	Dr Kobus Herbst , Director of Population Science , AHRI Dickman Gareta , Data Scientist , AHRI

<b>Data Collection</b>	
<b>Data Collection Dates</b>	Mortality data: start 2000-01-01 Mortality data: end 2017-12-31 Incidence data: start 2005-04-10 Incidence data: end 2017-02-19
<b>Data Collection Mode</b>	Face-to-face [f2f]
<b>Data Collection Notes</b> ALPHA data are harmonised datasets from ALPHA member sites and this dataset is from the Africa Health Research Institute.	
<b>Questionnaires</b> The original unharmonised data was collected by the Africa Health Research Institute	
<b>Data Collector(s)</b>	Africa Health Research Institute

<b>Accessibility</b>	
<b>Access Authority</b>	DataFirst (University of Cape Town) , <a href="mailto:support@data1st.org">support@data1st.org</a>
<b>Contact(s)</b>	DataFirst helpdesk (University of Cape Town) , <a href="http://www.support.data1st.org">www.support.data1st.org</a> , <a href="mailto:support@data1st.org">support@data1st.org</a>
<b>Access Conditions</b> Licensed access data, available under conditions	
<b>Citation Requirements</b> Herbst, K. et al. (2022) “ALPHA HIV Incidence and Mortality data, uMkhanyakude, South Africa.” Africa Health Research Institute. doi: 10.23664/ZAR-UMKHANYAKUDE-ALPHA-HIM-2002-2017.	

# Files Description

Dataset contains 2 file(s)

incidence_uMkhanyakude	
# Cases	59996
# Variable(s)	15
<b>Notes</b> uMkhanyakude J:/ALPHA\data/uMkhanyakude/ALPHA6_spec1_uMkhanyakude.dta saved 23 Nov 2018 15:50 uMkhanyakude J:/ALPHA\data/uMkhanyakude/ALPHA6_spec2b_uMkhanyakude.dta saved 11 Jul 2018 08:36	

mortality_uMkhanyakude	
# Cases	367505
# Variable(s)	14
<b>Notes</b> Kisesa J:/ALPHA\data/Kisesa/ALPHA6_spec1_Kisesa.dta saved 18 Jul 2018 17:08 sero4clean was last created on 12:09:16 22 Mar 2007 21 March 2007 created two new variables orig_studynr4 and orig_id4 added 21 March 2007 retained extra variables q4list other_activ other_places end of sero4clean.do notes sero123link was last created on 12:09:11 22 Mar 2007 21 March 2007 created new variables orig_studynr3, orig_id3, orig_studynr2, orig_id2, orig_studynr1 orig_id1 to keep all information on identifiers before they were changed during cleaning this change was made replace id3= 220102701 if studynr== 21818 end of notes from weakserolink.do last created on 12:04:58 22 Mar 2007 last created on 12:04:59 22 Mar 2007 extra variables about marriage have been retained spouline14 wifekijiji14 husbandkijiji14 living14 married14 demrnd14 was last created on 12:04:59 22 Mar 2007 extra variables about migration have been retained xlalanje16 xsablala16 xlalawapi16 xhamalini16 xsabhama16 xhamasab16 xhamawapi16 xalıklala16 xalikhama16 xaliktoka16 demrnd15 was last created on 12:05:01 22 Mar 2007 extra variables about migration have been retained xlalanje16 xsablala16 xlalawapi16 xhamalini16 xsabhama16 xhamasab16 hamawapi16 aliklala16 alikhama16 aliktoka16 the wapi variable has been recoded (replace wapi16=9 if wapi==10) so that 10 is now added to 9. Those with value 10 where actually people who had migrated further than mwanza but somehow the code became 10 demrnd16 was last created on 12:05:03 22 Mar 2007 extra variables about migration have been retained xlalanje17 xsablala17 xlalawapi17 xhamalini17 xsabhama17 xhamasab17 hamawapi17 aliklala17 alikhama17 aliktoka17 the wapi variable has been recoded (replace wapi17=9 if wapi==10) so that 10 is now added to 9. Those with value 10 where actually people who had migrated further than mwanza but somehow the code became 10 demrnd17 last created on 12:05:05 22 Mar 2007 extra variables about migration have been retained xlalanje18 xsablala18 xlalawapi18 xhamalini18 xsabhama18 xhamasab18 hamawapi18 aliklala18 alikhama18 aliktoka18 the wapi variable has been recoded (replace wapi18=9 if wapi==10) so that 10 is now added to 9. Those with value 10 where actually people who had migrated further than mwanza but somehow the code became 10 demrnd18 was last created on 12:05:06 22 Mar 2007 extra variables about migration have been retained xlalanje19 xsablala19 xlalawapi19 xhamalini19 xsabhama19 xhamasab19 hamawapi19 aliklala19 alikhama19 aliktoka19 extra variables about schools has been added wilaya kata schoolname schoolcode	

the wapi variable has been recoded (replace wapi19=9 if wapi==10) so that 10 is now added to 9. Those with value 10 where actually people who had migrated further than mwanza but somehow the code became 10

demrnd19 was last created on 12:05:09 22 Mar 2007

link0119 was last created on 12:08:59 22 Mar 2007

end of notes from linkrnds.do

sero1234link was last created on 12:09:23 22 Mar 2007

end of notes from weaksero1234.do

remember kij2sero3 is a file for sero3 kijiji2

last created on 10:21:08 26 Mar 2007

last created on 10:21:09 26 Mar 2007

demrnd14 was last created on 10:21:10 26 Mar 2007

demrnd15 was last created on 10:21:11 26 Mar 2007

demrnd16 was last created on 10:21:13 26 Mar 2007

demrnd17 last created on 10:21:14 26 Mar 2007

demrnd18 was last created on 10:21:16 26 Mar 2007

demrnd19 was last created on 10:21:18 26 Mar 2007

link0119 was last created on 10:31:54 26 Mar 2007

Created at 10:16:45 on 7 Oct 2009

Created at 17:30:56 on 19 Mar 2008

Created at 15:10:21 on 16 Mar 2009

merged sero12345 and demogr1to24 files

Kisesa J:/ALPHA\data/Kisesa/ALPHA6\_spec2b\_Kisesa.dta saved 18 Jul 2018 17:08

Manicaland J:/ALPHA\data/Manicaland/ALPHA6\_spec1\_Manicaland.dta saved

Manicaland J:/ALPHA\data/Manicaland/ALPHA6\_spec2b\_Manicaland.dta saved

Masaka J:/ALPHA\data/Masaka/ALPHA6\_spec1\_Masaka.dta saved 10 Sep 2018 12:53

Masaka J:/ALPHA\data/Masaka/ALPHA6\_spec2b\_Masaka.dta saved 12 Feb 2018 12:13

Ifakara J:/ALPHA\data/Ifakara/ALPHA6\_spec1\_Ifakara.dta saved 11 Apr 2016 22:57

Ifakara J:/ALPHA\data/Ifakara/ALPHA6\_spec2b\_Ifakara.dta saved 15 Apr 2016 11:31

Karonga J:/ALPHA\data/Karonga/ALPHA6\_spec1\_Karonga.dta saved 12 Jan 2018 11:49

Karonga J:/ALPHA\data/Karonga/ALPHA6\_spec2b\_Karonga.dta saved 12 Jan 2018 11:49

Rakai J:/ALPHA\data/Rakai/ALPHA6\_spec1\_Rakai.dta saved 4 Jul 2018 16:43

Rakai c:/emma/ALPHA\data/Rakai/ALPHA6\_spec2b\_Rakai.dta saved 26 Jun 2018 13:01

uMkhanyakude J:/ALPHA\data/uMkhanyakude/ALPHA6\_spec1\_uMkhanyakude.dta saved 23 Nov 2018 15:50

uMkhanyakude J:/ALPHA\data/uMkhanyakude/ALPHA6\_spec2b\_uMkhanyakude.dta saved 11 Jul 2018 08:36

Agincourt J:/ALPHA\data/Agincourt/ALPHA6\_spec1\_Agincourt.dta saved 12 Apr 2016 15:59

Agincourt J:/ALPHA\data/Agincourt/ALPHA6\_spec2b\_Agincourt.dta saved 12 Apr 2016 15:59

# Variables List

Dataset contains 29 variable(s)

File incidence_uMkhanyakude							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	<a href="#">study_name</a>	Study name	discrete	numeric-12.0	59996	0	-
2	<a href="#">idno</a>	Participant ID number	continuous	numeric-9.0	59996	0	-
3	<a href="#">sex</a>	Sex	discrete	numeric-12.0	59996	0	-
4	<a href="#">fouryear</a>	Calendar year, grouped in 4 years post 2005	discrete	numeric-9.0	59996	0	-
5	<a href="#">agegrp</a>	Five year age group	discrete	numeric-9.0	59996	0	-
6	<a href="#">timein</a>	Start of episode (age)	continuous	numeric-10.0	59996	0	-
7	<a href="#">timeout</a>	End of episode (age)	continuous	numeric-10.0	59996	0	-
8	<a href="#">timelast_..</a>	Age at last negative test	continuous	numeric-10.0	59996	0	-
9	<a href="#">timefirs_..</a>	Age at first positive test	continuous	numeric-10.0	10352	49644	-
10	<a href="#">fail</a>	Seroconversion occurred at the end of this episode	discrete	numeric-9.0	57637	2359	-
11	<a href="#">t0</a>	Stata variable: age at start of episode	continuous	numeric-10.0	59996	0	-
12	<a href="#">t</a>	Stata variable: age at end of episode	continuous	numeric-10.0	59996	0	-
13	<a href="#">st</a>	Stata variable: in survival analysis	discrete	numeric-8.0	59996	0	-
14	<a href="#">d</a>	Stata variable: failure (seroconversion)	discrete	numeric-8.0	59996	0	-
15	<a href="#">collapse_..</a>	-	continuous	numeric-9.0	59996	0	-

File mortality_uMkhanyakude							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	<a href="#">study_name</a>	Study name	discrete	numeric-12.0	367505	0	-
2	<a href="#">idno</a>	Participant ID number	continuous	numeric-9.0	367505	0	-
3	<a href="#">sex</a>	Sex	discrete	numeric-8.0	367505	0	-
4	<a href="#">fouryear</a>	Calendar year, grouped in 4 years post 2005	discrete	numeric-13.0	367505	0	-
5	<a href="#">agegrp</a>	Five year age group	discrete	numeric-9.0	367505	0	-
6	<a href="#">hivstatu_..</a>	HIV status	discrete	numeric-9.0	367505	0	-
7	<a href="#">timein</a>	Start of episode (age)	continuous	numeric-10.0	367505	0	-
8	<a href="#">timeout</a>	End of episode (age)	continuous	numeric-10.0	367505	0	-
9	<a href="#">fail</a>	Seroconversion occurred at the end of this episode	discrete	numeric-9.0	353799	13706	-
10	<a href="#">t0</a>	Stata variable: age at start of episode	continuous	numeric-10.0	367505	0	-
11	<a href="#">t</a>	Stata variable: age at end of episode	continuous	numeric-10.0	367505	0	-



<b>File mortality_uMkhanyakude</b>							
#	Name	Label	Type	Format	Valid	Invalid	Question
12	<a href="#">st</a>	Stata variable: in survival analysis	discrete	numeric-8.0	367505	0	-
13	<a href="#">d</a>	Stata variable: failure (seroconversion)	discrete	numeric-8.0	367505	0	-
14	<a href="#">collapse ..</a>	-	continuous	numeric-9.0	367505	0	-

# Variables Description

Dataset contains 29 variable(s)

File : incidence\_uMkhanyakude

# study\_name: Study name

Information

[Type= discrete] [Format=numeric] [Range= 1-10] [Missing=\*]

Statistics [NW/ W]

[Valid=59996 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Karonga	0	
2	Kisesa	0	
3	Manicaland	0	
4	Masaka	0	
5	Rakai	0	
6	uMkhanyakude	59996	<div>100.0%</div>
7	Agincourt	0	
8	Kisumu	0	
9	Ifakara	0	
10	Nairobi	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.

# idno: Participant ID number

Information

[Type= continuous] [Format=numeric] [Range= 1-20652] [Missing=\*]

Statistics [NW/ W]

[Valid=59996 /-] [Invalid=0 /-] [Mean=9856.716 /-] [StdDev=5784.316 /-]

# sex: Sex

Information

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

Statistics [NW/ W]

[Valid=59996 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Men	23701	<div>39.5%</div>
2	Women	36295	<div>60.5%</div>

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.

# fouryear: Calendar year, grouped in 4 years post 2005

Information

[Type= discrete] [Format=numeric] [Range= 0-6] [Missing=\*]

Statistics [NW/ W]

[Valid=59996 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	1995-99	0	
1	2000-04	7700	<div>12.8%</div>
2	2005-08	18620	<div>31.0%</div>
3	2009-12	17797	<div>29.7%</div>
4	2013-16	15879	<div>26.5%</div>
5	2017-20	0	
6	2021-24	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.

# agegrp: Five year age group

Information

[Type= discrete] [Format=numeric] [Range= 0-18] [Missing=\*]

Statistics [NW/ W]

[Valid=59996 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	0-4	0	

## File : incidence\_uMkhanyakude

### # agegrp: Five year age group

Value	Label	Cases	Percentage
1	5-9	0	
2	10-14	0	
3	15-19	19506	32.5%
4	20-24	15285	25.5%
5	25-29	7232	12.1%
6	30-34	4345	7.2%
7	35-39	3949	6.6%
8	40-44	4488	7.5%
9	45-49	5191	8.7%
10	50-54	0	
11	55-59	0	
12	60-64	0	
13	65-69	0	
14	70-74	0	
15	75-79	0	
16	80-84	0	
17	85-89	0	
18	90+	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.

### # timein: Start of episode (age)

Information	[Type= continuous] [Format=numeric] [Range= 15-49.9931553730322] [Missing=*]
Statistics [NW/ W]	[Valid=59996 /-] [Invalid=0 /-] [Mean=26.169 /-] [StdDev=9.881 /-]

### # timeout: End of episode (age)

Information	[Type= continuous] [Format=numeric] [Range= 15.0088980150582-50] [Missing=*]
Statistics [NW/ W]	[Valid=59996 /-] [Invalid=0 /-] [Mean=27.553 /-] [StdDev=10.009 /-]

### # timelastneg: Age at last negative test

Information	[Type= continuous] [Format=numeric] [Range= 13.7905544147844-63.854893908282] [Missing=*]
Statistics [NW/ W]	[Valid=59996 /-] [Invalid=0 /-] [Mean=30.215 /-] [StdDev=11.676 /-]

### # timefirstpos: Age at first positive test

Information	[Type= continuous] [Format=numeric] [Range= 16.1916495550992-61.3990417522245] [Missing=*]
Statistics [NW/ W]	[Valid=10352 /-] [Invalid=49644 /-] [Mean=29.461 /-] [StdDev=8.19 /-]

### # fail: Seroconversion occurred at the end of this episode

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=57637 /-] [Invalid=2359 /-]

Value	Label	Cases	Percentage
0	No	54965	95.4%
1	Yes	2672	4.6%
Sysmiss		2359	

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.

<b>File : incidence_uMkhanyakude</b>	
<b># t0: Stata variable: age at start of episode</b>	
<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 15-49.9931553730322] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=59996 /-] [Invalid=0 /-] [Mean=26.169 /-] [StdDev=9.881 /-]
<b># t: Stata variable: age at end of episode</b>	
<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 15.0088980150582-50] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=59996 /-] [Invalid=0 /-] [Mean=27.553 /-] [StdDev=10.009 /-]
<b># st: Stata variable: in survival analysis</b>	
<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=59996 /-] [Invalid=0 /-]
<b># d: Stata variable: failure (seroconversion)</b>	
<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=59996 /-] [Invalid=0 /-]
<b># collapsegrp</b>	
<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 0-32] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=59996 /-] [Invalid=0 /-] [Mean=0.932 /-] [StdDev=3.288 /-]

File : mortality\_uMkhanyakude

# study\_name: Study name

Information

[Type= discrete] [Format=numeric] [Range= 1-10] [Missing=\*]

Statistics [NW/ W]

[Valid=367505 -/] [Invalid=0 -/]

Value	Label	Cases	Percentage
1	Karonga	0	
2	Kisesa	0	
3	Manicaland	0	
4	Masaka	0	
5	Rakai	0	
6	uMkhanyakude	367505	100.0%
7	Agincourt	0	
8	Kisumu	0	
9	Ifakara	0	
10	Nairobi	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.

# idno: Participant ID number

Information

[Type= continuous] [Format=numeric] [Range= 210170-301565] [Missing=\*]

Statistics [NW/ W]

[Valid=367505 -/] [Invalid=0 -/] [Mean=251606.32 -/] [StdDev=25250.931 -/]

# sex: Sex

Information

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

Statistics [NW/ W]

[Valid=367505 -/] [Invalid=0 -/]

Value	Label	Cases	Percentage
1	Men	161374	43.9%
2	Women	206131	56.1%

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.

# fouryear: Calendar year, grouped in 4 years post 2005

Information

[Type= discrete] [Format=numeric] [Range= 0-6] [Missing=\*]

Statistics [NW/ W]

[Valid=367505 -/] [Invalid=0 -/]

Value	Label	Cases	Percentage
0	earliest-1999	0	
1	2000-04	97701	26.6%
2	2005-08	89551	24.4%
3	2009-12	92298	25.1%
4	2013-16	87955	23.9%
5	2017-20	0	
6	2021-24	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.

# agegrp: Five year age group

Information

[Type= discrete] [Format=numeric] [Range= 0-18] [Missing=\*]

Statistics [NW/ W]

[Valid=367505 -/] [Invalid=0 -/]

Value	Label	Cases	Percentage
0	0-4	0	

## File : mortality\_uMkhanyakude

### # agegrp: Five year age group

Value	Label	Cases	Percentage
1	5-9	0	
2	10-14	0	
3	15-19	105214	28.6%
4	20-24	80997	22.0%
5	25-29	56471	15.4%
6	30-34	41567	11.3%
7	35-39	32636	8.9%
8	40-44	27387	7.5%
9	45-49	23233	6.3%
10	50-54	0	
11	55-59	0	
12	60-64	0	
13	65-69	0	
14	70-74	0	
15	75-79	0	
16	80-84	0	
17	85-89	0	
18	90+	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.

### # hivstatus\_broad: HIV status

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]		
Statistics [NW/ W]	[Valid=367505 /-] [Invalid=0 /-]		
Value	Label	Cases	Percentage
1	Negative	85437	23.2%
2	Positive	35714	9.7%
3	Unknown	246354	67.0%

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.

### # timein: Start of episode (age)

Information	[Type= continuous] [Format=numeric] [Range= 15-49.9986310746064] [Missing=*]
Statistics [NW/ W]	[Valid=367505 /-] [Invalid=0 /-] [Mean=26.508 /-] [StdDev=9.472 /-]

### # timeout: End of episode (age)

Information	[Type= continuous] [Format=numeric] [Range= 15.0006844626968-50] [Missing=*]
Statistics [NW/ W]	[Valid=367505 /-] [Invalid=0 /-] [Mean=27.93 /-] [StdDev=9.572 /-]

### # fail: Seroconversion occurred at the end of this episode

Information	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=353799 /-] [Invalid=13706 /-]

Value	Label	Cases	Percentage
0	No	347590	98.2%
1	Yes	6209	1.8%
Sysmiss		13706	

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.

<b>File : mortality_uMkhanyakude</b>	
<b># t0: Stata variable: age at start of episode</b>	
<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 15-49.9986310746064] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=367505 /-] [Invalid=0 /-] [Mean=26.508 /-] [StdDev=9.472 /-]
<b># t: Stata variable: age at end of episode</b>	
<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 15.0006844626968-50] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=367505 /-] [Invalid=0 /-] [Mean=27.93 /-] [StdDev=9.572 /-]
<b># st: Stata variable: in survival analysis</b>	
<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 1-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=367505 /-] [Invalid=0 /-]
<b># d: Stata variable: failure (seroconversion)</b>	
<b>Information</b>	[Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=367505 /-] [Invalid=0 /-]
<b># collapsegrp</b>	
<b>Information</b>	[Type= continuous] [Format=numeric] [Range= 0-45] [Missing=*]
<b>Statistics [NW/ W]</b>	[Valid=367505 /-] [Invalid=0 /-] [Mean=3.773 /-] [StdDev=6.846 /-]