

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

Sweetness H Dube

Report generated on: July 27, 2022

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

Overview

Identification

ID NUMBER

zar-umkhanyakude-alpha-him-2002-2017-v1

Version

VERSION DESCRIPTION

Version 1: Edited, anonymised data available with restrictions

Overview

ABSTRACT

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.

The data from the Umkhanyakude site in South Africa include two data files:

HIV incidence data data file:

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.

Mortality data

This data file was created as described here (<https://gatesopenresearch.org/articles/1-4>)

KIND OF DATA

Event History Data

UNITS OF ANALYSIS

Individuals

KEYWORDS

ALPHA Network Dataset

Coverage

GEOGRAPHIC COVERAGE

The data covers the Umkhanyakude study area in Kwazulu Natal, south Africa

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

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

FUNDING

Name	Abbreviation	Role
Wellcome Trust		
Bill and Melinda Gates Foundation		

OTHER ACKNOWLEDGEMENTS

Name	Affiliation	Role
Dr Kobus Herbst	AHRI	Director of Population Science
Dickman Gareta	AHRI	Data Scientist

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
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

DDI DOCUMENT ID

zar-umkhanyakude-alpha-him-2002-2017-v1

Sampling

No content available

Questionnaires

No content available

Data Collection

Data Collection Dates

Start	End	Cycle
2000-01-01	2017-12-31	Mortality data
2005-04-10	2017-02-19	Incidence data

Data Processing

No content available

Data Appraisal

No content available

File Description

Variable List

mortality_uMkhanyakude

Content

Cases 367505

Variable(s) 14

Structure Type:
Keys: ()

Version

Producer

Missing Data

Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V1	study_name	Study name	discrete	numeric	
V2	idno	Participant ID number	contin	numeric	
V3	sex	Sex	discrete	numeric	
V4	fouryear	Calendar year, grouped in 4 years post 2005	discrete	numeric	
V5	agegrp	Five year age group	discrete	numeric	
V6	hivstatus_broad	HIV status	discrete	numeric	
V7	timein	Start of episode (age)	contin	numeric	
V8	timeout	End of episode (age)	contin	numeric	
V9	fail	Seroconversion occurred at the end of this episode	discrete	numeric	
V10	t0	Stata variable: age at start of episode	contin	numeric	
V11	t	Stata variable: age at end of episode	contin	numeric	
V12	st	Stata variable: in survival analysis	discrete	numeric	
V13	d	Stata variable: failure (seroconversion)	discrete	numeric	
V14	collapsegrp		contin	numeric	

incidence_uMkhanyakude

Content

Cases 59996

Variable(s) 15

Structure
Type:
Keys: ()

Version

Producer

Missing Data

Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V15	study_name	Study name	discrete	numeric	
V16	idno	Participant ID number	contin	numeric	
V17	sex	Sex	discrete	numeric	
V18	fouryear	Calendar year, grouped in 4 years post 2005	discrete	numeric	
V19	agegrp	Five year age group	discrete	numeric	
V20	timein	Start of episode (age)	contin	numeric	
V21	timeout	End of episode (age)	contin	numeric	
V22	timelastneg	Age at last negative test	contin	numeric	
V23	timefirstpos	Age at first positive test	contin	numeric	
V24	fail	Seroconversion occurred at the end of this episode	discrete	numeric	
V25	t0	Stata variable: age at start of episode	contin	numeric	
V26	t	Stata variable: age at end of episode	contin	numeric	
V27	st	Stata variable: in survival analysis	discrete	numeric	
V28	d	Stata variable: failure (seroconversion)	discrete	numeric	
V29	collapsegrp		contin	numeric	

Study name (study_name)

File: mortality_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-10

Valid cases: 367505
Invalid: 0

Participant ID number (idno)

File: mortality_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 9
Decimals: 0
Range: 210170-301565

Valid cases: 367505
Invalid: 0
Minimum: 210170
Maximum: 301565
Mean: 251606.3
Standard deviation: 25250.9

Sex (sex)

File: mortality_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 8
Decimals: 0
Range: 1-2

Valid cases: 367505
Invalid: 0

Calendar year, grouped in 4 years post 2005 (fouryear)

File: mortality_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 13
Decimals: 0
Range: 0-6

Valid cases: 367505
Invalid: 0

Five year age group (agegrp)

File: mortality_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 0-18

Valid cases: 367505
Invalid: 0

HIV status (hivstatus_broad)

File: mortality_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 1-3

Valid cases: 367505
Invalid: 0

Start of episode (age) (timein)

File: mortality_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 15-49.9986310746064

Valid cases: 367505
Invalid: 0
Minimum: 15
Maximum: 50
Mean: 26.5
Standard deviation: 9.5

End of episode (age) (timeout)

File: mortality_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 15.0006844626968-50

Valid cases: 367505
Invalid: 0
Minimum: 15
Maximum: 50
Mean: 27.9
Standard deviation: 9.6

Seroconversion occurred at the end of this episode (fail)

File: mortality_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 0-1

Valid cases: 353799
Invalid: 13706

Stata variable: age at start of episode (t0)

File: mortality_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 15-49.9986310746064

Valid cases: 367505
Invalid: 0
Minimum: 15
Maximum: 50
Mean: 26.5
Standard deviation: 9.5

Stata variable: age at end of episode (t)

File: mortality_uMkhanyakude

Overview

Type: Continuous	Valid cases: 367505
Format: numeric	Invalid: 0
Width: 10	Minimum: 15
Decimals: 0	Maximum: 50
Range: 15.0006844626968-50	Mean: 27.9
	Standard deviation: 9.6

Stata variable: in survival analysis (st)

File: mortality_uMkhanyakude

Overview

Type: Discrete	Valid cases: 367505
Format: numeric	Invalid: 0
Width: 8	
Decimals: 0	
Range: 1-1	

Stata variable: failure (seroconversion) (d)

File: mortality_uMkhanyakude

Overview

Type: Discrete	Valid cases: 367505
Format: numeric	Invalid: 0
Width: 8	
Decimals: 0	
Range: 0-1	

(collapsegrp)

File: mortality_uMkhanyakude

Overview

Type: Continuous	Valid cases: 367505
Format: numeric	Invalid: 0
Width: 9	Minimum: 0
Decimals: 0	Maximum: 45
Range: 0-45	Mean: 3.8
	Standard deviation: 6.8

Study name (study_name)

File: incidence_uMkhanyakude

Overview

Type: Discrete	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 12	
Decimals: 0	
Range: 1-10	

Participant ID number (idno)

File: incidence_uMkhanyakude

Overview

Type: Continuous	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 9	Minimum: 1
Decimals: 0	Maximum: 20652
Range: 1-20652	Mean: 9856.7
	Standard deviation: 5784.3

Sex (sex)

File: incidence_uMkhanyakude

Overview

Type: Discrete	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 12	
Decimals: 0	
Range: 1-2	

Calendar year, grouped in 4 years post 2005 (fouryear)

File: incidence_uMkhanyakude

Overview

Type: Discrete	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 9	
Decimals: 0	
Range: 0-6	

Five year age group (agegrp)

File: incidence_uMkhanyakude

Overview

Type: Discrete	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 9	
Decimals: 0	
Range: 0-18	

Start of episode (age) (timein)

File: incidence_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 15-49.9931553730322

Valid cases: 59996
Invalid: 0
Minimum: 15
Maximum: 50
Mean: 26.2
Standard deviation: 9.9

End of episode (age) (timeout)

File: incidence_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 15.0088980150582-50

Valid cases: 59996
Invalid: 0
Minimum: 15
Maximum: 50
Mean: 27.6
Standard deviation: 10

Age at last negative test (timelastneg)

File: incidence_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 13.7905544147844-63.854893908282

Valid cases: 59996
Invalid: 0
Minimum: 13.8
Maximum: 63.9
Mean: 30.2
Standard deviation: 11.7

Age at first positive test (timefirstpos)

File: incidence_uMkhanyakude

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 16.1916495550992-61.3990417522245

Valid cases: 10352
Invalid: 49644
Minimum: 16.2
Maximum: 61.4
Mean: 29.5
Standard deviation: 8.2

Seroconversion occurred at the end of this episode (fail)

File: incidence_uMkhanyakude

Overview

Type: Discrete
Format: numeric
Width: 9
Decimals: 0
Range: 0-1

Valid cases: 57637
Invalid: 2359

Stata variable: age at start of episode (t0)

File: incidence_uMkhanyakude

Overview

Type: Continuous	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 10	Minimum: 15
Decimals: 0	Maximum: 50
Range: 15-49.9931553730322	Mean: 26.2
	Standard deviation: 9.9

Stata variable: age at end of episode (t)

File: incidence_uMkhanyakude

Overview

Type: Continuous	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 10	Minimum: 15
Decimals: 0	Maximum: 50
Range: 15.0088980150582-50	Mean: 27.6
	Standard deviation: 10

Stata variable: in survival analysis (st)

File: incidence_uMkhanyakude

Overview

Type: Discrete	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 8	
Decimals: 0	
Range: 1-1	

Stata variable: failure (seroconversion) (d)

File: incidence_uMkhanyakude

Overview

Type: Discrete	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 8	
Decimals: 0	
Range: 0-1	

(collapsegrp)

File: incidence_uMkhanyakude

Overview

Type: Continuous	Valid cases: 59996
Format: numeric	Invalid: 0
Width: 9	Minimum: 0
Decimals: 0	Maximum: 32
Range: 0-32	Mean: 0.9
	Standard deviation: 3.3

Documentation

Technical documents

DDI:ALPHA HIV Incidence and Mortality Data

Title	DDI:ALPHA HIV Incidence and Mortality Data
Author(s)	Sweetness H Dube
Date	13/05.2022
Country	South Africa
Language	English
Publisher(s)	Sweetness H Dube
Filename	zar-umkhanyakude-alpha-him-2002-2017-v1.pdf
