

Tanzania - Magu Household Demographic Surveillance System Dataset 1994 - 2010

John Chagalucha - National Institute For Medical Research

Report generated on: July 24, 2015

Visit our data catalog at: <http://indepth-ishare.org/index.php>

Overview

Identification

ID NUMBER

INDEPTH.TZ021.3.v5

Version

VERSION DESCRIPTION

v1: Raw Data, and second version for Internal Use Only

v5: Cleaned and anonymised dataset for public distribution

PRODUCTION DATE

2013-06-18

Overview

ABSTRACT

Magu Demographic Surveillance System

The TAZAMA DSS study was set up in 1994 as part of a Tanzania - Netherlands collaboration to provide a framework for interventions to combat the spread of HIV, and was then known as TANESA. The original focus was on measuring the effects of prevention activities. Since 2005 the project has broadened its focus to include provision of a frame for studies involving HIV treatment and care.

The demographic surveillance system collects information on births and deaths and movements in and out of the households. It helps to understand the population dynamics in the study area including fertility, mortality and migration patterns. It provides information on the structure of families that live together. The demographic surveillance system study is also used to identify people who are eligible to participate in the serological-surveys (the right age group, and continuously resident rather than just visiting). It provides the data for calculating the denominators for demographic rates.

The objectives of the Magu DSS

To measure child and adult mortality and fertility in the general population,

To lay a reliable foundation for HIV epidemiological surveys by registering people's mobility and mortality

To assess the leading causes of death through verbal autopsy

To assess the potential changes in family structure due to the adult mortality (e.g. orphan hood)

To provide Vital Data to the Ward and District authorities for planning purposes.

Verbal Autopsy

In the Magu DSS, Verbal autopsy work started in 1994 just after the DSS's baseline census. VA interviews are carried out after every round of the DSS. The available data are for verbal autopsies corresponding to the DSS rounds done till 2010

Three questionnaires are used; one for neonates 0-28 days, one for children aged 29 days-14years and the one for adults 15 years and over.

KIND OF DATA

Event history data

UNITS OF ANALYSIS

Individual

Scope

NOTES

Fertility, Mortality and migration

TOPICS

Topic	Vocabulary	URI
Demography [N01.224]	MeSH	http://www.ncbi.nlm.nih.gov/mesh

KEYWORDS

Fertility, Mortality, Migration

Coverage

GEOGRAPHIC COVERAGE

Magu HDSS

UNIVERSE

The survey covered currently resident household members

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
John Changalucha	National Institute For Medical Research

OTHER PRODUCER(S)

Name	Affiliation	Role
Baltazar Mtenga	Tanzania National Institute For Medical Research	Database Programmer & Data Manager
David Belckse	Tanzania National Institute For Medical Research	IT Consultant

FUNDING

Name	Abbreviation	Role
Global Fund		
WelcomeTrust		

OTHER ACKNOWLEDGEMENTS

Name	Affiliation	Role
Baltazar Mtenga	Tanzania National Institute For Medical Research - Mwanza Centre	Database Programmer & Data Manager
Jonas Aswile	Tanzania National Institute For Medical Research - Mwanza Centre	Data Manager
Mponjori Samwel	Tanzania National Institute For Medical Research - Mwanza Centre	Data Manager
David Belckse	Tanzania National Institute For Medical Research - Mwanza Centre	IT Consultant
Ally Abeid	Tanzania National Institute For Medical Research - Mwanza Centre	System Administrator

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
------	--------------	-------------	------

Name	Abbreviation	Affiliation	Role
Tazama Project	TZ021	Tanzania National Institute For Medical Research - Mwanza Centre	Documentation of INDEPTH Network Micro DataSet
	TZ021	London School of Tropical Hygen Medicine	Documentation of INDEPTH Network Micro DataSet
	TZ021	Mwanza Clinical Trial Unit	Documentation of INDEPTH Network Micro DataSet

DATE OF METADATA PRODUCTION

2015-07-21

DDI DOCUMENT VERSION

Version 1 (January 2013)

Version 5 (June 2013)

Version 9 (July 2013)

Version 10 (July 2015)

DDI DOCUMENT ID

DDI.INDEPTH.TZ021.3.v10

Sampling

Sampling Procedure

This dataset is not based on a sample but contains information from the complete demographic surveillance area.

Deviations from Sample Design

None

Response Rate

99.99%

Weighting

Not applicable

Questionnaires

Overview

List of questionnaires:

Household and individual registration

Pregnancy

Birth

Death

Migration (In and Out)

Data Collection

Data Collection Dates

Start	End	Cycle
1994-02-01	2010-12-31	N/A

Data Collection Mode

Proxy Respondent [proxy]

DATA COLLECTION NOTES

Enumerators were trained initially and then refresher trainings conducted prior to each surveillance round.

Data Collectors

Name	Abbreviation	Affiliation
Magu Household Demographic Surveillance System	TZ021	

SUPERVISION

Fieldworkers operated in teams supervised by a Fieldwork supervisor. Supervisors conduct supervised visits and quality control visits and review fieldworkers data collection.

Data Processing

Data Editing

Data editing took place at different stages of the data processing.

On data entry, data consistency and plausibility were checked against set validation rules at database level. For errors identified, the questionnaire was referred back to the field for revisit and correction. For data that could be corrected on consulting the data scientists were done internally.

Other Processing

To produce this micro-data set, the operational HDSS data is processed using Pentaho Kettle ETL program to produce this standard event-history format dataset.

Data Appraisal

Estimates of Sampling Error

Not applicable

File Description

Variable List

TZ021.3.v5

Content	Event History Micro Data Set
Cases	301311
Variable(s)	20
Structure	Type: Keys: ()
Version	3.v5
Producer	Magu HDSS
Missing Data	

Variables

ID	Name	Label	Type	Format	Question
V1	RecNr	RecNr	contin	numeric	
V2	CountryId	CountryId	discrete	numeric	
V3	CentreId	CentreId	discrete	character	
V4	IndividualId	IndividualId	contin	numeric	
V5	Sex	Sex	discrete	numeric	
V21	DoB	DoB	discrete	character	
V7	EventCount	EventCount	discrete	numeric	
V8	EventNr	EventNr	discrete	numeric	
V9	EventCode	EventCode	discrete	character	
V22	EventDate	EventDate	discrete	character	
V23	ObservationDate	ObservationDate	discrete	character	
V12	LocationId	LocationId	contin	numeric	
V13	MotherId	MotherId	contin	numeric	
V14	DeliveryId	DeliveryId	discrete	numeric	
V15	Cause1	Cause1	discrete	character	
V16	Cause2	Cause2	discrete	character	
V17	Cause3	Cause3	discrete	character	
V18	Likelihood1	Likelihood1	contin	numeric	
V19	Likelihood2	Likelihood2	contin	numeric	
V20	Likelihood3	Likelihood3	discrete	numeric	

RecNr (RecNr)

File: TZ021.3.v5

Overview

Type: Continuous	Valid cases: 301311
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 1
Range: 1-301311	Maximum: 301311
	Mean: 150656
	Standard deviation: 86981.1

Description

A sequential number uniquely identifying each record in the data file

CountryId (CountryId)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: numeric	Invalid: 0
Decimals: 0	
Range: 834-834	

Description

ISO 3166-1 numeric code of the country in which the surveillance site is situated

CentreId (CentreId)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: character	Invalid: 0
Width: 5	

Description

An identifier issued by INDEPTH to each member centre of the format CCCSS, where CCC is a sequential centre identifier and SS is a sequential identifier of the site within the centre in the case of multiple site centres

IndividualId (IndividualId)

File: TZ021.3.v5

Overview

Type: Continuous	Valid cases: 301311
Format: numeric	Invalid: 0
Decimals: 0	Minimum: 1
Range: 1-106723	Maximum: 106723
	Mean: 52802.9
	Standard deviation: 29992.7

Description

A number uniquely identifying all the records belonging to a specific individual in the data file. This number is not be the same as the identifier used by a contributing centre to identify the individual.

Sex (Sex)

File: TZ021.3.v5

Sex (Sex)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: numeric	Invalid: 0
Decimals: 0	
Range: 1-2	

Description

Sex of the individual. 1 for Male and 2 for Female

DoB (DoB)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: character	
Width: 23	

Description

The date of birth of the individual. Format: YYYY/MM/DD

EventCount (EventCount)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: numeric	Invalid: 0
Decimals: 0	
Range: 2-12	

Description

The total number of events associated with this individual in this data set

EventNr (EventNr)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: numeric	Invalid: 0
Decimals: 0	
Range: 1-12	

Description

A number increasing from 1 to EventCount for each event record in order of event occurrence

EventCode (EventCode)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: character	Invalid: 0
Width: 3	

Description

A code identifying the type of event that has occurred.

EventDate (EventDate)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: character	
Width: 23	

Description

The date on which the event occurred. Format: YYYY/MM/DD

ObservationDate (ObservationDate)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 301311
Format: character	
Width: 23	

Description

Date on which the event was observed (recorded), also known as surveillance visit date. Format: YYYY/MM/DD

LocationId (LocationId)

File: TZ021.3.v5

Overview

Type: Continuous	Valid cases: 236710
Format: numeric	Invalid: 64601
Decimals: 0	Minimum: 1
Range: 1-14318	Maximum: 14318
	Mean: 5781
	Standard deviation: 3720.3

Description

Unique identifier associated with a residential unit within the site and is the location where the individual was or became resident when the event occurred. This identifier is not be the same as the identifier used internally by the contributing centre.

MotherId (MotherId)

File: TZ021.3.v5

Overview

Type: Continuous	Valid cases: 3996
Format: numeric	Invalid: 297315
Decimals: 0	Minimum: 10
Range: 10-106701	Maximum: 106701
	Mean: 29090
	Standard deviation: 22409.4

Description

The IndividualId of the mother. Only provided for BTH events.

DeliveryId (DeliveryId)

File: TZ021.3.v5

Overview

DeliveryId (DeliveryId)

File: TZ021.3.v5

Type: Discrete
Format: numeric
Decimals: 0

Valid cases: 0
Invalid: 301311

Description

The RecNr of the delivery event associated with this birth

Cause1 (Cause1)

File: TZ021.3.v5

Overview

Type: Discrete
Format: character
Width: 40

Valid cases: 1051
Invalid: 0

Description

Cause of death as denoted by the InterVA4 program

Cause2 (Cause2)

File: TZ021.3.v5

Overview

Type: Discrete
Format: character
Width: 39

Valid cases: 82
Invalid: 0

Description

Cause of death as denoted by the InterVA4 program

Cause3 (Cause3)

File: TZ021.3.v5

Overview

Type: Discrete
Format: character
Width: 39

Valid cases: 5
Invalid: 0

Description

Cause of death as denoted by the InterVA4 program

Likelihood1 (Likelihood1)

File: TZ021.3.v5

Overview

Type: Continuous
Format: numeric
Decimals: 0
Range: 40-100

Valid cases: 1051
Invalid: 300260
Minimum: 40
Maximum: 100
Mean: 92.6
Standard deviation: 13.9

Description

The likelihood of Cause1 being the actual cause of death as calculated by the InterVA4 program.

Likelihood2 (Likelihood2)

File: TZ021.3.v5

Overview

Type: Continuous	Valid cases: 82
Format: numeric	Invalid: 301229
Decimals: 0	Minimum: 24
Range: 24-49	Maximum: 49
	Mean: 36.4
	Standard deviation: 5.3

Description

The likelihood of Cause2 being the actual cause of death as calculated by the InterVA4 program.

Likelihood3 (Likelihood3)

File: TZ021.3.v5

Overview

Type: Discrete	Valid cases: 5
Format: numeric	Invalid: 301306
Decimals: 0	
Range: 20-28	

Description

The likelihood of Cause3 being the actual cause of death as calculated by the InterVA4 program.

