South Africa

Sam Clark, The Ohio State University
Clara Calvert, LSHTM
Milly Marston, LSHTM
Yue Chu, The Ohio State University

Verbal Autopsy:Reimagining data and automated cause assignment (using ALPHA network data)

Study Documentation

June 8, 2022

Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Production Date	June 21, 2022
Version	v01
Identification	NIH.Verbal.Autopsy.Dataset.v1

Table of Contents

Verbal Autopsy:Reimagining data and automated cause assignment (using ALPHA network data) (ALPHA HIM - U 2000-2017)

Overview

Type	Health and Demographic Surveillance System (HDSS)
Identification	AHRI.NIH.Verbal.Autopsy.Dataset.v1
Version	Version 1.0.0
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.

Abstract

The overall aim of the grant is to develop affordable, robust, calibrated methods to assign causes to individual deaths and characterize the distribution of deaths by cause from verbal autopsy (VA) data. As part of ALPHA, all partners are being asked to provide their VA data (all ages) alongside any physician review data available, residency data and HIV data for secondary analysis to address the following objectives:

(i) Preparing VA data in a standardised format, and comparing sex- and age- specific cause-specific mortality fractions across different ALPHA partner study areas
(ii) Assessing the performance of automated methods for assigning causes of death based on verbal autopsy data, against physician review or HIV sero-survey data (where each of these data sources are available)

Once these secondary data have been shared with the ALPHA team, we will check the data for consistency and integrity and clean them with the help of the owners of the data and merge the clean data to form a pooled data set. In collaboration with the team at The Ohio State University, we will prepare distributions of deaths by causes using automated methods, and calculate the cause-specific mortality fractions and rates stratified by key socio-demographic characteristics of the population and look at how they change over time. The performance of automated methods will be assessed in comparison to physician review assignment of cause of death or HIV sero-survey data. These results and outputs will be shared with each partner, with opportunity to provide input, and be involved with writing any papers.

Kind of Data	Surveillance Data
Unit of Analysis	Each observation describes either an individual or an individual's period of residence in a household.

Scope & Coverage

Keywords	Mortality; Verbal Autopsy; InSilicoVA; cause of death; ALPHA; ALPHA network
Topics	Demography; Epidemiology; HIV; Mortality; Verbal Autopsy
Time Period(s)	2000-2020
Countries	South Africa

Geographic Coverage
South Africa

Universe
Household members of the Africa Health Research Institute Population Intervention Programme Study Area

Producers & Sponsors	
Primary Investigator(s)	Sam Clark, The Ohio State University Clara Calvert, LSHTM Milly Marston, LSHTM Yue Chu, The Ohio State University
Other Producer(s)	Africa Health Research Institute (AHRI)
Funding Agency/ies	National Institutes of Health (NIH)

Sampling
Sampling Procedure Household demographic Surveillance

Data Collection	
Data Collection Dates	start 2000-01-01 end 2020-12-31
Data Collection Mode	Face-to-face [f2f]
Data Collection Notes ALPHA data are harmonised datasets from ALPHA member sites and this dataset is from the Africa Health Research Institute.	
Questionnaires The original unharmonised data was collected by the Africa Health Research Institute	
Data Collector(s)	Africa Health Research Institute

Accessibility	
Access Authority	DataFirst (University of Cape Town) , support@data1st.org
Contact(s)	DataFirst helpdesk (University of Cape Town) , www.support.data1st.org , support@data1st.org
Access Conditions Licensed access data, available under conditions	
Citation Requirements Clark, S. et al. (2022) "Verbal Autopsy:Reimagining data and automated cause assignment (using ALPHA network data)." Africa Health Research Institute. doi: 10.23664/AHRI.NIH.VERBAL.AUTOPSY.DATASET.V1	

Files Description

Dataset contains 3 file(s)

AN02-02 Dataset 6.2b	
# Cases	224692
# Variable(s)	9

HIS01-01 HospitalDeaths	
# Cases	5753
# Variable(s)	51

VA01-01 PhyscianCoding Cause of Death	
# Cases	40442
# Variable(s)	28

Variables List

Dataset contains 88 variable(s)

File AN02-02 Dataset 6.2b

#	Name	Label	Type	Format	Valid	Invalid	Question
1	idno	Person ID number	continuous	numeric-12.0	224692	0	-
2	study_name	Name of study field site	discrete	character-12	224692	0	-
3	hiv_test ..	date of HIV test	discrete	character-11	202186	-	-
4	hiv_test ..	HIV test result	discrete	numeric-12.0	224692	0	-
5	informed ..	whether or not the participant was informed of test result	discrete	numeric-12.0	224692	0	-
6	source_o ..	Where the test information comes from	discrete	numeric-12.0	224692	0	-
7	test_ass ..	Test assumption	discrete	numeric-12.0	224692	0	-
8	original ..	Original hiv test result	discrete	numeric-12.0	224692	0	-
9	survey_r ..	Round name	discrete	numeric-12.0	201198	23494	-

File HIS01-01 HospitalDeaths

#	Name	Label	Type	Format	Valid	Invalid	Question
1	PatientId	Unique patient identifier	continuous	numeric-12.0	5753	0	-
2	Sex	Patient gender	discrete	numeric-12.0	5753	0	-
3	DoB	Date of birth	continuous	numeric-11.0	5753	0	-
4	LocalArea	Local area of residence	discrete	character-20	5753	0	-
5	Isigodi	Isigodi of residence	discrete	character-20	5751	0	-
6	NearestC ..	Nearest to place of residence	discrete	numeric-12.0	5753	0	-
7	Admissio ..	Unique admission identifier	continuous	numeric-12.0	5753	0	-
8	Admissio ..	Date of admission	continuous	numeric-11.0	5753	0	-
9	AgeAtAdm ..	Age in completed yrs on admission	discrete	numeric-12.0	0	5753	-
10	AgeDays	Age in days, if age=0	discrete	numeric-12.0	0	5753	-
11	Admissio ..	Admission ward	discrete	numeric-12.0	5753	0	-
12	Admissio ..	Admitting doctor code	continuous	numeric-12.0	5752	1	-
13	Discharg ..	Date of discharge	continuous	numeric-11.0	5753	0	-
14	Discharg ..	Status on discharge	discrete	numeric-12.0	5753	0	-
15	Discharg ..	Discharge doctor code	continuous	numeric-12.0	5743	10	-
16	LOS	Length of Stay (days)	continuous	numeric-12.0	5753	0	-
17	LOSH	Length of Stay (hours), if LOS=0	continuous	numeric-12.0	703	5050	-
18	HIVStatus	HIV Status on admission	discrete	numeric-12.0	5753	0	-
19	OnART	On ART prior to admission	discrete	numeric-12.0	5753	0	-
20	ARTDateS ..	Date on which ART was started	continuous	numeric-11.0	1893	3860	-

File HIS01-01 HospitalDeaths

#	Name	Label	Type	Format	Valid	Invalid	Question
21	HIVTest	Latest HIV test in hospital result	discrete	numeric-12.0	5753	0	-
22	LatestCD4	Latest in hospital CD4 count	continuous	numeric-12.0	170	5583	-
23	CD4Date	Date on which CD4 was done	continuous	numeric-11.0	170	5583	-
24	ADxType1	Admission diagnosis 1 type	discrete	numeric-12.0	5751	2	-
25	ADx1	Admission diagnosis 1 ICD10 code	discrete	character-6	5751	0	-
26	ADxType2	Admission diagnosis 2 type	discrete	numeric-12.0	4490	1263	-
27	ADx2	Admission diagnosis 2 ICD10 code	discrete	character-6	4490	0	-
28	ADxType3	Admission diagnosis 3 type	discrete	numeric-12.0	2068	3685	-
29	ADx3	Admission diagnosis 3 ICD10 code	discrete	character-6	2068	0	-
30	ADxType4	Admission diagnosis 4 type	discrete	numeric-12.0	507	5246	-
31	ADx4	Admission diagnosis 4 ICD10 code	discrete	character-6	507	0	-
32	ADxType5	Admission diagnosis 5 type	discrete	numeric-12.0	102	5651	-
33	ADx5	Admission diagnosis 5 ICD10 code	discrete	character-6	102	0	-
34	ADxType6	Admission diagnosis 6 type	discrete	numeric-12.0	11	5742	-
35	ADx6	Admission diagnosis 6 ICD10 code	discrete	character-5	11	0	-
36	DDxType1	Discharge diagnosis 1 type	discrete	numeric-12.0	5748	5	-
37	DDx1	Discharge diagnosis 1 ICD10 code	discrete	character-6	5748	0	-
38	DDxType2	Discharge diagnosis 2 type	discrete	numeric-12.0	4870	883	-
39	DDx2	Discharge diagnosis 2 ICD10 code	discrete	character-6	4870	0	-
40	DDxType3	Discharge diagnosis 3 type	discrete	numeric-12.0	2527	3226	-
41	DDx3	Discharge diagnosis 3 ICD10 code	discrete	character-6	2527	0	-
42	DDxType4	Discharge diagnosis 4 type	discrete	numeric-12.0	762	4991	-
43	DDx4	Discharge diagnosis 4 ICD10 code	discrete	character-6	762	0	-
44	DDxType5	Discharge diagnosis 5 type	discrete	numeric-12.0	160	5593	-
45	DDx5	Discharge diagnosis 5 ICD10 code	discrete	character-6	160	0	-
46	DDxType6	Discharge diagnosis 6 type	discrete	numeric-12.0	15	5738	-
47	DDx6	Discharge diagnosis 6 ICD10 code	discrete	character-5	15	0	-
48	ACDISIIn ..	Surveillance individual identifier	continuous	numeric-12.0	446	5307	-
49	ARTemisP ..	ARTemis individual identifier	continuous	numeric-12.0	1144	4609	-
50	TheatreP ..	Theatre Procedure Code	discrete	character-5	4	0	-

File HIS01-01 HospitalDeaths

#	Name	Label	Type	Format	Valid	Invalid	Question
51	Anaesthe ..	Anaesthetic Code	discrete	numeric-12.0	4	5749	-

File VA01-01 PhyscianCoding Cause of Death

#	Name	Label	Type	Format	Valid	Invalid	Question
1	QSessionID	VA Unique Internal Id	continuous	numeric-12.0	40442	0	-
2	IIntId	Unique Internal Id of Individual	continuous	numeric-12.0	40442	0	-
3	DateOfDe ..	Individual's Date of Death	discrete	character-11	40442	-	-
4	DateOfBi ..	Individual's Date of Birth	discrete	character-11	40442	-	-
5	Sex	Sex	discrete	numeric-12.0	40442	0	-
6	VDX_Unde ..	Underlying cause of death	discrete	character-90	39825	0	-
7	VDX_Imme ..	Immediate cause of death	discrete	character-62	14383	0	-
8	VDX_Asso ..	Contributory cause of death 1	discrete	character-50	6606	0	-
9	VDX_Asso ..	Contributory cause of death 2	discrete	character-51	1614	0	-
10	VDX_Doctor	Doctor	discrete	character-3	40295	0	-
11	VDX_DxDate	Date of diagnosis	discrete	character-11	40442	-	-
12	Immediat ..	Immediate cause of death	discrete	character-5	12926	0	-
13	Immediat ..	Immediate Cause Descriptin	discrete	character-185	12926	-	-
14	Immediat ..	Immediate Cause GBD code	discrete	character-4	12926	0	-
15	Immediat ..	Immediate Cause GBD Descriptin	discrete	character-42	12926	0	-
16	Underlyi ..	Underlying cause of death	discrete	character-5	40278	0	-
17	Underlyi ..	Underlying Cause Descriptin	discrete	character-185	40278	-	-
18	Underlyi ..	Underlying Cause GBD code	discrete	character-4	40278	0	-
19	Underlyi ..	Underlying Cause GBD Desc	discrete	character-42	40278	0	-
20	Associat ..	Associated cause of death 1	discrete	character-5	5542	0	-
21	Associat ..	Associated Cause 1 Descriptin	discrete	character-102	5542	-	-
22	Associat ..	Associated Cause 1 Cause GBD code	discrete	character-4	5542	0	-
23	Associat ..	Associated Cause 1 GBD Descriptin	discrete	character-40	5542	0	-
24	Associat ..	Associated cause of death 2	discrete	character-5	1284	0	-
25	Associat ..	Descriptive label for code above	discrete	character-72	1284	0	-
26	Associat ..	Associated Cause 2 Cause GBD code	discrete	character-4	1284	0	-
27	Associat ..	Associated Cause 2 GBD Descriptin	discrete	character-32	1284	0	-
28	DeathCau ..	Summary GBD Cause	discrete	character-8	40442	0	-

Variables Description

Dataset contains 88 variable(s)

File : AN02-02 Dataset 6.2b

idno: Person ID number

Information	[Type= continuous] [Format=numeric] [Range= 16-258195] [Missing=*]
Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-] [Mean=81852.305 /-] [StdDev=60715.191 /-]

study_name: Name of study field site

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-]

hiv_test_date: date of HIV test

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=202186 /-]

hiv_test_result: HIV test result

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-]

Value	**Label**	**Cases**	**Percentage**
0	negative	148199	66.0%
1	positive	76172	33.9%
2	indeterminate	321	0.1%
3	not reported	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.

informed_of_result: whether or not the participant was informed of test result

Information	[Type= discrete] [Format=numeric] [Range= 0-8] [Missing=*]
Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-]

Value	**Label**	**Cases**	**Percentage**
0	no	201198	89.5%
1	yes	23494	10.5%
8	Don?t know/not asked	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.

source_of_test_information: Where the test information comes from

Information	[Type= discrete] [Format=numeric] [Range= 1-8] [Missing=*]
Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-]

Value	**Label**	**Cases**	**Percentage**
1	part of a population-based study	201198	89.5%
2	part of a special research study	0	
3	clinical record- HIV clinic	23494	10.5%
4	self reported by respondent	0	
5	report by proxy respondent at VA	0	
6	clinical record- walk in VCT	0	
7	clinical record- PMTCT/ANC	0	
8	clinical record- other	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.

test_assumption: Test assumption

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]

File : AN02-02 Dataset 6.2b

test_assumption: Test assumption

Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage	
0	New test	224692		100.0%
1	Test from previous study used	0		
2	Previous positive HIV test used	0		
3	self reported instead of testing (Kisumu)	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.

original_hiv_test_result: Original hiv test result

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=224692 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage	
0	negative	148199		66.0%
1	positive	76172		33.9%
2	indeterminate	321	0.1%	
3	not reported	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.

survey_round_name: Round name

Information	[Type= discrete] [Format=numeric] [Range= 0-11] [Missing=*]
Statistics [NW/ W]	[Valid=201198 /-] [Invalid=23494 /-]
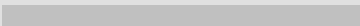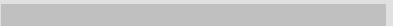
File : HIS01-01 HospitalDeaths

PatientId: Unique patient identifier

Information	[Type= continuous] [Format=numeric] [Range= 35-108138] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-] [Mean=36252.952 /-] [StdDev=32122.654 /-]

Sex: Patient gender

Information	[Type= discrete] [Format=numeric] [Range= 0-9] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Male	2775	48.2%
1	Female	2978	51.8%
9	Unknown	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.

DoB: Date of birth

Information	[Type= continuous] [Format=numeric] [Range= -1805069420000-1951344000000] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-] [Mean=316491009382.931 /-] [StdDev=733586226916.454 /-]

LocalArea: Local area of residence
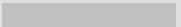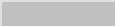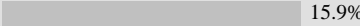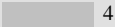
Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Isigodi: Isigodi of residence

| Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5751 /-] [Invalid=0 /-]

NearestClinic: Nearest to place of residence

Information	[Type= discrete] [Format=numeric] [Range= 1-22] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Kwamsane	533	9.3%
2	Somkhele	390	6.8%
3	Mtuba	917	15.9%
4	Sipho Zungu	282	4.9%
5	Zwenelisha	84	1.5%
6	Mpukunyoni	268	4.7%
7	Nkundusi	388	6.7%
8	Madwaleni	196	3.4%
9	Ntondweni	43	0.7%
10	Gunjaneni	117	2.0%
11	Macabuzela	191	3.3%
12	Makhowe	173	3.0%
13	Inhlwathi	153	2.7%
14	Mpembeni	204	3.5%
15	Philanjalo	0	
16	Gateway	1180	20.5%
17	Esiyembeni	25	0.4%
18	Machibini	136	2.4%

File : HIS01-01 HospitalDeaths

NearestClinic: Nearest to place of residence

Value	Label	Cases	Percentage
19	Unknown	207	3.6%
20	Cingci	18	0.3%
21	Baksdina	164	2.9%
22		84	1.5%

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.

AdmissionId: Unique admission identifier

Information	[Type= continuous] [Format=numeric] [Range= 36-123033] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-] [Mean=44717.25 /-] [StdDev=37841.626 /-]

AdmissionDate: Date of admission

Information	[Type= continuous] [Format=numeric] [Range= 1600010100000-1953082229000] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-] [Mean=1755257318324.56 /-] [StdDev=99876334388.546 /-]

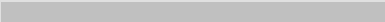## # AgeAtAdmission: Age in completed yrs on admission
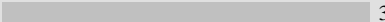
Information	[Type= discrete] [Format=numeric] [Missing=*]
Statistics [NW/ W]	[Valid=0 /-] [Invalid=5753 /-]

AgeDays: Age in days, if age=0

Information	[Type= discrete] [Format=numeric] [Missing=*]
Statistics [NW/ W]	[Valid=0 /-] [Invalid=5753 /-]

AdmissionWard: Admission ward

Information	[Type= discrete] [Format=numeric] [Range= 1-7020] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	C Female Medical	2118	36.8%
2	H Male Surgical	239	4.2%
3	High Care Unit	198	3.4%
4	Maternity / ANC	2	0.0%
7	Post-natal C/S	0	
8	Post-NVD	3	0.1%
9	TB male	110	1.9%
10	B Paediatric Medical	368	6.4%
11	S Female Surgical	326	5.7%
12	TB Female	90	1.6%
13	H Male Medical	2026	35.2%
14	B Paediatric Surgical	7	0.1%
15	Labour Ward	0	
17	Nursery Ward	186	3.2%
18		9	0.2%
19		19	0.3%
3019		14	0.2%
3020		37	0.6%
7020		1	0.0%

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.

File : HIS01-01 HospitalDeaths

AdmissionDoctor: Admitting doctor code

Information	[Type= continuous] [Format=numeric] [Range= 1-14123] [Missing=*]
Statistics [NW/ W]	[Valid=5752 /-] [Invalid=1 /-] [Mean=842.866 /-] [StdDev=2346.586 /-]

DischargeDate: Date of discharge

Information	[Type= continuous] [Format=numeric] [Range= 1602684231000-1953108000000] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-] [Mean=1755986204242.94 /-] [StdDev=99849653317.318 /-]

DischargeStatus: Status on discharge

Information	[Type= discrete] [Format=numeric] [Range= 0-9] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Current Inpatient	0	
1	Discharged	0	
2	Transferred	0	
3	Died	5753	100.0%
4	Absconded	0	
9	Unknown	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.

DischargeDoctor: Discharge doctor code

Information	[Type= continuous] [Format=numeric] [Range= 1-14123] [Missing=*]
Statistics [NW/ W]	[Valid=5743 /-] [Invalid=10 /-] [Mean=543.729 /-] [StdDev=1494.859 /-]

LOS: Length of Stay (days)

Information	[Type= continuous] [Format=numeric] [Range= 0-367] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-] [Mean=7.948 /-] [StdDev=12.246 /-]

LOSH: Length of Stay (hours), if LOS=0

Information	[Type= continuous] [Format=numeric] [Range= 0-23] [Missing=*]
Statistics [NW/ W]	[Valid=703 /-] [Invalid=5050 /-] [Mean=12.307 /-] [StdDev=7.106 /-]

HIVStatus: HIV Status on admission

Information	[Type= discrete] [Format=numeric] [Range= 0-9] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Undefined	11	0.2%
1	Positive	2628	45.7%
2	Negative	323	5.6%
3	Not recorded	2128	37.0%
4	Child unexposed	113	2.0%
5	Awaiting PCR result	39	0.7%
9	Unknown	511	8.9%

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.

OnART: On ART prior to admission

Information	[Type= discrete] [Format=numeric] [Range= 0-9] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

File : HIS01-01 HospitalDeaths

OnART: On ART prior to admission

Value	Label	Cases	Percentage
0	Undefined	69	1.2%
1	Yes	1881	32.7%
2	No	2297	39.9%
9	Unknown	1506	26.2%

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.

ARTDateStarted: Date on which ART was started

Information	[Type= continuous] [Format=numeric] [Range= -5396112000000-1938124800000] [Missing=*/10000000000001]
Statistics [NW/ W]	[Valid=1893 /-] [Invalid=3860 /-] [Mean=-1642040808874.8 /-] [StdDev=1429772861453.64 /-]

Value	Label	Cases	Percentage
10000000000001	..		

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.

HIVTest: Latest HIV test in hospital result

Information	[Type= discrete] [Format=numeric] [Range= 0-3] [Missing=*]
Statistics [NW/ W]	[Valid=5753 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
0	Not done	5480	95.3%
1	Positive	187	3.3%
2	Negative	81	1.4%
3	PCR Taken	5	0.1%

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.

LatestCD4: Latest in hospital CD4 count

Information	[Type= continuous] [Format=numeric] [Range= 3-1178] [Missing=*]
Statistics [NW/ W]	[Valid=170 /-] [Invalid=5583 /-] [Mean=156.782 /-] [StdDev=188.194 /-]

CD4Date: Date on which CD4 was done

Information	[Type= continuous] [Format=numeric] [Range= -2556057600000-2029536000000] [Missing=*/10000000000001]
Statistics [NW/ W]	[Valid=170 /-] [Invalid=5583 /-] [Mean=-3741628235.294 /-] [StdDev=2065226536692.56 /-]

Value	Label	Cases	Percentage
10000000000001	..		

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.

ADxType1: Admission diagnosis 1 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=5751 /-] [Invalid=2 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	5613	97.6%
2	Underlying Cause	26	0.5%
3	Associated Cause	9	0.2%
9	Symptom/Sign	103	1.8%
Sysmiss		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.

File : HIS01-01 HospitalDeaths

ADx1: Admission diagnosis 1 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5751 /-] [Invalid=0 /-]

ADxType2: Admission diagnosis 2 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=4490 /-] [Invalid=1263 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	67	1.5%
2	Underlying Cause	1992	44.4%
3	Associated Cause	2379	53.0%
9	Symptom/Sign	52	1.2%
Sysmiss		1263	

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.

ADx2: Admission diagnosis 2 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=4490 /-] [Invalid=0 /-]

ADxType3: Admission diagnosis 3 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=2068 /-] [Invalid=3685 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	6	0.3%
2	Underlying Cause	30	1.5%
3	Associated Cause	1984	95.9%
9	Symptom/Sign	48	2.3%
Sysmiss		3685	

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.

ADx3: Admission diagnosis 3 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=2068 /-] [Invalid=0 /-]

ADxType4: Admission diagnosis 4 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=507 /-] [Invalid=5246 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	1	0.2%
2	Underlying Cause	1	0.2%
3	Associated Cause	495	97.6%
9	Symptom/Sign	10	2.0%
Sysmiss		5246	

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.

ADx4: Admission diagnosis 4 ICD10 code

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

File : HIS01-01 HospitalDeaths

ADx4: Admission diagnosis 4 ICD10 code

Statistics [NW/ W]	[Valid=507 /-] [Invalid=0 /-]

ADxType5: Admission diagnosis 5 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=102 /-] [Invalid=5651 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	0	
2	Underlying Cause	0	
3	Associated Cause	98	96.1%
9	Symptom/Sign	4	3.9%
Sysmiss		5651	

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.

ADx5: Admission diagnosis 5 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=102 /-] [Invalid=0 /-]

ADxType6: Admission diagnosis 6 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=11 /-] [Invalid=5742 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	0	
2	Underlying Cause	0	
3	Associated Cause	10	90.9%
9	Symptom/Sign	1	9.1%
Sysmiss		5742	

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.

ADx6: Admission diagnosis 6 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=11 /-] [Invalid=0 /-]

DDxType1: Discharge diagnosis 1 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=5748 /-] [Invalid=5 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	5626	97.9%
2	Underlying Cause	29	0.5%
3	Associated Cause	9	0.2%
9	Symptom/Sign	84	1.5%
Sysmiss		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.

DDx1: Discharge diagnosis 1 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5748 /-] [Invalid=0 /-]

File : HIS01-01 HospitalDeaths

DDxType2: Discharge diagnosis 2 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=4870 /-] [Invalid=883 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	104	2.1%
2	Underlying Cause	2647	54.4%
3	Associated Cause	2093	43.0%
9	Symptom/Sign	26	0.5%
Sysmiss		883	

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.

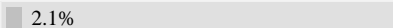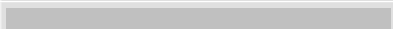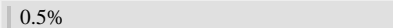## # DDx2: Discharge diagnosis 2 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=4870 /-] [Invalid=0 /-]

DDxType3: Discharge diagnosis 3 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=2527 /-] [Invalid=3226 /-]

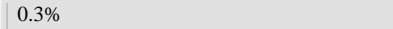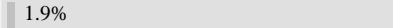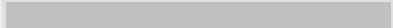| Value	Label	Cases	Percentage
1	Immediate Cause	7	0.3%
2	Underlying Cause	49	1.9%
3	Associated Cause	2446	96.8%
9	Symptom/Sign	25	1.0%
Sysmiss		3226	

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.

DDx3: Discharge diagnosis 3 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=2527 /-] [Invalid=0 /-]

DDxType4: Discharge diagnosis 4 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=762 /-] [Invalid=4991 /-]

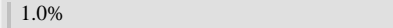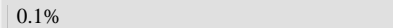| Value	Label	Cases	Percentage
1	Immediate Cause	1	0.1%
2	Underlying Cause	3	0.4%
3	Associated Cause	747	98.0%
9	Symptom/Sign	11	1.4%
Sysmiss		4991	

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.

DDx4: Discharge diagnosis 4 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=762 /-] [Invalid=0 /-]

DDxType5: Discharge diagnosis 5 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]

File : HIS01-01 HospitalDeaths

DDxType5: Discharge diagnosis 5 type

Statistics [NW/ W]	[Valid=160 /-] [Invalid=5593 /-]

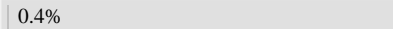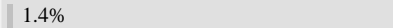| Value	Label	Cases	Percentage
1	Immediate Cause	0	
2	Underlying Cause	0	
3	Associated Cause	154	96.2%
9	Symptom/Sign	6	3.8%
Sysmiss		5593	

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.

DDx5: Discharge diagnosis 5 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=160 /-] [Invalid=0 /-]

DDxType6: Discharge diagnosis 6 type

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=15 /-] [Invalid=5738 /-]

Value	Label	Cases	Percentage
1	Immediate Cause	0	
2	Underlying Cause	0	
3	Associated Cause	14	93.3%
9	Symptom/Sign	1	6.7%
Sysmiss		5738	

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.

DDx6: Discharge diagnosis 6 ICD10 code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=15 /-] [Invalid=0 /-]

ACDISIIntId: Surveillance individual identifier

Information	[Type= continuous] [Format=numeric] [Range= 149-172512] [Missing=*]
Statistics [NW/ W]	[Valid=446 /-] [Invalid=5307 /-] [Mean=70121.504 /-] [StdDev=46772.652 /-]

ARTemisPersonNo: ARTemis individual identifier

Information	[Type= continuous] [Format=numeric] [Range= 154-211127] [Missing=*]
Statistics [NW/ W]	[Valid=1144 /-] [Invalid=4609 /-] [Mean=97125.941 /-] [StdDev=63174.562 /-]

TheatreProcedureCode: Theatre Procedure Code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=4 /-] [Invalid=0 /-]

# **AnaestheticCode: Anaesthetic Code**				
Information	[Type= discrete] [Format=numeric] [Range= 0-9] [Missing=*]			
Statistics [NW/ W]	[Valid=4 /-] [Invalid=5749 /-]			

Value	Label	Cases	Percentage	
0	Undefined	1		25.0%
1	General	1		25.0%
2	Spinal block	0		
3	Nerve block	0		
4	Local	1		25.0%
9	Unknown	1		25.0%
Sysmiss		5749		

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.

File : VA01-01 PhyscianCoding Cause of Death

QSessionID: VA Unique Internal Id

Information	[Type= continuous] [Format=numeric] [Range= 20854-155668] [Missing=*]
Statistics [NW/ W]	[Valid=40442 /-] [Invalid=0 /-] [Mean=60450.121 /-] [StdDev=29969.001 /-]

IIntId: Unique Internal Id of Individual

Information	[Type= continuous] [Format=numeric] [Range= 11-176326] [Missing=*]
Statistics [NW/ W]	[Valid=40442 /-] [Invalid=0 /-] [Mean=58908.789 /-] [StdDev=37154.011 /-]

DateOfDeath: Individual's Date of Death

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40442 /-]

DateOfBirth: Individual's Date of Birth

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40442 /-]

Sex: Sex

Information	[Type= discrete] [Format=numeric] [Range= 1-9] [Missing=*]
Statistics [NW/ W]	[Valid=40442 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage	
1	Male	20362		50.3%
2	Female	20077		49.6%
9	Unknown	3	0.0%	

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.

VDX_UnderlyingDx: Underlying cause of death

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=39825 /-] [Invalid=0 /-]

VDX_ImmediateDx: Immediate cause of death

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=14383 /-] [Invalid=0 /-]

VDX_AssocDx1: Contributory cause of death 1

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=6606 /-] [Invalid=0 /-]

VDX_AssocDx2: Contributory cause of death 2

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=1614 /-] [Invalid=0 /-]

VDX_Doctor: Doctor

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40295 /-] [Invalid=0 /-]

VDX_DxDate: Date of diagnosis

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40442 /-]

File : VA01-01 PhyscianCoding Cause of Death

ImmediateCauseCode: Immediate cause of death

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=12926 /-] [Invalid=0 /-]

ImmediateCauseDesc: Immediate Cause Descriptin

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=12926 /-]

ImmediateCauseGBDCode: Immediate Cause GBD code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=12926 /-] [Invalid=0 /-]

ImmediateCauseGBDDesc: Immediate Cause GBD Descriptin

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=12926 /-] [Invalid=0 /-]

UnderlyingCauseCode: Underlying cause of death

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40278 /-] [Invalid=0 /-]

UnderlyingCauseDesc: Underlying Cause Descriptin

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40278 /-]

UnderlyingCauseGBDCode: Underlying Cause GBD code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40278 /-] [Invalid=0 /-]

UnderlyingCauseGBDDesc: Underlying Cause GBD Desc

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=40278 /-] [Invalid=0 /-]

AssociatedCause1Code: Associated cause of death 1

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5542 /-] [Invalid=0 /-]

AssociatedCause1Desc: Associated Cause 1 Descriptin

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5542 /-]

AssociatedCause1GBDCode: Associated Cause 1 Cause GBD code

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5542 /-] [Invalid=0 /-]

AssociatedCause1GBDDesc: Associated Cause 1 GBD Descriptin

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=5542 /-] [Invalid=0 /-]

AssociatedCause2Code: Associated cause of death 2

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

File : VA01-01 PhyscianCoding Cause of Death

AssociatedCause2Code: Associated cause of death 2

| Statistics [NW/ W] | [Valid=1284 /-] [Invalid=0 /-] |

AssociatedCause2Desc: Descriptive label for code above

| Information | [Type= discrete] [Format=character] [Missing=*] |
| Statistics [NW/ W] | [Valid=1284 /-] [Invalid=0 /-] |

AssociatedCause2GBDCode: Associated Cause 2 Cause GBD code

| Information | [Type= discrete] [Format=character] [Missing=*] |
| Statistics [NW/ W] | [Valid=1284 /-] [Invalid=0 /-] |

AssociatedCause2GBDDesc: Associated Cause 2 GBD Descriptin

| Information | [Type= discrete] [Format=character] [Missing=*] |
| Statistics [NW/ W] | [Valid=1284 /-] [Invalid=0 /-] |

DeathCauseGBDgroup: Summary GBD Cause

| Information | [Type= discrete] [Format=character] [Missing=*] |
| Statistics [NW/ W] | [Valid=40442 /-] [Invalid=0 /-] |