

Africa, Asia, Oceania

Osman A. Sankoh, INDEPTH Network
Peter Byass, WHO Collaborating Centre for Verbal Autopsy, Umeå University
P. Kim Streatfield, Matlab, Bangladesh
Wasif A. Khan, Bandarban, Bangladesh
Abba Bhuiya, Chakaria, Bangladesh
Nurul Alam, AMK, Bangladesh
Ali Sie, Nouna, Burkina Faso
Abdramane Soura, Ouagadougou, Burkina Faso
Bassirou Bonfoh, Taabo, Côte d'Ivoire
Berhe Weldearegawi, Kilite-Awlaelo, Ethiopia
Abraham Oduro, Navrongo, Ghana
Margaret Gyapong, Dodowa, Ghana
Momodou Jasseh, Farafenni, The Gambia
Siswanto Wilopo, Purworejo, Indonesia
Shashi Kant, Ballabgarh, India
Sanjay Juvekar, Vadu, India
Thomas N. Williams, Kilifi, Kenya
Frank O. Odhiambo, Kisumu, Kenya
Alex Ezeh, Nairobi, Kenya
Amelia Crampin, Karonga, Malawi
Valérie Delaunay, Niakhar, Senegal
Stephen M. Tollman, Agincourt, South Africa
Abraham J. Herbst, Africa Centre, South Africa
Nguyen T.K. Chuc, FilaBavi, Vietnam
Marcel Tanner, Swiss Tropical and Public Health Institute

INDEPTH Network Cause-Specific Mortality - Release 2014

Study Documentation

October 28, 2014

Metadata Production

Metadata Producer(s)	Kobus Herbst (KHe) , INDEPTH Network , Study documentation Peter Byass (PBy) , INDEPTH Network , Study documentation
Production Date	October 13, 2014
Version	Version 3 (22 Sep 2014) Version 2 (15 Sep 2014) Version 1 (13 Aug 2014)
Identification	DDI.INDEPTH.GH003.COD2014.v2

Table of Contents

Overview	4
Scope & Coverage	5
Producers & Sponsors	5
Sampling	6
Data Collection	7
Data Processing & Appraisal	7
Accessibility	7
Rights & Disclaimer	8
Files Description	9
CODA_2013_v7_Anonymised	9
Variables List	10
CODA_2013_v7_Anonymised	10
Variables Description	11
CODA_2013_v7_Anonymised	12

INDEPTH Network Cause-Specific Mortality - Release 2014

Overview	
Type	Demographic Surveillance
Identification	INDEPTH.GH003.COD2014.v1
Version	Production Date: 2014-09-15 Version 1 Notes Version 1. Base Version
Series	INDEPTH Cause of Death

Abstract

Cause of death data based on VA interviews were contributed by fourteen INDEPTH HDSS sites in sub-Saharan Africa and eight sites in Asia. The principles of the Network and its constituent population surveillance sites have been described elsewhere [1]. Each HDSS site is committed to long-term longitudinal surveillance of circumscribed populations, typically each covering around 50,000 to 100,000 people. Households are registered and visited regularly by lay field-workers, with a frequency varying from once per year to several times per year. All vital events are registered at each such visit, and any deaths recorded are followed up with verbal autopsy interviews, usually 147 undertaken by specially trained lay interviewers. A few sites were already operational in the 1990s, but in this dataset 95% of the person-time observed related to the period from 2000 onwards, with 58% from 2007 onwards. Two sites, in Nairobi and Ouagadougou, followed urban populations, while the remainder covered areas that were generally more rural in character, although some included local urban centres. Sites covered entire populations, although the Karonga, Malawi, site only contributed VAs for deaths of people aged 12 years and older. Because the sites were not located or designed in a systematic way to be representative of national or regional populations, it is not meaningful to aggregate results over sites.

All cause of death assignments in this dataset were made using the InterVA-4 model version 4.02 [2]. InterVA-4 uses probabilistic modelling to arrive at likely cause(s) of death for each VA case, the workings of the model being based on a combination of expert medical opinion and relevant available data. InterVA-4 is the only model currently available that processes VA data according to the WHO 2012 standard and categorises causes of death according to ICD-10. Since the VA data reported here were collected before the WHO 2012 standard was formulated, they were all retrospectively transformed into the WHO 2012 and InterVA-4 input format for processing.

The InterVA-4 model was applied to the data from each site, yielding, for each case, up to three possible causes of death or an indeterminate result. Each cause for a case is a single record in the dataset. In a minority of cases, for example where symptoms were vague, contradictory or mutually inconsistent, it was impossible for InterVA-4 to determine a cause of death, and these deaths were attributed as entirely indeterminate. For the remaining cases, one to three likely causes and their likelihoods were assigned by InterVA-4, and if the sum of their likelihoods was less than one, the residual component was then assigned as being indeterminate. This was an important process for capturing uncertainty in cause of death outcome(s) from the model at the individual level, thus avoiding over-interpretation of specific causes. As a consequence there were three sources of unattributed cause of death: deaths registered for which VAs were not successfully completed; VAs completed but where the cause was entirely indeterminate; and residual components of deaths attributed as indeterminate.

In this dataset each case has between one and four records, each with its own cause and likelihood. Cases for which VAs were not successfully completed has a single record with the cause of death recorded as "VA not completed" and a likelihood of one. Thus the overall sum of the likelihoods equated to the total number of deaths. Each record also contains a population weighting factor reflecting the ratio of the population fraction for its site, age group, sex and year to the corresponding age group and sex fraction in the standard population (see section on weighting).

In this context, all of these data are secondary datasets derived from primary data collected separately by each participating site. In all cases the primary data collection was covered by site-level ethical approvals relating to on-going demographic surveillance in those specific locations. No individual identity or household location data are included in this secondary data.

1. Sankoh O, Byass P. The INDEPTH Network: filling vital gaps in global epidemiology. *International Journal of Epidemiology* 2012; 41:579-588.

2. Byass P, Chandramohan D, Clark SJ, D'Ambruso L, Fottrell E, Graham WJ, et al. Strengthening standardised interpretation of verbal autopsy data: the new InterVA-4 tool. *Global Health Action* 2012; 5:19281.

Kind of Data	Verbal autopsy-based cause of death data
Unit of Analysis	Death Cause

Scope & Coverage

Scope

Cause of death derived from verbal autopsy interviews using automated analysis by a computer program, InterVA 4.

Topics	Cause of Death [N01.224.935.698.100], Mortality [N01.224.935.698]
Time Period(s)	1992-2012
Countries	Africa, Asia, Oceania

Geographic Coverage

Demographic surveillance areas of the following HDSSs:

```

<pre>Code Country INDEPTH Centre
BD011 Bangladesh ICDDR-B : Matlab
BD012 Bangladesh ICDDR-B : Bandarban
BD013 Bangladesh ICDDR-B : Chakaria
BD014 Bangladesh ICDDR-B : AMK
BF031 Burkina Faso Nouna
BF041 Burkina Faso Ouagadougou
CI011 Côte d'Ivoire Taabo
ET031 Ethiopia Kilite Awlalelo
GH011 Ghana Navrongo
GH031 Ghana Dodowa
GM011 The Gambia Farafenni
ID011 Indonesia Purworejo
IN011 India Ballabgarh
IN021 India Vadu
KE011 Kenya Kilifi
KE021 Kenya Kisumu
KE031 Kenya Nairobi
MW011 Malawi Karonga
SN011 Senegal IRD : Bandafassi
VN012 Vietnam Hanoi Medical University : Filabavi
ZA011 South Africa Agincourt
ZA031 South Africa Africa Centre </pre>

```

Universe

Surveillance population
Deceased individuals
Cause of death

Producers & Sponsors

Primary Investigator(s)	Osman A. Sankoh, INDEPTH Network Peter Byass, WHO Collaborating Centre for Verbal Autopsy, Umeå University P. Kim Streatfield, Matlab, Bangladesh Wasif A. Khan, Bandarban, Bangladesh
--------------------------------	---

	<p>Abba Bhuiya, Chakaria, Bangladesh Nurul Alam, AMK, Bangladesh Ali Sie, Nouna, Burkina Faso Abdramane Soura, Ouagadougou, Burkina Faso Bassirou Bonfoh, Taabo, Côte d'Ivoire Berhe Weldearegawi, Kilite-Awlaelo, Ethiopia Abraham Oduro, Navrongo, Ghana Margaret Gyapong, Dodowa, Ghana Momodou Jasseh, Farafenni, The Gambia Siswanto Wilopo, Purworejo, Indonesia Shashi Kant, Ballabgarh, India Sanjay Juvekar, Vadu, India Thomas N. Williams, Kilifi, Kenya Frank O. Odhiambo, Kisumu, Kenya Alex Ezeh, Nairobi, Kenya Amelia Crampin, Karonga, Malawi Valérie Delaunay, Niakhar, Senegal Stephen M. Tollman, Agincourt, South Africa Abraham J. Herbst, Africa Centre, South Africa Nguyen T.K. Chuc, FilaBavi, Vietnam Marcel Tanner, Swiss Tropical and Public Health Institute</p>
Funding Agency/ies	<p>Bill & Melinda Gates Foundation IDRC Rockefeller Foundation SIDA / Research Cooperation Swiss TPH Wellcome Trust WHO/HMN William and Flora Hewlett Foundation</p>
Other Acknowledgment(s)	<p>Kobus Herbst , Dataset Production , INDEPTH Network Peter Byass , Dataset Production & Analysis , WHO Collaborating Centre for Verbal Autopsy, Umeå University Samuelina Arthur , Dataset Coordination , INDEPTH Network</p>

Sampling

Sampling Procedure

No sampling, covers total population in demographic surveillance area

Weighting

The number of deaths by sex and age group were weighted using the INDEPTH 2013 standard population structure for low- and middle-income countries (LMICs) in Africa and Asia [1], as shown in Table 1. This public-domain standard population has been presented in relation to other global standards such as Segi and WHO, from which it differs in reflecting the higher fertility and younger-age mortality rates commonly seen in LMIC populations [1].

Each record contains a population weighting factor (wt) reflecting the ratio of the population fraction for its site, age group, sex and year to the corresponding age group and sex fraction in the standard population described in Table 1, for the purposes of standardisation. A further factor (lik_wt) is calculated for each record as the product of the VA cause likelihood and the population standard weighting (both described above), which can be used as the basis for calculating age-sex-time standardised CSMFs and CSMRs.

Table 1:

<pre>Age Group	INDEPTH 2013 standard	
Male	Female	
0-28 days	0.11%	0.10%

1-11 months	1.49%	1.38%
1-4 years	6.01%	5.57%
5-14 years	13.76%	12.57%
15-49 years	22.54%	23.50%
50-64 years	3.87%	4.36%
65+ years	2.22%	2.52%

1. Sankoh O, Sharrow D, Herbst K, Kabudula CW, Alam N, Kant S, et al. The INDEPTH standard population for low- and middle-income countries, 2013. *Global Health Action* 2014; 7:23286.

Data Collection

Data Collection Dates	start 1992-01-01 end 2012-12-31
------------------------------	------------------------------------

Time Period(s)	start 1992-01-01 end 2012-12-31
-----------------------	------------------------------------

Data Collection Mode	Face-to-face [f2f]
-----------------------------	--------------------

Data Collection Notes

Households are registered and visited regularly by lay field-workers, with a frequency varying from once per year to several times per year. All vital events are registered at each such visit, and any deaths recorded are followed up with verbal autopsy interviews, usually undertaken by specially trained lay interviewers.

Questionnaires

The Verbal Autopsy Questionnaires used by the various sites differed, but in most cases they were a derivation from the original WHO Verbal Autopsy questionnaire.

<http://www.who.int/healthinfo/statistics/verbalautopsystandards/en/index1.html>

Data Processing & Appraisal

Data Editing

One cause of death record was inserted for every death where a verbal autopsy was not conducted. The cause of death assigned in these cases is "XX VA not completed"

Other Processing

All records have been anonymised by associating a randomly generated serial number with each case. To facilitate linkage to other data sets released on the INDEPTH Data Repository a linkage file can be requested from the investigators.

Accessibility

Access Authority	Osman Sankoh (INDEPTH Network) , osman.sankoh@indepth-network.org Kobus Herbst (INDEPTH Network) , kherbst@afriacentre.ac.za
-------------------------	--

Contact(s)	Kobus Herbst (INDEPTH Network) , kherbst@afriacentre.ac.za Peter Byass (WHO Collaborating Centre for Verbal Autopsy, Umeå University) , peter.byass@epiph.umu.se
-------------------	---

Access Conditions

This data is made available for licensed access under the following conditions:

1. Data and other material provided by INDEPTH will not be redistributed or sold to other individuals, institutions or organisations without INDEPTH's written agreement.
2. In the case of multi-centre datasets, data originating from a single contributing member centre of the INDEPTH Network may not be analysed or reported on in isolation without the express permission of the member centre concerned.

3. No attempt will be made to re-identify respondents, and there will be no use of the identity of any person or establishment discovered inadvertently. Any such discovery will be reported immediately to INDEPTH.
4. No attempt will be made to produce links between datasets provided by INDEPTH or between INDEPTH data and other datasets that could identify individuals.
5. Any books, articles, conference papers, theses, dissertations, reports or other publications employing data obtained from INDEPTH will cite the source, in line with the citation requirement provided with the dataset.
6. An electronic copy of all publications based on the requested data will be sent to INDEPTH.
7. The original collector of the data, INDEPTH, and the relevant funding agencies bear no responsibility for the data's use or interpretation or inferences based upon it.
8. Any published use of the data must cite the dataset as specified under the citation requirements.

Citation Requirements

INDEPTH Network. INDEPTH Network Cause-Specific Mortality - Release 2014. Oct 2014. Provided by the INDEPTH Network Data Repository. www.indepth-network.org <<http://www.indepth-network.org>>.10.7796/INDEPTH.GH003.COD2014.v1 <<http://dx.doi.org/10.7796/INDEPTH.GH003.COD2014.v1>>

Rights & Disclaimer

Disclaimer

The user of the data acknowledges that the original collector of the data, INDEPTH, and the relevant funding agencies bear no responsibility for the data's use or interpretation or inferences based upon it.

Files Description

Dataset contains 1 file(s)

CODA_2013_v7_Anonymised	
# Cases	176834
# Variable(s)	17

Variables List

Dataset contains 17 variable(s)

File CODA_2013_v7_Anonymised							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	id	Anonymised individual identifier	continuous	numeric-12.0	176834	0	-
2	site	INDEPTH Centre Code	discrete	character-5	176834	0	-
3	year	Year of death	discrete	numeric-8.0	176834	0	-
4	agegp	Age group at death	discrete	numeric-8.0	176834	0	-
5	sex	Sex	discrete	numeric-8.0	176834	0	-
6	py	Person years	continuous	numeric-9.0	176834	0	-
7	va_comp	Verbal autopsy completed	discrete	numeric-8.0	176834	0	-
8	mal_level	Malaria level	discrete	character-1	163334	0	-
9	hiv_level	HIV level	discrete	character-1	163334	0	-
10	cause	Cause of death	discrete	character-40	176834	0	-
11	wt	Case weight	continuous	numeric-9.0	176834	0	-
12	lik	Likelihood of cause	continuous	numeric-9.0	176834	0	-
13	lik_wt	Weighted likelihood	continuous	numeric-9.0	176834	0	-
14	causen	Numeric cause of death code	discrete	numeric-40.0	176834	0	-
15	site_n	INDEPTH Centre Code	discrete	numeric-8.0	176834	0	-
16	cause_gp	Cause group	discrete	numeric-20.0	176834	0	-
17	age_cat	Age category	discrete	numeric-11.0	176834	0	-

Variables Description

Dataset contains 17 variable(s)

File : CODA_2013_v7_Anonymised

id: Anonymised individual identifier

Information [Type= continuous] [Format=numeric] [Range= 1-98458] [Missing=*]

Statistics [NW/ W] [Valid=176834 /-] [Invalid=0 /-] [Mean=48547.427 /-] [StdDev=28505.009 /-]

Definition Anonymised individual identifier

Recoding and Derivation An anonymised individual identifier that maps to a site specific identifier for the individual

site: INDEPTH Centre Code

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

Statistics [NW/ W] [Valid=176834 /-] [Invalid=0 /-]

Definition INDEPTH Centre Code

Value	Label	Cases	Percentage
BD011		19629	11.1%
BD012		538	0.3%
BD013		2333	1.3%
BD014		5079	2.9%
BF031		11201	6.3%
BF041		1674	0.9%
CI011		1472	0.8%
ET031		881	0.5%
GH011		21524	12.2%
GH031		6239	3.5%
GM011		5024	2.8%
ID011		2564	1.4%
IN011		4214	2.4%
IN021		1258	0.7%
KE011		8160	4.6%
KE021		34693	19.6%
KE031		7193	4.1%
MW011		3236	1.8%
SN011		3349	1.9%
VN012		1367	0.8%
ZA011		19171	10.8%
ZA031		16035	9.1%

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.

year: Year of death

Information [Type= discrete] [Format=numeric] [Range= 1992-2012] [Missing=*]

Statistics [NW/ W] [Valid=176834 /-] [Invalid=0 /-]

Definition Year of death

Value	Label	Cases	Percentage
1992		292	0.2%
1993		527	0.3%
1994		491	0.3%
1995		581	0.3%
1996		543	0.3%

File : CODA_2013_v7_Anonymised

year: Year of death

Value	Label	Cases	Percentage
1997		495	0.3%
1998		1708	1.0%
1999		1774	1.0%
2000		4148	2.3%
2001		4567	2.6%
2002		5056	2.9%
2003		13303	7.5%
2004		16681	9.4%
2005		17418	9.8%
2006		18134	10.3%
2007		17232	9.7%
2008		19923	11.3%
2009		20434	11.6%
2010		19375	11.0%
2011		11660	6.6%
2012		2492	1.4%

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.

agegp: Age group at death

Information	[Type= discrete] [Format=numeric] [Range= 1-7] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-]
Definition	Age group at death

Value	Label	Cases	Percentage
1	neonate	8500	4.8%
2	infant	14306	8.1%
3	1-4 yr	17779	10.1%
4	5-14 yr	6554	3.7%
5	15-49 yr	47532	26.9%
6	50-64 yr	25984	14.7%
7	65+ yr	56179	31.8%

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.

sex: Sex

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-]
Definition	Sex

Value	Label	Cases	Percentage
1	Male	91169	51.6%
2	Female	85665	48.4%

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.

py: Person years

Information	[Type= continuous] [Format=numeric] [Range= 12-62112.7890625] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-] [Mean=9881.931 /-] [StdDev=11556.887 /-]

File : CODA_2013_v7_Anonymised

py: Person years

Definition Total person years of exposure for all individuals under surveillance in the particular year, age group and sex.

va_comp: Verbal autopsy completed

Information [Type= discrete] [Format=numeric] [Range= 0-1] [Missing=*]

Statistics [NW/ W] [Valid=176834 /-] [Invalid=0 /-]

Definition Whether a verbal autopsy has been completed for this case

Value	Label	Cases	Percentage
0	No	13493	7.6%
1	Yes	163341	92.4%

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.

mal_level: Malaria level

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

Statistics [NW/ W] [Valid=163334 /-] [Invalid=0 /-]

Definition Malaria exposure level in the geographic area where the site is situated. The InterVA-4 “high” (value=H) malaria setting was used for all the West African sites, plus the East African sites (with the exceptions, on the grounds of high altitude, of Nairobi, Kenya and Kilite-Awlaelo, Ethiopia), on the basis of local experience. All other sites used the “low” (value=L) setting; the “very low” setting was not used. The InterVA-4 guideline is that the “high” setting is appropriate for an expected malaria CSMF higher than about 1%, though the setting chosen does not result in any great dichotomisation of outputs; the clinical equivalent would be a physician’s knowledge that his/her current case comes from a setting where malaria is more or less likely, irrespective of particular symptoms

Value	Label	Cases	Percentage
H		87873	53.8%
L		75461	46.2%

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.

hiv_level: HIV level

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

Statistics [NW/ W] [Valid=163334 /-] [Invalid=0 /-]

Definition HIV level. The InterVA-4 “high” HIV/AIDS setting (value=H) was used for sites in Kenya, Malawi and South Africa. All other sites used the “low” setting (value=L); the “very low” setting was not used. The InterVA-4 guideline is that the “high” setting is appropriate for an expected HIV/AIDS cause-specific mortality fraction (CSMF) higher than about 1%, though it does not result in any great dichotomisation of outputs; the clinical equivalent is a physician’s knowledge that his/her current case comes from a setting where HIV/AIDS is more or less likely, irrespective of that current case’s particular symptoms.

Value	Label	Cases	Percentage
H		81667	50.0%
L		81667	50.0%

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.

cause: Cause of death

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

Statistics [NW/ W] [Valid=176834 /-] [Invalid=0 /-]

Definition Cause of death from the WHO Verbal Autopsy cause list. See: <http://www.who.int/healthinfo/statistics/verbalautopsystandards/en/> title="WHO">Verbal autopsy standards: ascertaining and attributing causes of death

Value	Label	Cases	Percentage
01.01 Sepsis (non-obstetric)		807	0.5%

File : CODA_2013_v7_Anonymised

cause: Cause of death

Value	Label	Cases	Percentage
01.02	Acute resp infect incl pneumonia	12136	6.9%
01.03	HIV/AIDS related death	11972	6.8%
01.04	Diarrhoeal diseases	2789	1.6%
01.05	Malaria	8149	4.6%
01.06	Measles	288	0.2%
01.07	Meningitis and encephalitis	1312	0.7%
01.08 & 10.05	Tetanus	8	0.0%
01.09	Pulmonary tuberculosis	14141	8.0%
01.10	Pertussis	177	0.1%
01.11	Haemorrhagic fever	25	0.0%
01.99	Other and unspecified infect dis	1170	0.7%
02.01	Oral neoplasms	221	0.1%
02.02	Digestive neoplasms	4879	2.8%
02.03	Respiratory neoplasms	2896	1.6%
02.04	Breast neoplasms	294	0.2%
02.05 & 02.06	Reproductive neoplasms MF	1295	0.7%
02.99	Other and unspecified neoplasms	2190	1.2%
03.01	Severe anaemia	331	0.2%
03.02	Severe malnutrition	1101	0.6%
03.03	Diabetes mellitus	1820	1.0%
04.01	Acute cardiac disease	1909	1.1%
04.02	Stroke	6675	3.8%
04.03	Sickle cell with crisis	288	0.2%
04.99	Other and unspecified cardiac dis	4436	2.5%

File : CODA_2013_v7_Anonymised

cause: Cause of death

Value	Label	Cases	Percentage
05.01	Chronic obstructive pulmonary dis	2012	1.1%
05.02	Asthma	1238	0.7%
06.01	Acute abdomen	4260	2.4%
06.02	Liver cirrhosis	752	0.4%
07.01	Renal failure	764	0.4%
08.01	Epilepsy	464	0.3%
09.01	Ectopic pregnancy	21	0.0%
09.02	Abortion-related death	71	0.0%
09.03	Pregnancy-induced hypertension	120	0.1%
09.04	Obstetric haemorrhage	274	0.2%
09.05	Obstructed labour	14	0.0%
09.06	Pregnancy-related sepsis	214	0.1%
09.07	Anaemia of pregnancy	49	0.0%
09.08	Ruptured uterus	7	0.0%
09.99	Other and unspecified maternal CoD	109	0.1%
10.01	Prematurity	795	0.4%
10.02	Birth asphyxia	1118	0.6%
10.03	Neonatal pneumonia	1444	0.8%
10.04	Neonatal sepsis	457	0.3%
10.06	Congenital malformation	200	0.1%
10.99	Other and unspecified neonatal CoD	674	0.4%
12.01	Road traffic accident	1906	1.1%
12.02	Other transport accident	85	0.0%
12.03	Accid fall	661	0.4%

File : CODA_2013_v7_Anonymised

cause: Cause of death

Value	Label	Cases	Percentage
12.04	Accid drowning and submersion	833	0.5%
12.05	Accid expos to smoke fire & flame	297	0.2%
12.06	Contact with venomous plant/animal	189	0.1%
12.07	Accid poisoning & noxious subs	86	0.0%
12.08	Intentional self-harm	892	0.5%
12.09	Assault	1586	0.9%
12.10	Exposure to force of nature	35	0.0%
12.99	Other and unspecified external CoD	426	0.2%
98	Other and unspecified NCD	976	0.6%
99	Indeterminate	59003	33.4%
XX	VA not completed	13493	7.6%

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.

wt: Case weight

Information	[Type= continuous] [Format=numeric] [Range= 0.42280301451683-5.73882818222046] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-] [Mean=1.007 /-] [StdDev=0.334 /-]
Definition	Case weight. Calculated as the ratio of the proportion of the standard population in the relevant age/sex group to the proportion of the total person years of exposure of the relevant age/sex group in the site.

lik: Likelihood of cause

Information	[Type= continuous] [Format=numeric] [Range= 0-1] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-] [Mean=0.633 /-] [StdDev=0.383 /-]
Definition	Likelihood of cause being the true cause of death as calculated by the InterVA4 program.

lik_wt: Weighted likelihood

Information	[Type= continuous] [Format=numeric] [Range= 0-5.73882818222046] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-] [Mean=0.637 /-] [StdDev=0.459 /-]
Definition	Weighted likelihood. lik times wt.

causen: Numeric cause of death code

Information	[Type= discrete] [Format=numeric] [Range= 10-71] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-]
Definition	Numeric cause of death code from the WHO Verbal Autopsy cause list.

File : CODA_2013_v7_Anonymised

causen: Numeric cause of death code

Value	Label	Cases	Percentage
10	01.01 Sepsis (non-obstetric)	807	0.5%
11	01.02 Acute resp infect, incl pneumonia	12136	6.9%
12	01.03 HIV/AIDS related death	11972	6.8%
13	01.04 Diarrhoeal diseases	2789	1.6%
14	01.05 Malaria	8149	4.6%
15	01.06 Measles	288	0.2%
16	01.07 Meningitis and encephalitis	1312	0.7%
17	01.08, 10.05 Tetanus	8	0.0%
18	01.09 Pulmonary tuberculosis	14141	8.0%
19	01.10 Pertussis	177	0.1%
20	01.11 Haemorrhagic fever	25	0.0%
21	01.99 Other and unspecified infect dis	1170	0.7%
22	02.01 Oral neoplasms	221	0.1%
23	02.02 Digestive neoplasms	4879	2.8%
24	02.03 Respiratory neoplasms	2896	1.6%
25	02.04 Breast neoplasms	294	0.2%
26	02.05, 02.06 Reproductive neoplasms M,F	1295	0.7%
27	02.99 Other and unspecified neoplasms	2190	1.2%
28	03.01 Severe anaemia	331	0.2%
29	03.02 Severe malnutrition	1101	0.6%
30	03.03 Diabetes mellitus	1820	1.0%
31	04.01 Acute cardiac disease	1909	1.1%
32	04.03 Sickle cell with crisis	288	0.2%
33	04.02 Stroke	6675	3.8%
34	04.99 Other and unspecified cardiac dis	4436	2.5%
35	05.01 Chronic obstructive pulmonary dis	2012	1.1%
36	05.02 Asthma	1238	0.7%
37	06.01 Acute abdomen	4260	2.4%
38	06.02 Liver cirrhosis	752	0.4%
39	07.01 Renal failure	764	0.4%
40	08.01 Epilepsy	464	0.3%
41	98 Other and unspecified NCD	976	0.6%
42	10.06 Congenital malformation	200	0.1%
43	10.01 Prematurity	795	0.4%
44	10.02 Birth asphyxia	1118	0.6%
45	10.03 Neonatal pneumonia	1444	0.8%
46	10.04 Neonatal sepsis	457	0.3%
47	10.99 Other and unspecified neonatal CoD	674	0.4%
48	11.01 Fresh stillbirth	0	
49	11.02 Macerated stillbirth	0	
50	12.01 Road traffic accident	1906	1.1%
51	12.02 Other transport accident	85	0.0%
52	12.03 Accid fall	661	0.4%

File : CODA_2013_v7_Anonymised

causen: Numeric cause of death code

Value	Label	Cases	Percentage
53	12.04 Accid drowning and submersion	833	0.5%
54	12.05 Accid expos to smoke, fire & flame	297	0.2%
55	12.06 Contact with venomous plant/animal	189	0.1%
56	12.10 Exposure to force of nature	35	0.0%
57	12.07 Accid poisoning and noxious subs	86	0.0%
58	12.08 Intentional self-harm	892	0.5%
59	12.09 Assault	1586	0.9%
60	12.99 Other and unspecified external CoD	426	0.2%
61	09.01 Ectopic pregnancy	21	0.0%
62	09.02 Abortion-related death	71	0.0%
63	09.03 Pregnancy-induced hypertension	120	0.1%
64	09.04 Obstetric haemorrhage	274	0.2%
65	09.05 Obstructed labour	14	0.0%
66	09.06 Pregnancy-related sepsis	214	0.1%
67	09.07 Anaemia of pregnancy	49	0.0%
68	09.08 Ruptured uterus	7	0.0%
69	09.99 Other and unspecified maternal CoD	109	0.1%
70	99 Indeterminate	59003	33.4%
71	XX Va not completed	13493	7.6%

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.

site_n: INDEPTH Centre Code

Information	[Type= discrete] [Format=numeric] [Range= 1-22] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-]
Definition	INDEPTH Centre Code numerically encoded

Value	Label	Cases	Percentage
1	BD011	19629	11.1%
2	BD012	538	0.3%
3	BD013	2333	1.3%
4	BD014	5079	2.9%
5	BF031	11201	6.3%
6	BF041	1674	0.9%
7	CI011	1472	0.8%
8	ET031	881	0.5%
9	GH011	21524	12.2%
10	GH031	6239	3.5%
11	GM011	5024	2.8%
12	ID011	2564	1.4%
13	IN011	4214	2.4%
14	IN021	1258	0.7%
15	KE011	8160	4.6%
16	KE021	34693	19.6%
17	KE031	7193	4.1%
18	MW011	3236	1.8%

File : CODA_2013_v7_Anonymised

site_n: INDEPTH Centre Code

Value	Label	Cases	Percentage
19	SN011	3349	1.9%
20	VN012	1367	0.8%
21	ZA011	19171	10.8%
22	ZA031	16035	9.1%

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.

cause_gp: Cause group

Information	[Type= discrete] [Format=numeric] [Range= 1-6] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-]
Definition	Causes grouped into larger categories of cause-specific mortality

Value	Label	Cases	Percentage
1	01 Communicable	52974	30.0%
2	02 Neoplasms	11775	6.7%
3	03 Non-communicable	27026	15.3%
4	04 Maternal/Neonatal	5567	3.1%
5	05 Trauma	6996	4.0%
6	06 Indeterminate	72496	41.0%

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.

age_cat: Age category

Information	[Type= discrete] [Format=numeric] [Range= 1-3] [Missing=*]
Statistics [NW/ W]	[Valid=176834 /-] [Invalid=0 /-]
Definition	Age groups classified into larger categories

Value	Label	Cases	Percentage
1	01 Children	47139	26.7%
2	02 Adults	47532	26.9%
3	03 Elderly	82163	46.5%

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.