



Encontro sobre a Melhoria da qualidade da informação sobre causas de morte no Brasil

1 a 6 de outubro de 2017 • Porto de Galinhas/PE



D4H CRVS INITIATIVE

Ten steps to better data quality: the sequential structure of ANACONDA

Dr. Lene Mikkelsen
2 October, 2017
Recife Brazil

Welcome to the ANACONDA Tool: Getting started

Welcome to ANACONDA

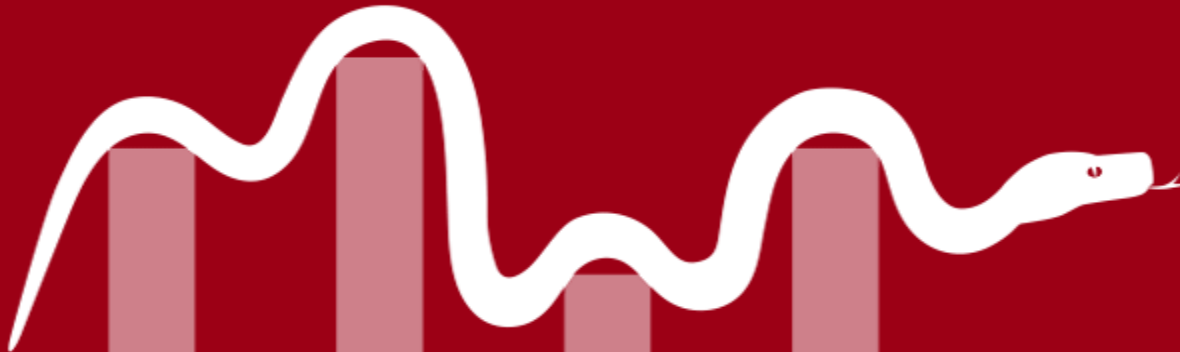
English ▾

ANACONDA is a tool for checking the quality of mortality data. It provides a step-by-step approach to enable users to quickly conduct a comprehensive review of data on mortality levels and causes of death. In particular, ANACONDA will help you identify probable errors, misdiagnoses and inconsistencies in your mortality and cause of death data.

Learn more about ANACONDA...

Start data entry and import into ANACONDA


Open existing ANACONDA file from workspace







Version 3.1.0







Data input and general background checks

- Step 1:**  **Input Data**
Provides a summary overview of the input data and performs initial data validity checks


Mortality data

- Step 2:**  **Crude Death Rate and completeness of death reporting**
Calculation of Crude Death Rate as an initial check on data completeness using comparator data
- Step 3:**  **Age-specific mortality rates**
Review of plausibility of age-specific mortality rates using established demographic patterns
- Step 4:**  **Age- sex- distribution of deaths**
Comparison of the distribution of deaths by sex and age with a regional comparator
- Step 5:**  **Completeness of child mortality data**
Compares the level of under-five mortality rates based on registered data with estimates from censuses and surveys


Cause of death data

- Step 6:**  **Classification of deaths into broad cause of death (COD) groups**
Provides an initial assessment of plausibility of broad causes of death patterns
- Step 7:**  **Quality of cause of death data**
Analyses the nature and extent of "unusable" codes in the data and relative importance of different types of ill-defined deaths
- Step 8:**  **Age pattern of mortality for broad disease and injury groups**
Assesses the plausibility of age patterns of reported death for broad cause groups
- Step 9:**  **Leading causes of death**
Compares the 20 leading causes of death with the regional pattern to assess plausibility


Summary index of mortality data quality

- Step 10:**  **Vital statistics performance index - VSPI(Q)**
Calculates and disaggregates an overall summary index of mortality data quality

Review summary report

-  **Review**
Consolidated summary report based on the application of the 10 steps to the input data

References

-  **References**
Provides a list of references cited in the tool

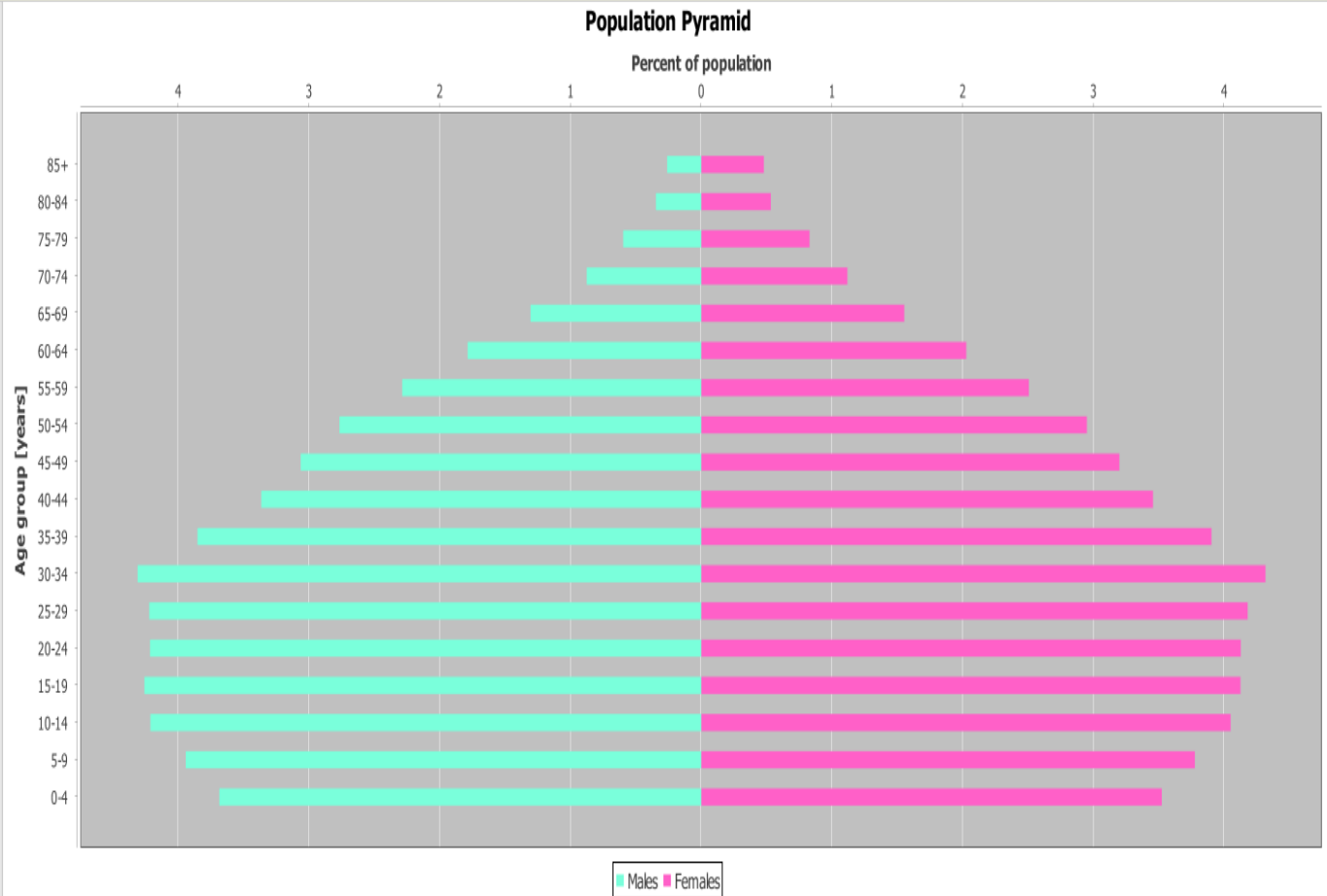


- **STEP 1:**
- **Input data**

- Performs numerical checks on the population and mortality input data
- Draws population and death pyramids
- Tabulates the data according to different classifications
- Shows death distributions for different age groups

Step 1.3 checks your input population data

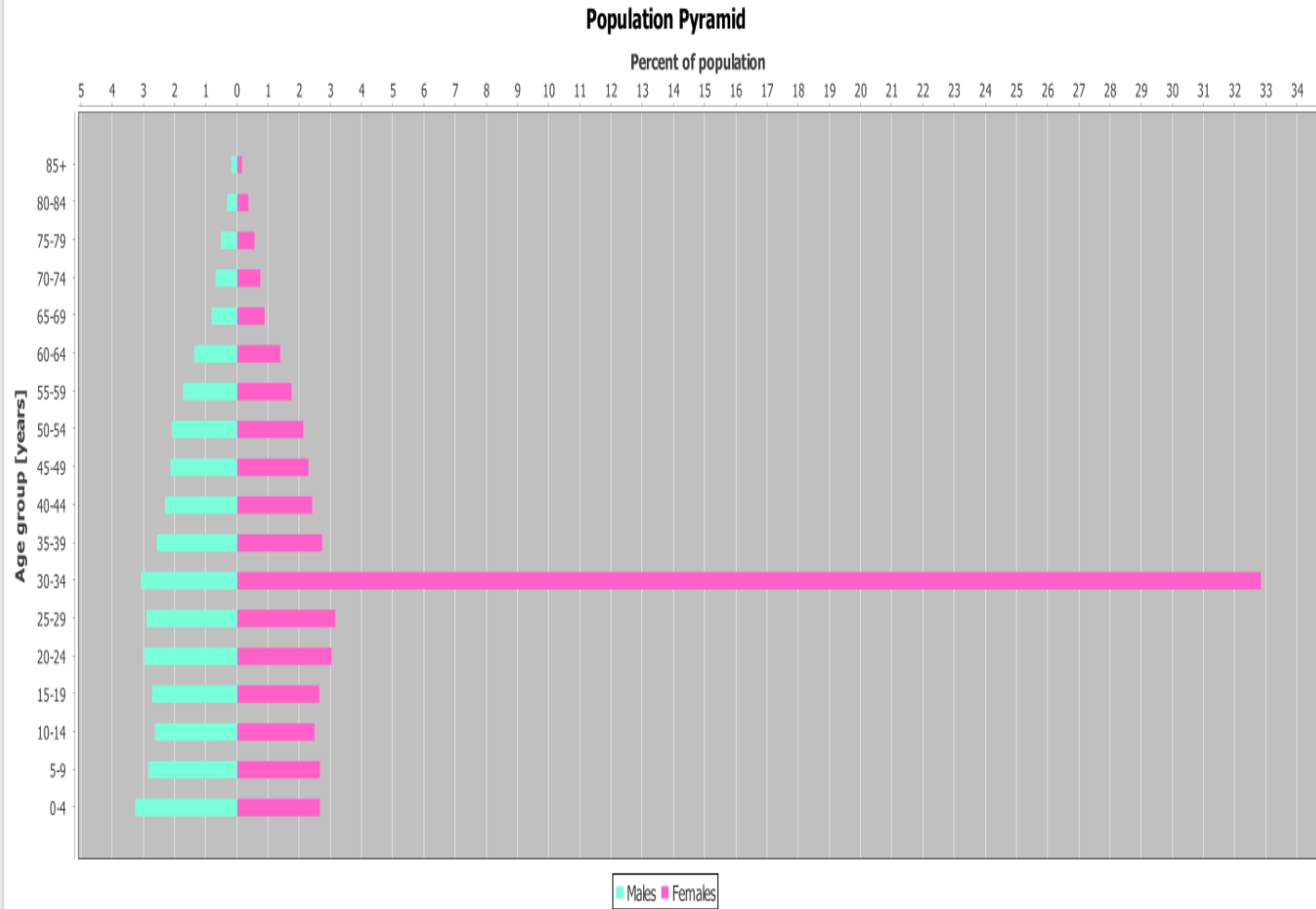
Population				
Age groups	Males	% of total	Females	% of total
0 years	1,472,923	0.7	1,405,421	0.7
1-4 years	6,065,132	3.0	5,794,264	2.8
5-9 years	8,062,852	3.9	7,716,257	3.8
10-14 years	8,616,189	4.2	8,276,054	4.0
15-19 years	8,710,123	4.3	8,430,077	4.1
20-24 years	8,622,007	4.2	8,434,416	4.1
25-29 years	8,634,055	4.2	8,542,753	4.2
30-34 years	8,816,331	4.3	8,821,076	4.3
35-39 years	7,879,629	3.9	7,976,626	3.9
40-44 years	6,882,205	3.4	7,062,021	3.5
45-49 years	6,266,080	3.1	6,536,317	3.2
50-54 years	5,659,620	2.8	6,027,724	2.9
55-59 years	4,678,733	2.3	5,120,879	2.5
60-64 years	3,655,012	1.8	4,142,038	2.0
65-69 years	2,672,038	1.3	3,172,665	1.6
70-74 years	1,793,495	0.9	2,283,016	1.1
75-79 years	1,222,314	0.6	1,691,282	0.8
80-84 years	710,926	0.3	1,085,523	0.5
85+ years	535,858	0.3	976,718	0.5
All ages	100,955,522	49.4	103,495,127	50.6



Brazil, 2015

Spot the error!

Population				
Age groups	Males	% of total	Females	% of total
0 years	115000	0.7	108000	0.7
1-4 years	391600	2.5	301300	1.9
5-9 years	441600	2.9	409300	2.6
10-14 years	408800	2.6	382500	2.5
15-19 years	422600	2.7	406700	2.6
20-24 years	460800	3.0	466600	3.0
25-29 years	450900	2.9	485800	3.1
30-34 years	478400	3.1	5076600	32.8
35-39 years	398500	2.6	420300	2.7
40-44 years	358200	2.3	371300	2.4
45-49 years	332000	2.1	352500	2.3
50-54 years	325200	2.1	326900	2.1
55-59 years	269000	1.7	267400	1.7
60-64 years	213400	1.4	213000	1.4
65-69 years	127400	0.8	135400	0.9
70-74 years	107300	0.7	113800	0.7
75-79 years	81900	0.5	85300	0.6
80-84 years	50850	0.3	55250	0.4
85+ years	29684	0.2	23052	0.1
All ages	5463134	35.3	10001002	64.7

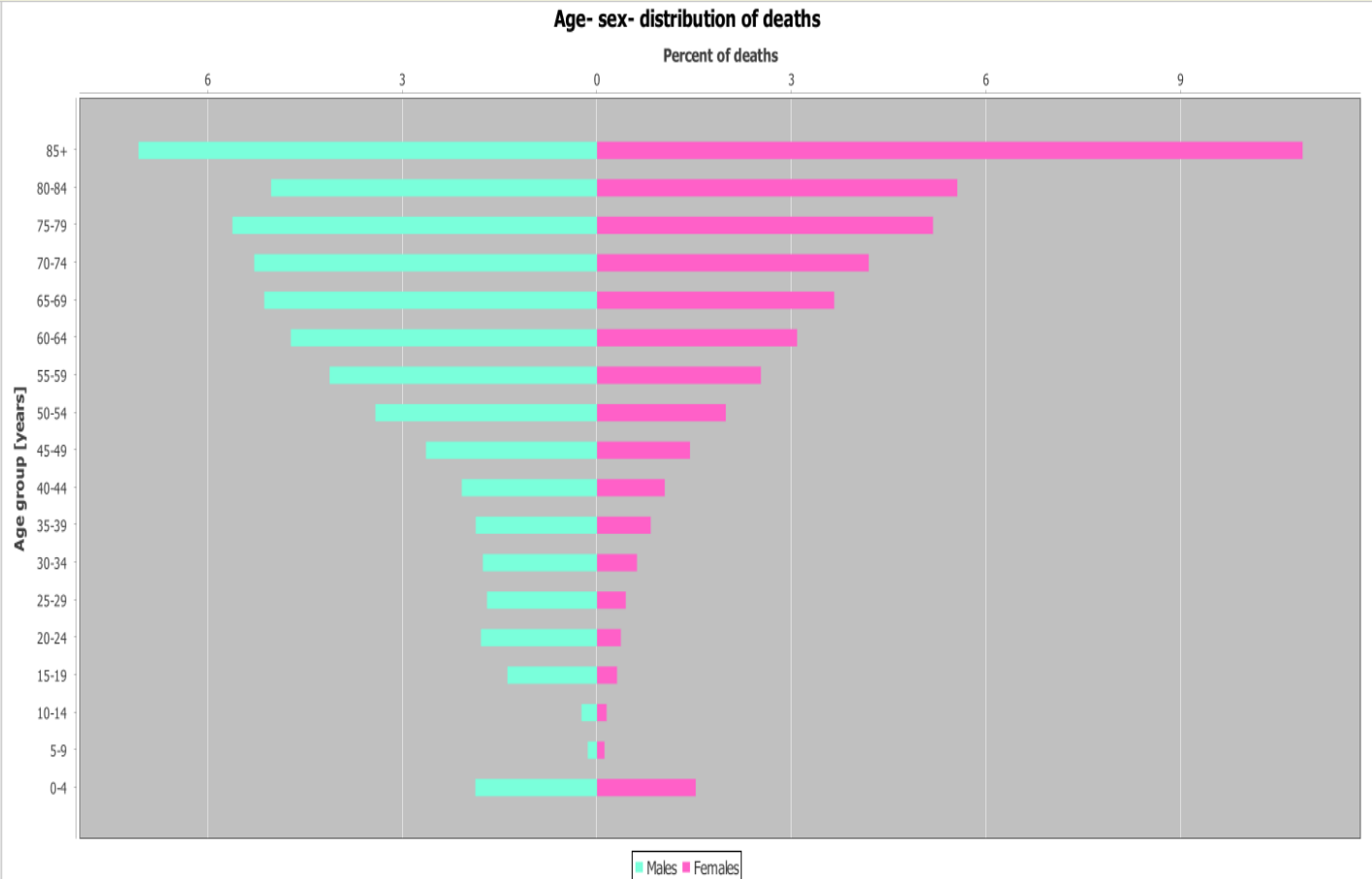


Explain your population pyramid



Step 1.4 Age and sex-distribution of deaths

Deaths				
Age groups	Males	% of total	Females	% of total
0 years	20,653	1.6	16,592	1.3
1-4 years	3,014	0.2	2,563	0.2
5-9 years	1,813	0.1	1,435	0.1
10-14 years	3,036	0.2	1,841	0.1
15-19 years	17,430	1.4	3,869	0.3
20-24 years	22,576	1.8	4,607	0.4
25-29 years	21,381	1.7	5,562	0.4
30-34 years	22,204	1.8	7,740	0.6
35-39 years	23,588	1.9	10,414	0.8
40-44 years	26,290	2.1	13,136	1.0
45-49 years	33,246	2.6	18,044	1.4
50-54 years	43,082	3.4	25,002	2.0
55-59 years	51,973	4.1	31,819	2.5
60-64 years	59,507	4.7	38,848	3.1
65-69 years	64,660	5.1	46,063	3.7
70-74 years	66,616	5.3	52,769	4.2
75-79 years	70,851	5.6	65,263	5.2
80-84 years	63,338	5.0	69,973	5.6
85+ years	89,098	7.1	137,066	10.9
All ages	706,931	56.1	552,973	43.9



Brazil, 2015

Useful ways of checking data is by tabulating them differently

- By sex and age groups
- Aggregate the individual diseases into **groups of public health interest** (e.g. transport accidents, assaults, pregnancy&childbirths, etc.)
- ANACONDA aggregates the input data into two standard tabulation lists:
 - **WHO 103 cause list** (useful for general comparative mortality analysis)
 - **GBD 300 list** (useful for more detailed specific disease comparisons, e.g. alzheimer, breast cancer, measles, etc.)

Overview tabulation of input data

Code	Description	Total	<5 Males	<5 Females	5+ Males	5+ Females
A02.0	Salmonella enteritis	5	0	0	1	4
A02.2	Localized salmonella infections	0	0	0	0	0
A02.9	Salmonella infection, unspecified	3	0	0	2	1
A04.7	Enterocolitis due to Clostridium difficile	24	11	3	6	4
A04.9	Bacterial intestinal infection, unspecified	60	3	3	28	26
A05.9	Bacterial foodborne intoxication, unspecified	1	0	0	1	0
A06.0	Acute amoebic dysentery	2	0	0	1	1
A06.9	Amoebiasis, unspecified	2	0	0	0	2
A08.4	Viral intestinal infection, unspecified	5	0	0	3	2
A09.-	Diarrhoea and gastroenteritis of presumed infectious origin	0	0	0	0	0
A15.0	Tuberculosis of lung, confirmed by sputum microscopy with or without culture	6	0	0	6	0
A15.1	Tuberculosis of lung, confirmed by culture only	2	0	0	2	0
A15.2	Tuberculosis of lung, confirmed histologically	2	0	0	1	1
A15.3	Tuberculosis of lung, confirmed by unspecified means	6	0	0	4	2
A16.2	Tuberculosis of lung, without mention of bacteriological or histological confirmation	685	3	2	488	192
A16.5	Tuberculous pleurisy, without mention of bacteriological or histological confirmation	22	0	0	14	8
A16.8	Other respiratory tuberculosis, without mention of bacteriological or histological confirmation	1	0	0	1	0
A16.9	Respiratory tuberculosis unspecified, without mention of bacteriological or histological confirmation	114	0	0	85	29

WHO 103 cause tabulation list

Grouping of total deaths according to the ICD-10 mortality tabulation list 1 (103 aggregated codes) - Males

Aggregated code	Description	ICD code range	<5	5-9	10-19	20-24	25-29	30-49	50-69	70+	Total
1-002	Diseases of the skin and subcutaneous tissue	L00-L99	0	0	3	1	2	34	00	201	339
1-083	Diseases of the musculoskeletal system and connective tissue	M00-M99	7	9	29	11	11	53	121	367	608
1-084	Diseases of the genitourinary system	N00-N99	25	2	21	13	26	151	701	2350	3290
1-085	Glomerular and renal tubulointerstitial diseases	N00-N15	4	2	6	1	1	8	16	77	115
1-086	Remainder of diseases of the genitourinary< system	N17-N98	21	0	15	12	25	143	685	2273	3175
1-087	Pregnancy, childbirth and the puerperium	O00-O99	0	0	0	0	0	0	0	0	0
1-088	Pregnancy with abortive outcome	O00-O07	0	0	0	0	0	0	0	0	0
1-089	Other direct obstetric deaths	O10-O92	0	0	0	0	0	0	0	0	0
1-090	Indirect obstetric deaths	O98-O99	0	0	0	0	0	0	0	0	0
1-091	Remainder of pregnancy, childbirth and the puerperium	O95-O97	0	0	0	0	0	0	0	0	0
1-092	Certain conditions originating in the perinatal period	P00-P96	2195	10	3	2	1	1	0	0	2212
1-093	Congenital malformations, deformations and chromosomal abn...	Q00-Q99	1072	25	41	28	15	41	30	13	1266
1-094	Symptoms, signs and abnormal clinical and laboratory findings, ...	R00-R99	117	21	57	67	74	342	552	965	2213
1-095	External causes of morbidity and mortality	V01-Y89	222	163	2680	3694	3026	7347	3512	1759	22568
1-096	Transport accidents	V01-V99	49	52	556	848	715	1735	1122	607	5704
1-097	Falls	W00-W19	6	13	17	44	37	154	288	321	883
1-098	Accidental drowning and submersion	W65-W74	51	45	136	84	53	129	79	25	608
1-099	Exposure to smoke, fire and flames	X00-X09	4	2	2	6	6	9	15	16	60
1-100	Accidental poisoning by and exposure to noxious substances	X40-X49	1	1	3	6	7	24	11	1	54
1-101	Intentional self-harm	X60-X84	0	1	232	219	186	535	359	134	1678
1-102	Assault	X85-Y09	15	17	1552	2271	1837	4143	1106	152	11197
1-103	All other external causes	W20-W64,W75-W99,X10-X39,X50-X59,Y10-Y89	96	32	182	216	185	618	532	503	2384
1-901	SARS	U04	0	0	0	0	0	0	0	0	0
Total			4707	513	3594	4456	3955	14200	28419	54257	114302

GBD tabulation list (300+ causes)

COD data compiled according to the GBD tabulation list - Males

GBD classification	Description	<5	5-9	10-19	20-24	25-29	30-49	50-69	70+	Total
_gc	Garbage Code	771	138	525	566	511	2144	4917	12554	22170
cirrhosis	Cirrhosis and other chronic liver diseases	12	5	7	10	10	207	855	612	1719
ckd	Chronic kidney disease	3	0	7	5	16	82	385	1054	1552
ckd_diabetes	Chronic kidney disease due to diabetes mellitus	0	0	0	1	3	48	322	453	827
ckd_glomerulo	Chronic kidney disease due to glomerulonephritis	2	2	3	0	0	1	5	6	19
ckd_htn	Chronic kidney disease due to hypertension	0	0	3	6	8	67	267	928	1279
ckd_other	Chronic kidney disease due to other causes	8	1	0	0	0	2	4	3	18
cong_chromo	Other chromosomal abnormalities	36	1	2	1	1	1	0	0	42
cong_cleft	Cleft lip and cleft palate	1	0	0	0	0	0	0	0	1
cong_digestive	Digestive congenital anomalies	114	3	3	3	0	2	0	4	129
cong_downs	Down syndrome	32	0	2	5	0	7	9	0	56
cong_heart	Congenital heart anomalies	525	10	21	15	10	22	12	4	619
cong_msk	Congenital musculoskeletal and limb anomalies	59	0	0	0	0	0	1	1	61
cong_neural	Neural tube defects	21	0	1	1	0	2	0	0	25
cong_other	Other congenital birth defects	248	10	12	3	4	3	4	1	285
cong_turner	Turner syndrome	0	0	0	0	0	0	0	0	0
cong_urogenital	Urogenital congenital anomalies	45	0	0	0	0	1	0	0	46
cvd_afib	Atrial fibrillation and flutter	0	0	0	0	0	0	0	0	0
cvd_aortic	Aortic aneurysm	0	0	3	1	4	57	247	486	798
cvd_cmp_myocarditis	Myocarditis	4	0	4	0	0	5	4	2	19
cvd_cmp_other	Other cardiomyopathy	9	0	2	3	0	6	11	20	51
cvd_endo	Endocarditis	5	0	5	5	3	24	118	110	270
cvd_htn	Hypertensive heart disease	1	0	1	1	0	29	302	1340	1675
cvd_ihd	Ischemic heart disease	0	0	32	31	48	887	5687	12969	19657

All lists allow you to rank the data and extract these for reports or presentations

Code	Description	Total▼	<5 Males	<5 Females	5+ Males	5+ Females
Total		96460	3092	2413	48695	42260
J18.9	Pneumonia, unspecified	9755	249	191	4763	4552
A41.9	Septicaemia, unspecified	7158	245	211	3177	3525
I21.9	Acute myocardial infarction, unspecified	3929	1	0	2241	1687
I50.0	Congestive heart failure	3406	0	0	1740	1666
J84.1	Other interstitial pulmonary diseases with fibrosis	3100	0	2	1595	1503
I10X	(definition not found)	3049	0	0	1499	1550
X59.9	Exposure to unspecified factor, unspecified place. As from 2006: Exposure to unspecified ...	2967	95	84	1947	841
J18.0	Bronchopneumonia, unspecified	2803	88	83	1357	1275
C16.9	Stomach, unspecified	2740	1	2	1460	1277
K74.6	Other and unspecified cirrhosis of liver	2466	0	2	1514	950

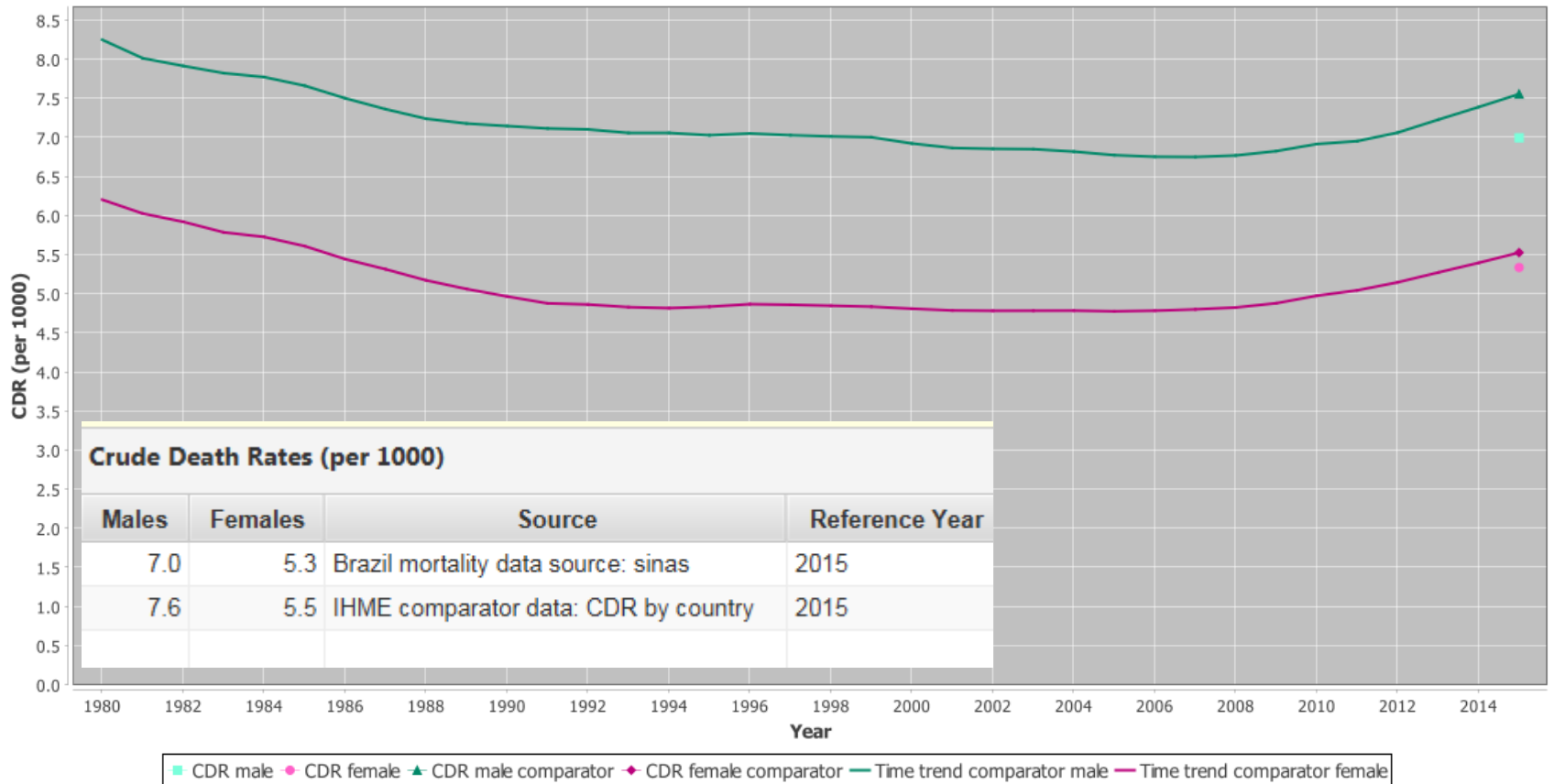


STEP 2: Crude death rates

- Calculates the CDR
- Compares the observed level with other estimates of CDR to assess potential under reporting
- Estimates the true completeness from the input data using a modelled approach

CDR and completeness of death recording

Time-trend in comparator data



Brazil, 2015

Interpreting the CDR

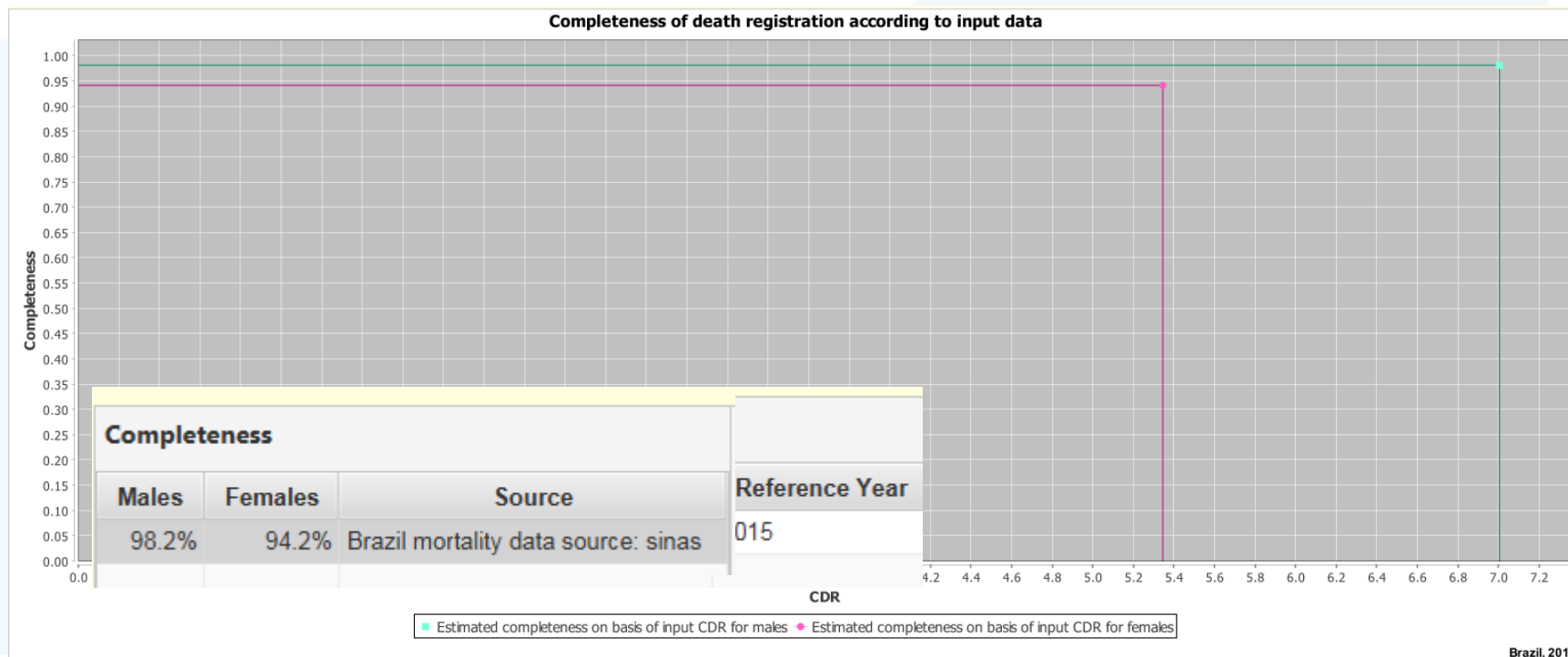
- Knowing the population age structure will help you interpret whether a low CDR is true or indicative of under-registration of deaths
- Which type of populations have high CDR?

Estimating completeness from your own data

What are the three indicators that you need to estimate the true completeness of deaths registration?

1. Crude death rate (CDR) – i.e. registered deaths divided by population
2. % of population aged 65+
3. Under-five mortality rate (${}_5q_0$)

Predicting completeness in Brazil on 2015 data





STEP 3: Age-sex specific mortality rates (ASMR)

- Identifies potential age-sex problems in your data
- Checks whether your ASMR of males and females are plausible
- Or are they affected by under-registration or by bad age reporting

Standard age-sex patterns

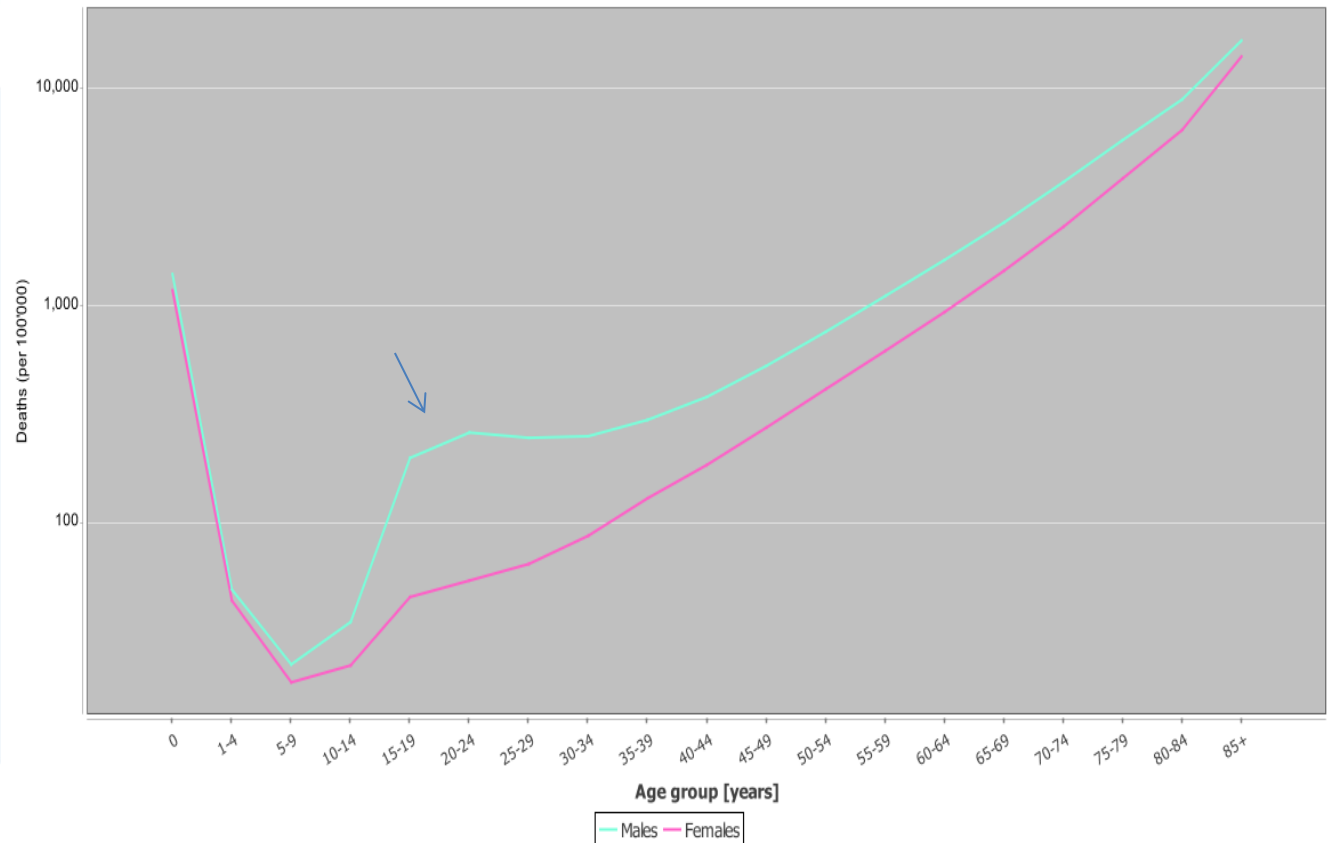
- **AGE:**
 - At what ages do we expect to have the highest mortality rates?
 - At what ages are the ASMR the lowest?
 - From what age do the rates rise EXPONENTIALLY?
- **SEX:**
 - Which sex has the highest mortality rates?
 - Any exception to this?

ASMR pattern for males and females in Brazil 2015.

Deaths per 100'000

Age groups	Males	Females
0 years	1,402.2	1,180.6
1-4 years	49.7	44.2
5-9 years	22.5	18.6
10-14 years	35.2	22.2
15-19 years	200.1	45.9
20-24 years	261.8	54.6
25-29 years	247.6	65.1
30-34 years	251.9	87.7
35-39 years	299.4	130.6
40-44 years	382.0	186.0
45-49 years	530.6	276.1
50-54 years	761.2	414.8
55-59 years	1,110.8	621.4
60-64 years	1,628.1	937.9
65-69 years	2,419.9	1,451.9
70-74 years	3,714.3	2,311.4
75-79 years	5,796.5	3,858.8
80-84 years	8,909.2	6,446.0
85+ years	16,627.2	14,033.3
All ages	700.2	534.3

Age-specific death rates



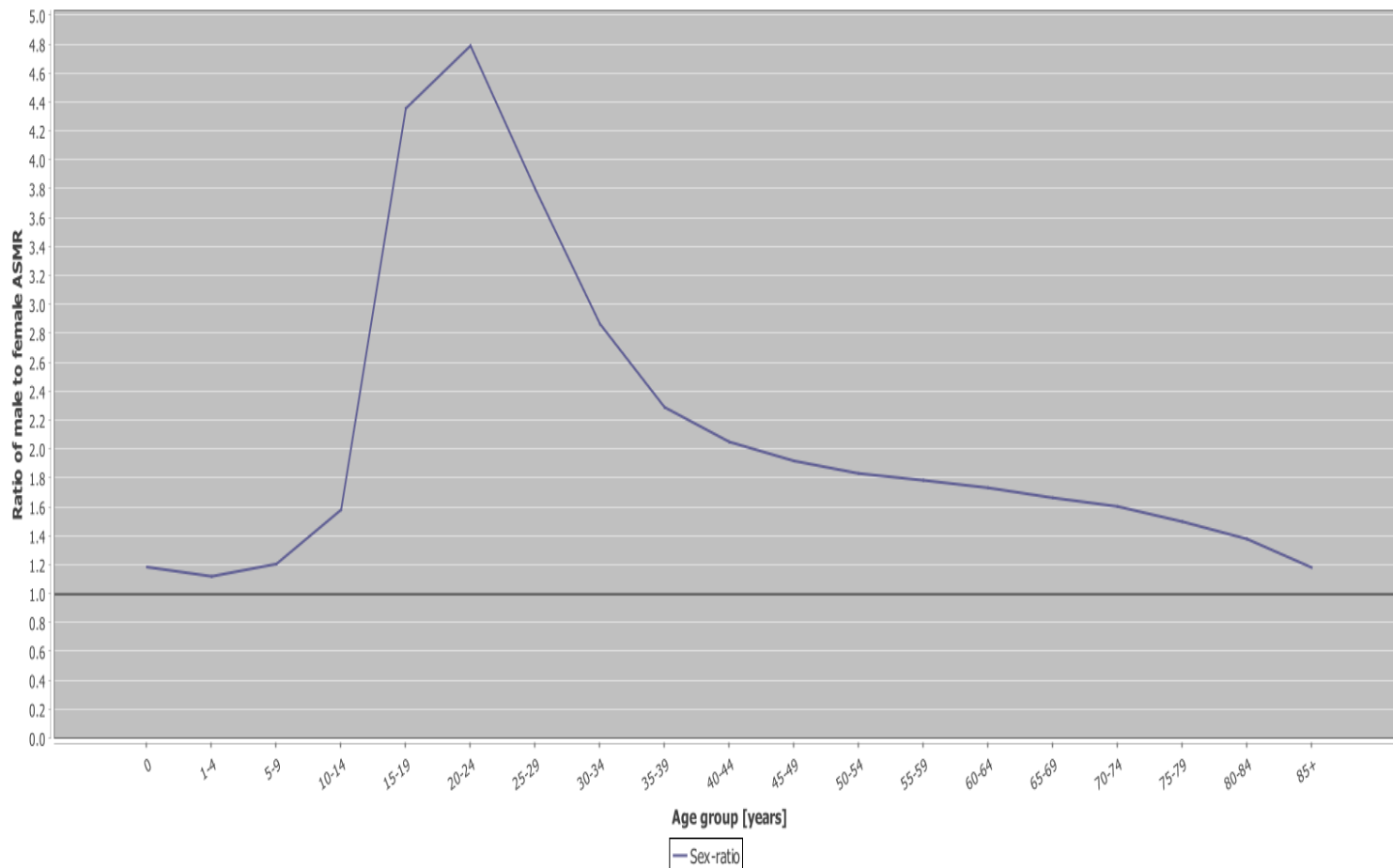
Brazil, 2015

Excess male mortality in Brazil

Sex ratio of age specific mortality rate (ASMR)

Age groups	Sex-ratio
0 years	1.2
1-4 years	1.1
5-9 years	1.2
10-14 years	1.6
15-19 years	4.4
20-24 years	4.8
25-29 years	3.8
30-34 years	2.9
35-39 years	2.3
40-44 years	2.1
45-49 years	1.9
50-54 years	1.8
55-59 years	1.8
60-64 years	1.7
65-69 years	1.7
70-74 years	1.6
75-79 years	1.5
80-84 years	1.4
85+ years	1.2
All ages	1.3

Sex ratio of age specific mortality rate (ASMR)



Brazil, 2015



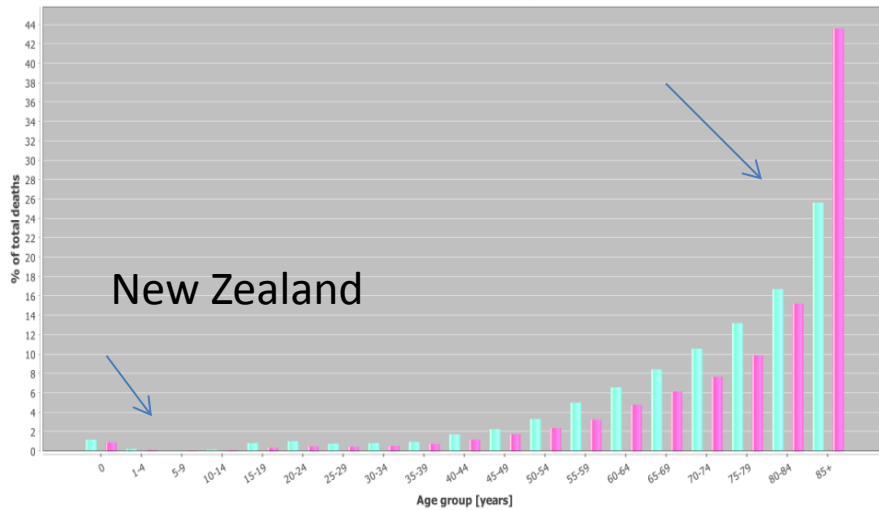
STEP 4:

Age and sex distribution of deaths

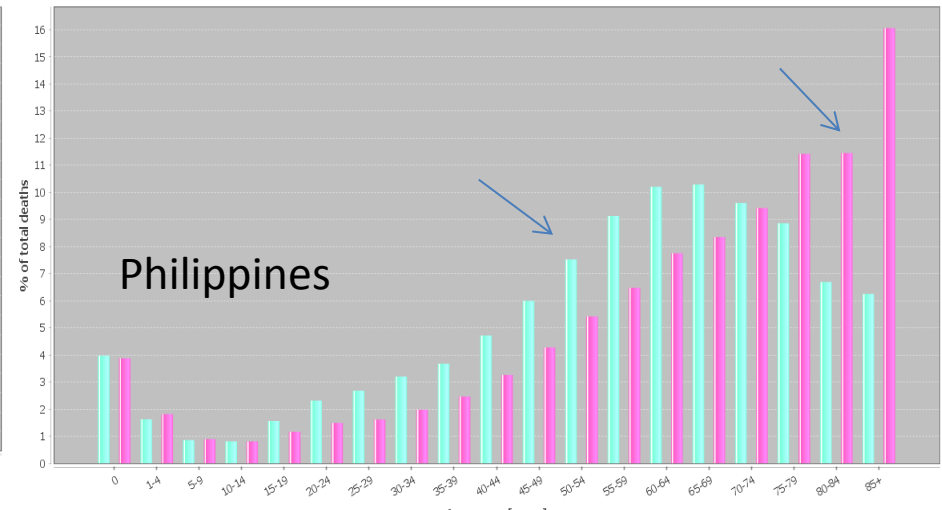
- Looks at the number of deaths and checks how the reported male and female deaths are distributed by age
- Do these patterns look plausible or are there signs of misreporting?
- Compares your data to a regional comparator

Different death distributions. Comparison between countries (NZ, PH, Myanmar)

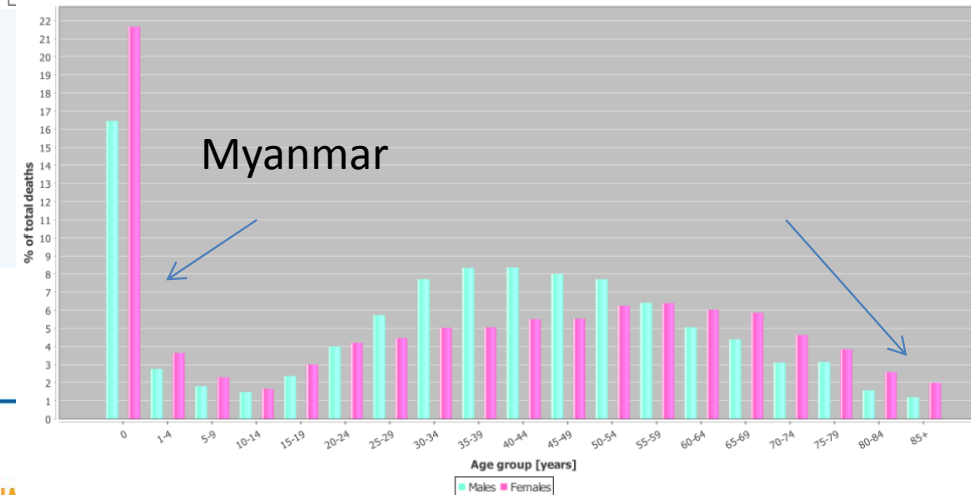
Distribution of deaths by age



Distribution of deaths by age



Distribution of deaths by age



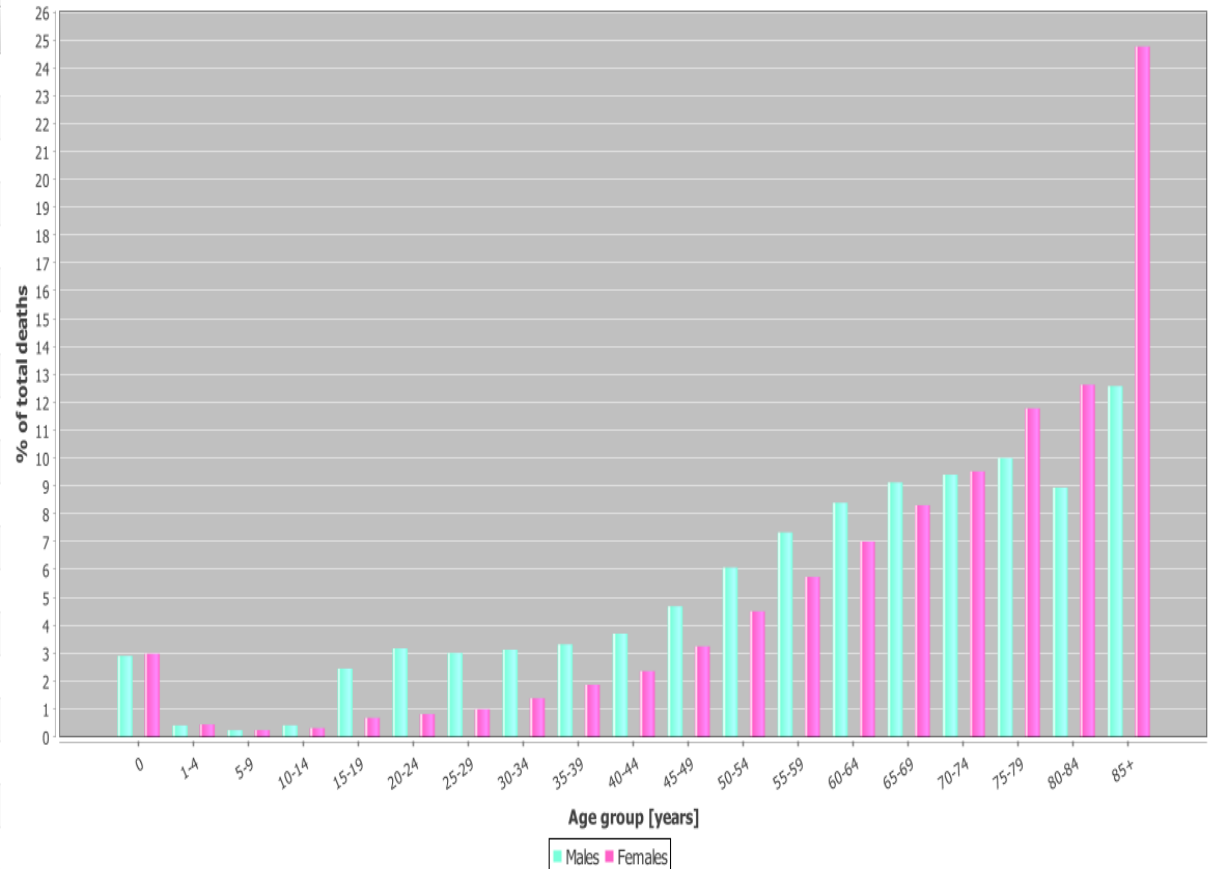
Philippines, 2014

Distribution of deaths by age and sex in Brazil

Age- sex- distribution of deaths

Age groups	Males	Males %	Females	Females %
0 years	20,653	2.9	16,592	3.0
1-4 years	3,014	0.4	2,563	0.5
5-9 years	1,813	0.3	1,435	0.3
10-14 years	3,036	0.4	1,841	0.3
15-19 years	17,430	2.5	3,869	0.7
20-24 years	22,576	3.2	4,607	0.8
25-29 years	21,381	3.0	5,562	1.0
30-34 years	22,204	3.1	7,740	1.4
35-39 years	23,588	3.3	10,414	1.9
40-44 years	26,290	3.7	13,136	2.4
45-49 years	33,246	4.7	18,044	3.3
50-54 years	43,082	6.1	25,002	4.5
55-59 years	51,973	7.4	31,819	5.8
60-64 years	59,507	8.4	38,848	7.0
65-69 years	64,660	9.1	46,063	8.3
70-74 years	66,616	9.4	52,769	9.5
75-79 years	70,851	10.0	65,263	11.8
80-84 years	63,338	9.0	69,973	12.7
85+ years	89,098	12.6	137,066	24.8
All ages	706,931	100.0	552,973	100.0

Distribution of deaths by age



Brazil, 2015



STEP 5: Under- reporting of child mortality

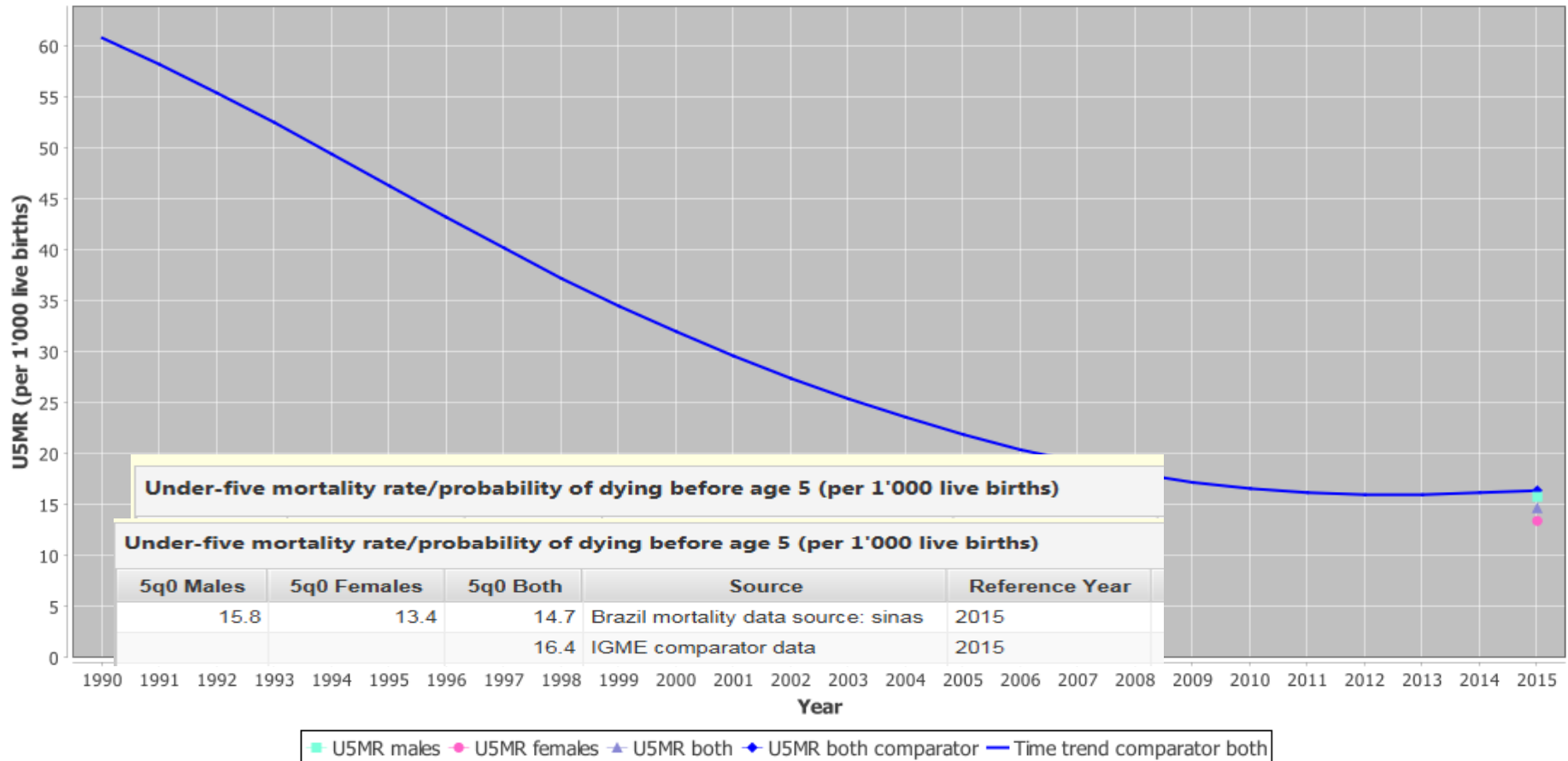
Most common problem with CRVS systems is to undercount child (<5years) deaths.

- Calculates the probability of a newborn dying before age 5 based on reported child deaths
- Compares this to what is considered best estimated of child deaths globally
- Assesses the probable extent of undercount of child deaths by comparing these rates



Child mortality and estimated undercounting (11%?)

Time-trend in estimated under-five mortality rate from comparator data



Brazil, 2015

Summary: key things to look out for when checking mortality data for plausibility of levels/patterns

1. Does the age-distribution of deaths fit expectations for populations with your level of overall mortality? Missing deaths at certain ages? (Step 1)
2. How complete is your registration? – compare your CDR with the CDR trend line provided. Does the calculated completeness fit your expectations? If not why? (Step 2)
3. Does the age-pattern of ASMR look right? - check that $\log(m_x)$ changes linearly with age after about age 35. Does the male/female ratio show male excess mortality, particularly for the ages 15-29?(Step 3)
4. Is the sex-pattern of recorded death rates plausible? - check that Male $m_x >$ Female m_x at all ages. Is the female pattern more skewed towards the older ages? If not, why not? (Step 4)
5. How badly are you missing child deaths? Compare child mortality level with data from best global estimates from censuses/surveys

Time for a mini break?





STEP 6:

Does a first broad assessment of the 'usability' of the COD data

- Tabulates the data by three broad cause groups (to check whether we are under/over-counting broad disease types or injuries?)
- Checks the reliability of this distribution
- Looks at the ratio between these as a measure of epidemiological transition (does this ratio accord with what you think/know about health status in the population??)

Step 6 uses the GBD broad cause of death groups to do a plausibility check on your COD data

Group I:

- Infections & Parasitic diseases (e.g. TB, pneumonia, diarrhoea, malaria, measles)
- Maternal/perinatal causes (e.g. maternal haemorrhage, birth trauma)
- Malnutrition

Group II:

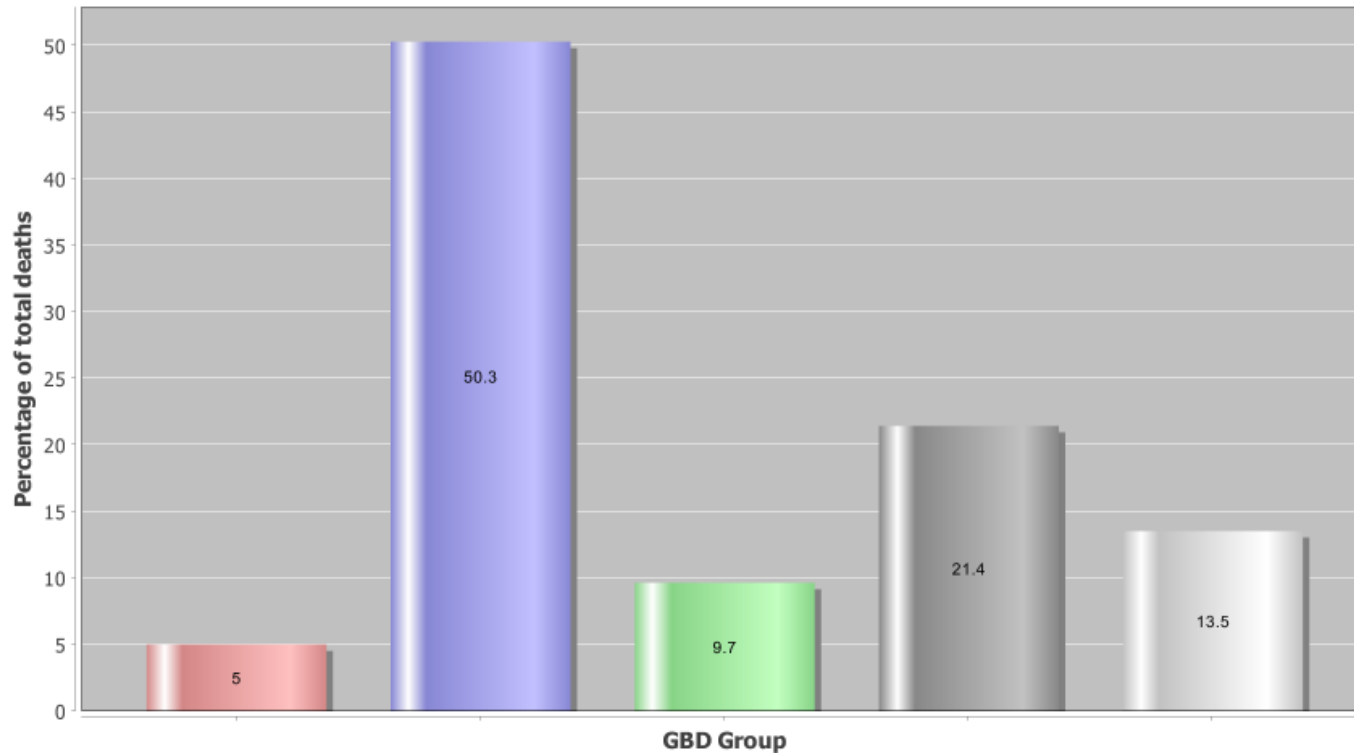
- Non-communicable diseases (e.g. Cancer, diabetes, heart disease, stroke)
- Mental health conditions (e.g. Schizophrenia)

Group III:

- Injuries (e.g. Accidents, homicide, suicide)

Distribution of deaths by broad cause groups (including garbage)

Distribution of deaths by three GBD broad cause groups

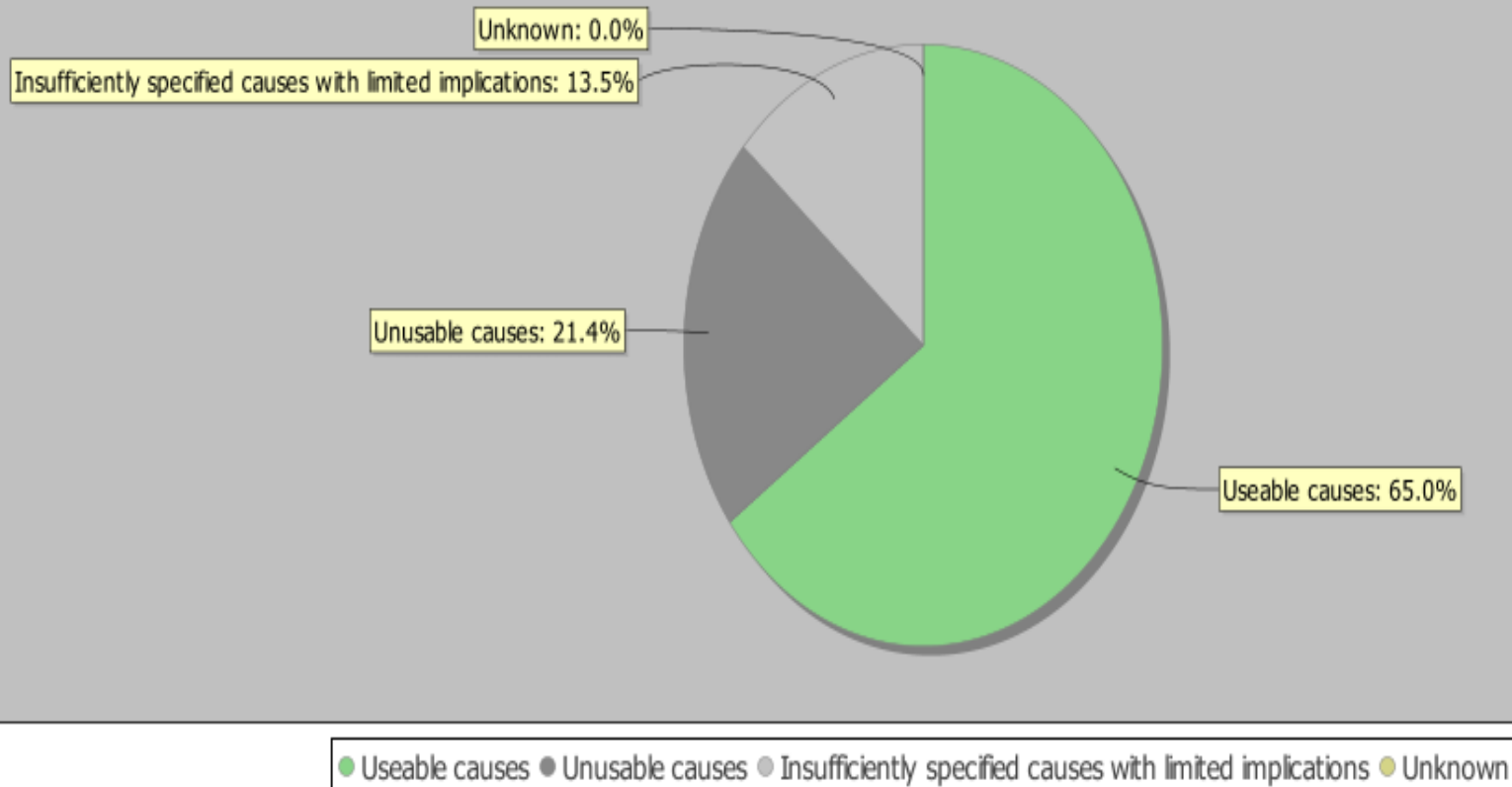


■ Group 1 (Communicable diseases) ■ Group 2 (Non-communicable diseases) ■ Group 3 (External causes) ■ Unusable causes
■ Insufficiently specified causes with limited implications

Brazil, 2015

Overall usability of the COD data

Distribution of deaths by usability



Brazil, 2015



STEP 7: Detailed quality analysis of the COD data

- Breaks down the total unusable causes into different typologies: where, what and how bad?
- How are they distributed by age and sex?
- **Redistributes** the unusable data across the three broad cause groups (what changes?)
- Checks for non-standard ICD codes and biologically implausible causes

Where do we find the unusable codes?

Breakdown by ICD chapter

Total deaths distributed into ICD chapters	
ICD chapter	Description
1	Chapter I: Certain infectious and parasitic diseases
2	Chapter II: Neoplasms
3	Chapter III: Diseases of the blood and blood-forming organs and certain disorders
4	Chapter IV: Endocrine, nutritional and metabolic diseases
5	Chapter V: Mental and behavioural disorders
6	Chapter VI: Diseases of the nervous system
7	Chapter VII: Diseases of the eye and adnexa
8	Chapter VIII: Diseases of the ear and mastoid process
9	Chapter IX: Diseases of the circulatory system
10	Chapter X: Diseases of the respiratory system
11	Chapter XI: Diseases of the digestive system
12	Chapter XII: Diseases of the skin and subcutaneous tissue
13	Chapter XIII: Diseases of the musculoskeletal system and connective tissue
14	Chapter XIV: Diseases of the genitourinary system
15	Chapter XV: Pregnancy, childbirth and the puerperium
16	Chapter XVI: Certain conditions originating in the perinatal period
17	Chapter XVII: Congenital malformations, deformations and chromosomal abnormal
18	Chapter XVIII: Symptoms, signs and abnormal clinical and laboratory findings, not
19	Chapter XIX: Injury, poisoning and certain other consequences of external causes
20	Chapter XX: External causes of morbidity and mortality
21	Chapter XXI: Factors influencing health status and contact with health services
22	Chapter XXII: Codes for special purposes
Total	

% of total unusable causes	
	4.5
	5.1
	0.6
	0.8
	0.7
	1.4
	0.0
	0.0
	35.3
	22.4
	2.3
	0.0
	0.6
	2.2
	0.0
	0.2
	0.1
	17.3
	0.0
	6.5
	0.0
	0.0
	100.0

Five category classification of all non-usable codes based on ICD concepts

- Cat. 1: Symptoms, signs and uncertain diagnoses (Chapter 18 of ICD/R-codes)
- Cat. 2: Impossible as underlying COD
- Cat.3: Intermediate COD
- Cat.4: Immediate COD
- Cat.5: Insufficiently specified COD within ICD chapters

Cat.1: Symptoms & signs and uncertain diagnosis (Chapter 18 of ICD)

- Because ICD is used both for morbidity and mortality coding, it contains many codes that should not be used for mortality coding
- The contents of Chapter 18 (R-codes) are mainly **symptoms and signs** of disease or describe some **health encounter** and should **not** be used as causes of death

Cat.2: “Impossible as underlying causes of death”?

- Causes that you don't die from, e.g. broken leg, tooth and skin disorders, migraines and bipolar disorders, etc.
- When these conditions are used on a death certificate they are considered as impossible and discarded as an UCOD

Cat.3 and 4: Intermediate and immediate causes of death

- A death certificate that reports only the immediate or intermediate COD is useless for public health uses. Why?
- Together they constitute a large group of errors found on death certificates

Cat.5: “Insufficiently specified”. What does it mean?

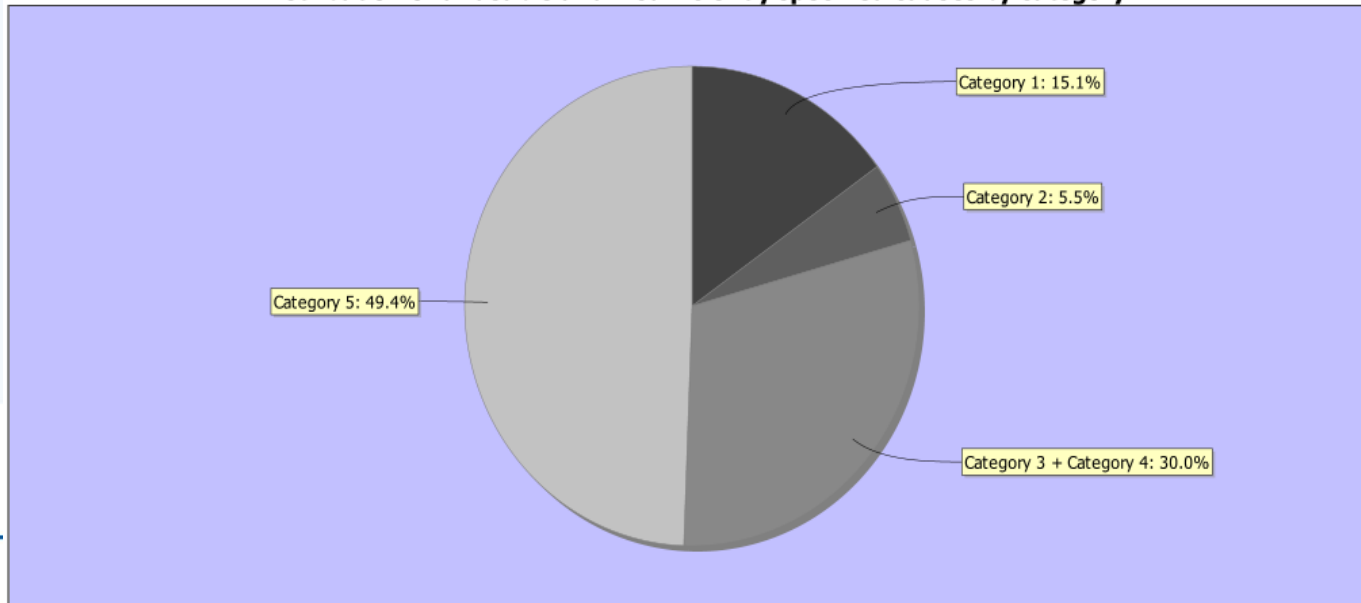
- This group is usually the largest and consists of poorly specified causes, e.g. ill-defined cancers (without the site/malignancy), stroke and pneumonia without type, or unspecified vehicle accidents, injury with undetermined intent
- All these are ***insufficiently specified causes within ICD chapters*** and can be **somewhat informative** for public health decision making
- You may argue that including them into ‘unusable’ is being very strict

Five category classification of all unusable codes in Brazil

Classification of unusable and insufficiently specified causes

Unusable codes classification	No. of deaths with unusable codes	% of total causes	% of total unusable and insufficiently specified causes
Category 1: "Symptoms, signs and ill-defined conditions"	66,697	5.3	15.1
Category 2: "Impossible as underlying causes of death"	24,053	1.9	5.5
Category 3: "Intermediate causes of death"	131,109	10.4	29.7
Category 4: "Immediate causes of death"	1,129	0.1	0.3
Category 5: "Insufficiently specified causes within ICD Chapters"	217,883	17.3	49.4
Total unusable and insufficiently specified causes	440,871	35.0	100.0

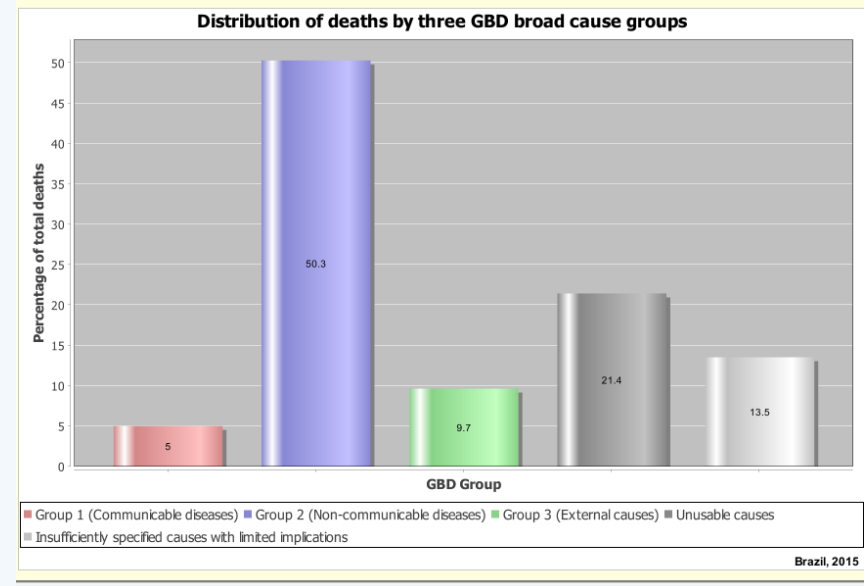
Distribution of unusable and insufficiently specified causes by category



● Category 1 ● Category 2 ● Category 3 + Category 4 ● Category 5

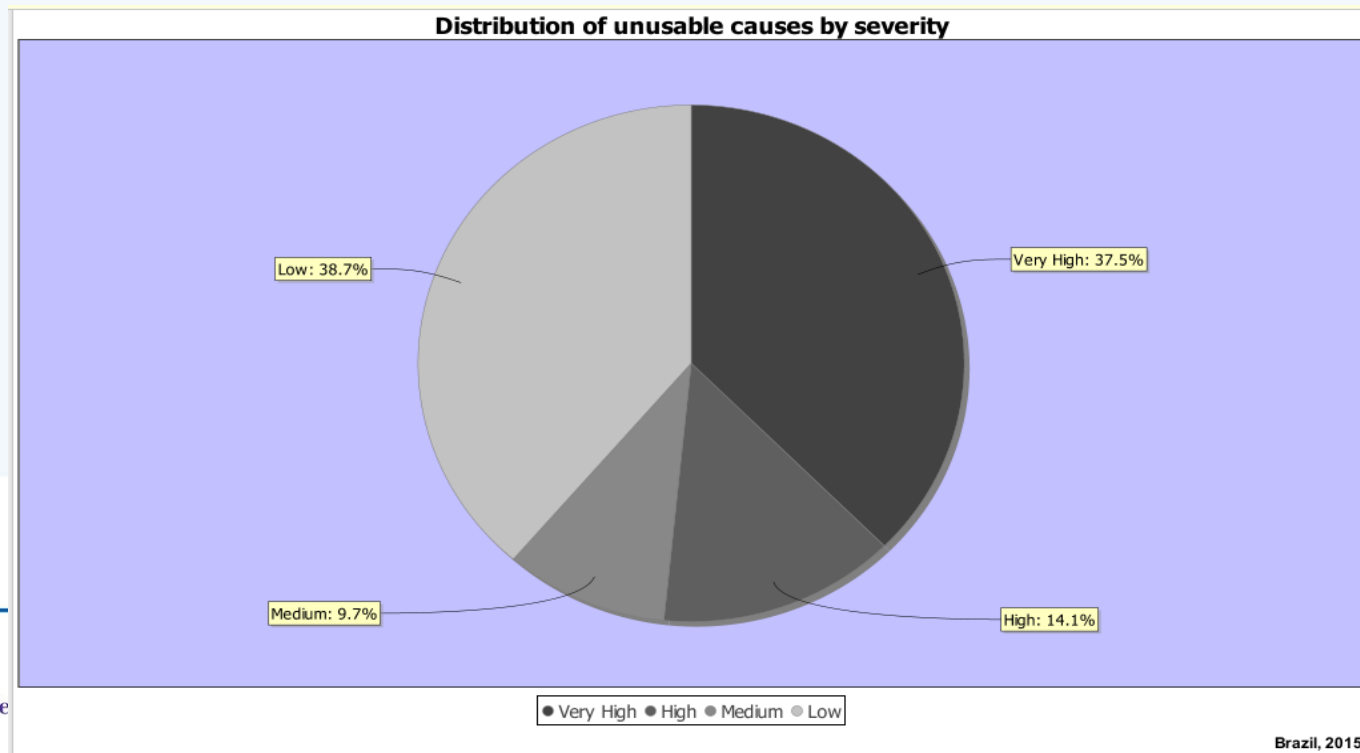
Four category classification of unusable codes based on severity/impact on distribution or policy

- **Level 1:** the true UCOD could be in *more than one* of the three broad COD groups
- **Level 2:** the true UCOD is *within* only one broad cause group (the same as the GC)
- **Level 3:** the true UCOD is *within the same ICD chapter* as the GC
- **Level 4:** the true UCOD is *within the same disease or injury group*



Four category classification of unusable codes by severity for Brazil

Classification of unusable causes by severity			
Unusable codes severity	No. of deaths with unusable codes	% of total causes	% of total unusable and insufficiently specified causes
Very High	165,317	13.1	37.5
High	62,176	4.9	14.1
Medium	42,727	3.4	9.7
Low	170,651	13.5	38.7
Total unusable and insufficiently specified ca...	440,871	35.0	100.0



Further breakdown of each level into packages of similar unusable causes. Example of *Level 1 unusable codes* considered to have “very high” negative impact

Leading packages - severity Very High				Leading ICD-10 codes - severity Very High - package: All, ill defined			
Rank	Package name	Total causes	%	Rank	ICD code	Name of category	Total causes
1	All, ill defined	68,989	^	1	R99.-	Other ill-defined and unspecified causes of mortality	39,582
2	Left HF	29,376		2	R98.-	Unattended death	18,828
3	Sepsis	19,530		3	R68.8	Other specified general symptoms and signs	2,346
4	Renal failure	7,702		4	R96.0	Instantaneous death	1,637
5	Pulmonary embolism	7,057		5	F17.2	Mental and behavioural disorders due to use of tobacco, dependence syndrome	1,358
6	Respiratory failure Acute	4,539		6	M62.3	Immobility syndrome (paraplegic)	740
7	Pneumonitis	4,468		7	R96.1	Death occurring less than 24 hours from onset of symptoms, not otherwise explained	521
8	Senility	3,592		8	F32.9	Depressive episode, unspecified	408
9	Shock & Cardiac arrest	2,658		9	M84.4	Pathological fracture, not elsewhere classified	273
10	Anemia unspecified	2,340		10	G62.9	Polyneuropathy, unspecified	258
11	Peritonitis	2,183		11	F99.-	Mental disorder, not otherwise specified	188
12	Dehydration	1,675		12	M19.9	Arthrosis, unspecified	101
13	cerebral palsy	1,384		13	N42.9	Disorder of prostate, unspecified	74
14	Other and unspecified disorders of fluid, electrolyte and acid-base balance	994		14	M16.9	Coxarthrosis, unspecified	73
15	Brain Anoxia	891		15	M62.8	Other specified disorders of muscle	73
16	hydrocephalus	790		16	F17.1	Mental and behavioural disorders due to use of tobacco, harmful use	64
17	Osteomyelitis	523		17	H66.9	Otitis media, unspecified	62

Example of *Level 4 unusable codes* considered to have “low” negative impact

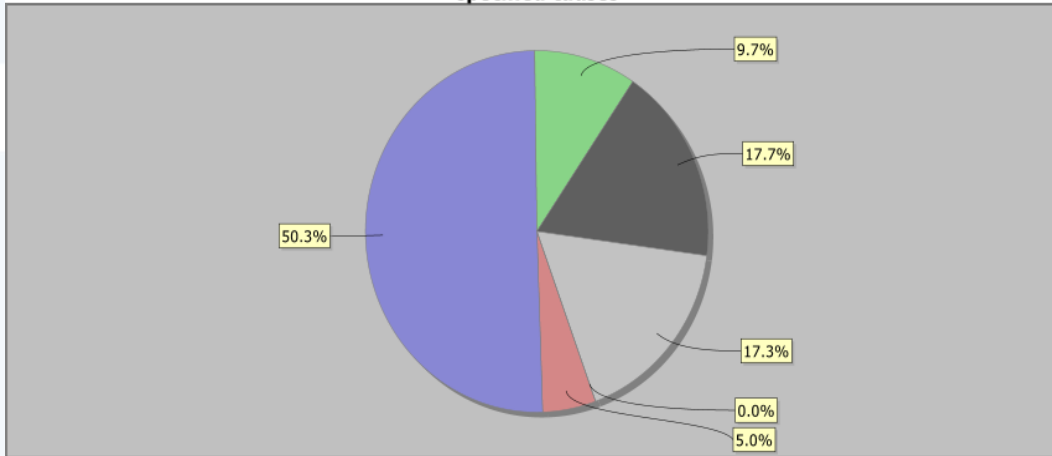
Leading packages - severity Low				Leading ICD-10 codes - severity Low - package: Unspecified acute lower respiratory infectious diseases			
Rank	Package name	Total causes	% of causes in level	Rank	ICD code	Name of category	Total causes
1	Unspecified acute lower respiratory infectious dis...	76,772	44.99	1	J18.9	Pneumonia, unspecified	52,262
2	Unspecified Stroke	71,837	42.10	2	J18.0	Bronchopneumonia, unspecified	14,519
3	Cardiomyopathy	10,078	5.91	3	J15.9	Bacterial pneumonia, unspecified	7,483
4	Unspecified road Inj	7,177	4.21	4	J18.1	Lobar pneumonia, unspecified	965
5	Unspecified Assault	2,125	1.25	5	J18.8	Other pneumonia, organism unspecified	514
6	Unspecified meningitis	791	0.46	6	P23.9	Congenital pneumonia, unspecified	369
7	Unspecified transport Inj	780	0.46	7	J18.2	Hypostatic pneumonia, unspecified	339
8	Unspecified traffic Inj	580	0.34	8	J22.-	Unspecified acute lower respiratory infection	308
9	Myelodysplastic syn	205	0.12	9	P23.8	Congenital pneumonia due to other organisms	8
10	Unspecified pneumoconiosis	159	0.09	10	P23.6	Congenital pneumonia due to other bacterial agents	5
11	Myocardial degeneration	68	0.04				
12	Unspecified Leishmaniasis	53	0.03				
13	Unspecified intestinal parasit	16	0.01				
14	Unspecified malaria	10	0.01				

What are the advantages of this new classification?

- Identifies the comparative importance of four different types of unusable codes according to the impact they can have on a dataset
- Helps to identify and direct focus on those codes with the most harmful impact for misguiding public policy, i.e. 1+2
- Aggregates the universe of unusable codes into 'packages' (164) with similar characteristics/impact
- Within each specific 'package' identifies the most important sources of diagnostic error

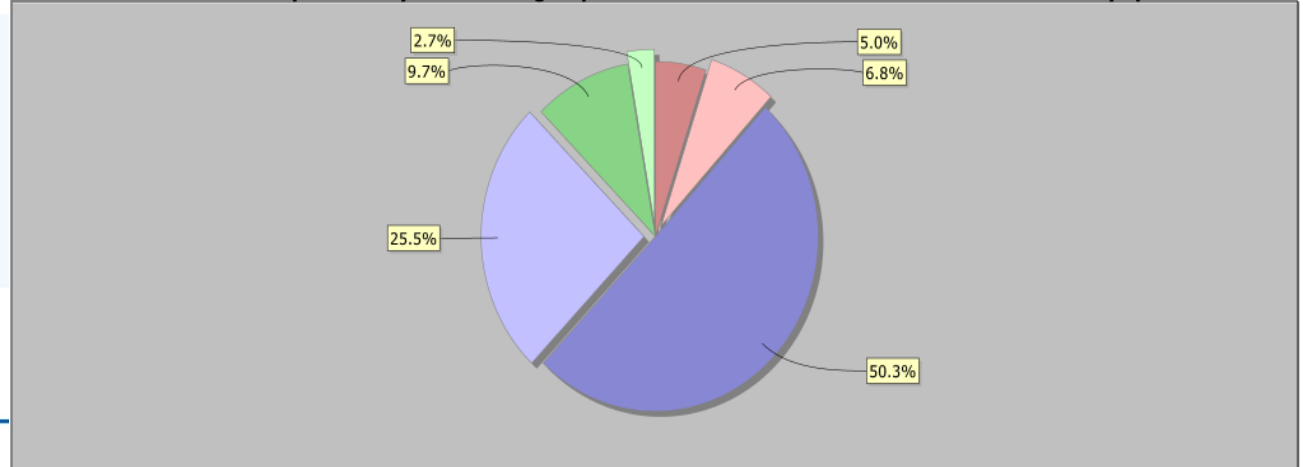
Redistribution of unusable codes (using GBD algorithms) – Brazil 2015

Distribution of input data according to the three broad GBD groups and unusable and insufficiently specified causes



● Group 1 (Communicable diseases) ● Group 2 (Non-communicable diseases) ● Group 3 (External causes) ● Unusable causes
● Insufficiently specified causes ● Unknown

Probable distribution of input data by broad GBD groups after redistribution of unusable and insufficiently specified causes



● Group 1 (Communicable diseases) ● Unusable and insufficiently specified causes redistributed to Group 1 ● Group 2 (Non-communicable diseases)
● Unusable and insufficiently specified causes redistributed to Group 2 ● Group 3 (External causes)
● Unusable and insufficiently specified causes redistributed to Group 3

Checking biological plausibility of cause of death diagnoses. Brazil 2015

Description of cause	Total observations	Unlikely observations	% Unlikely observations	Reason(s) for implausibility
Inflammatory diseases of cervix, vagina, and vulva	66	1	1.5	under 10 years
Malignant bladder cancer	3,918	3	0.1	under 15 years
Malignant breast cancer	15,603	1	0.0	under 15 years
Malignant colon and rectum cancer	16,692	9	0.1	under 15 years
Malignant esophageal cancer	8,376	1	0.0	under 15 years
Malignant larynx cancer	4,408	2	0.0	under 15 years
Liver cancer	9,660	34	0.4	under 15 years
Malignant tracheal, bronchus, and lung cancer	26,593	16	0.1	under 15 years
Malignant skin melanoma	1,789	3	0.2	under 15 years
Malignant skin melanoma (malignant)	28	1	3.6	under 15 years
Malignant lip and oral cavity cancer	3,929	4	0.1	under 15 years
Malignant nasopharynx cancer	313	7	2.2	under 15 years
Non-melanoma skin cancer (squamous-cell carcinoma)	1,985	2	0.1	under 15 years
Malignant other pharynx cancer	2,476	1	0.0	under 15 years
Malignant ovarian cancer	3,546	6	0.2	under 15 years
Malignant pancreatic cancer	9,439	1	0.0	under 15 years
Malignant prostate cancer	14,416	2	0.0	under 15 years
Malignant stomach cancer	14,249	2	0.0	under 15 years
Malignant testicular cancer	357	3	0.8	under 15 years
Neonatal encephalopathy due to birth asphyxia and trauma	4,655	30	0.6	above 10 years
Hemolytic disease and other neonatal jaundice	189	5	2.6	above 10 years
Other neonatal disorders	4,576	7	0.2	above 10 years
Neonatal preterm birth complications	7,708	3	0.0	above 10 years
Neonatal sepsis and other neonatal infections	3,738	8	0.2	above 10 years
	819,030	152	0.0	

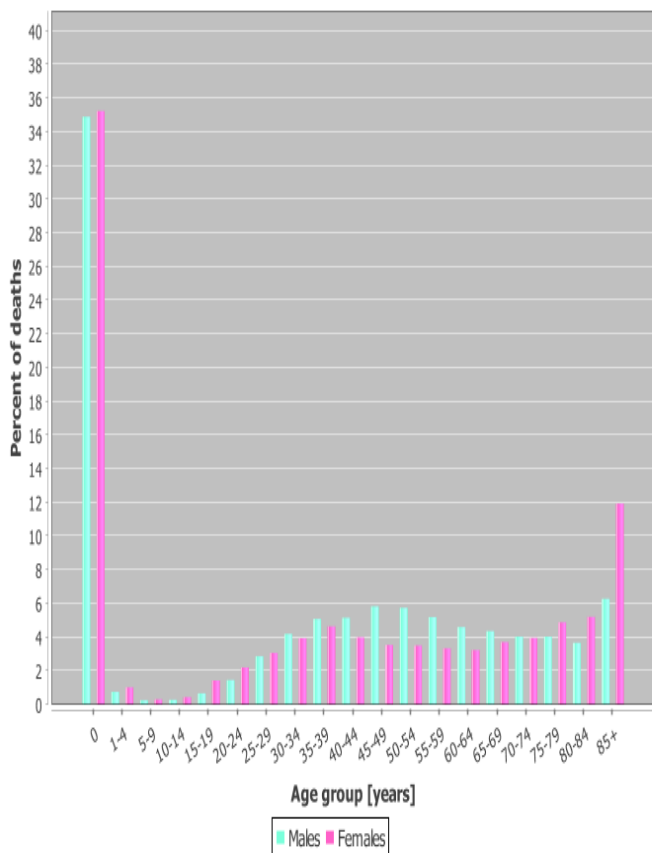


Step 8: Age pattern of Broad Causes of Death

- Based on decades of observations we know that all diseases and injuries generally follow a well established age pattern
- This step checks whether the age pattern looks plausible for the three broad disease groups

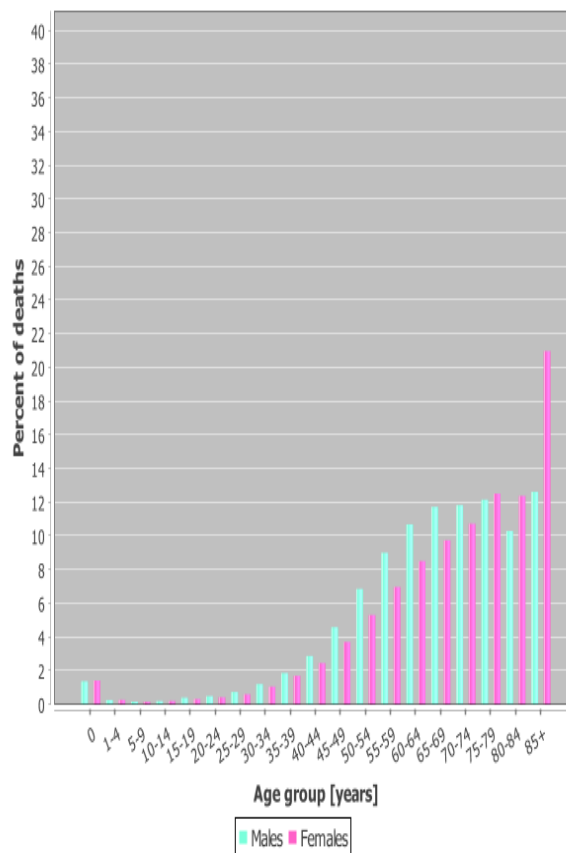
Age pattern of the three broad disease groups

Age pattern of GBD Group 1



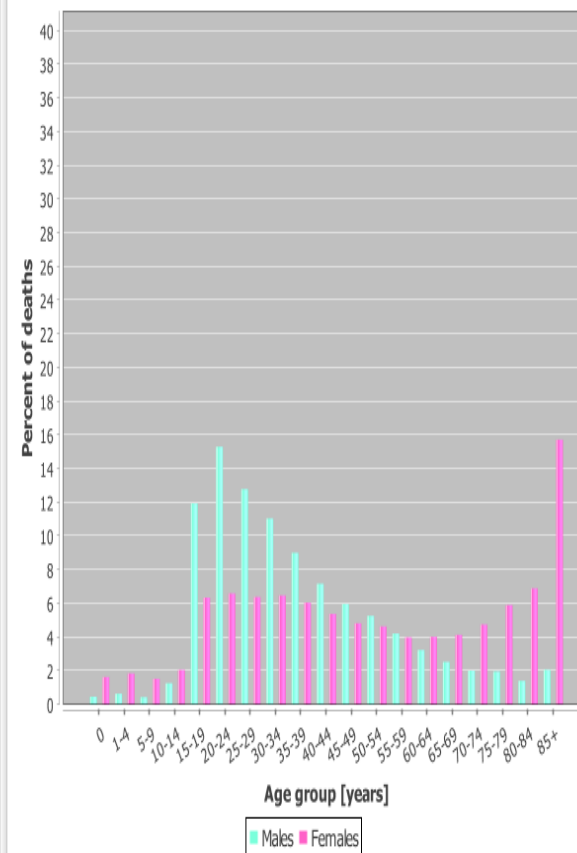
Brazil, 2015

Age pattern of GBD Group 2



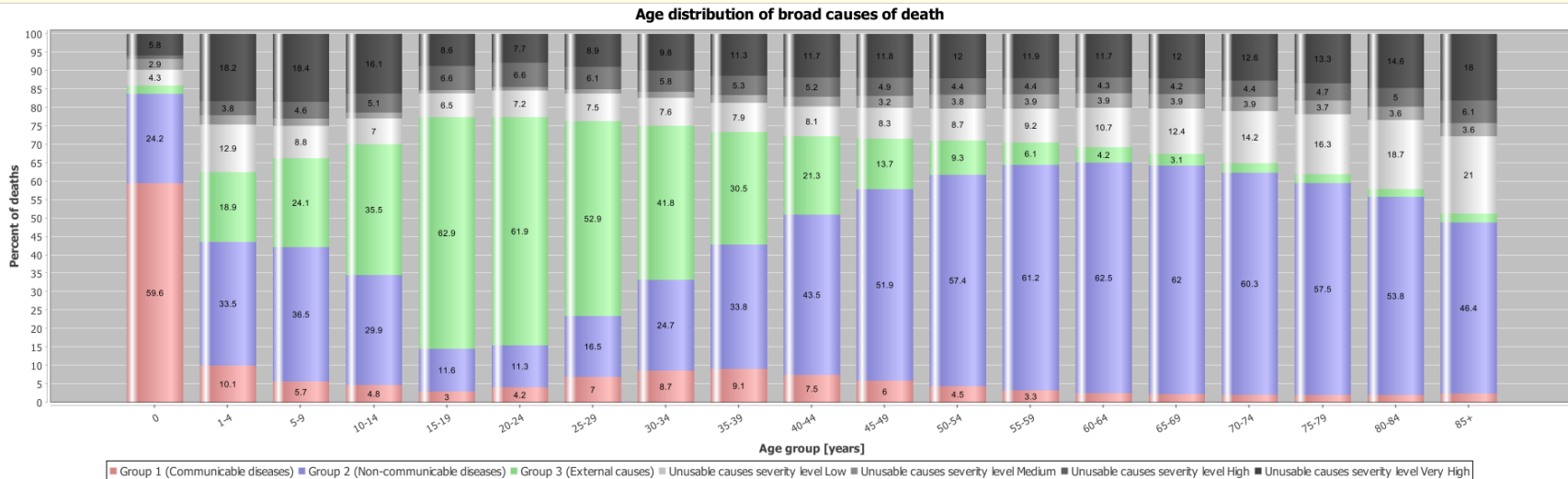
Brazil, 2015

Age pattern of GBD Group 3



Brazil, 2015

Distribution of broad cause of death groups and unusable codes, by age



Brazil, 2015

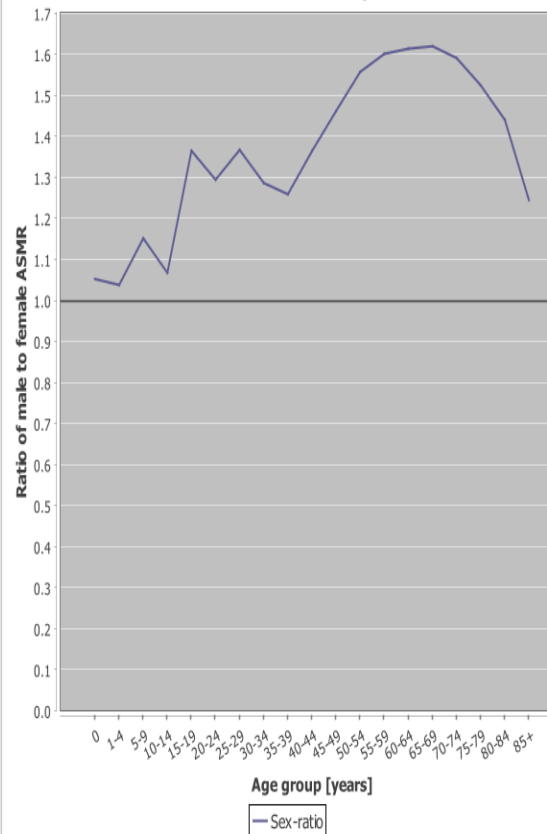
Sex ratio by age for three broad GBD disease groups

Sex ratio of GBD Group 1



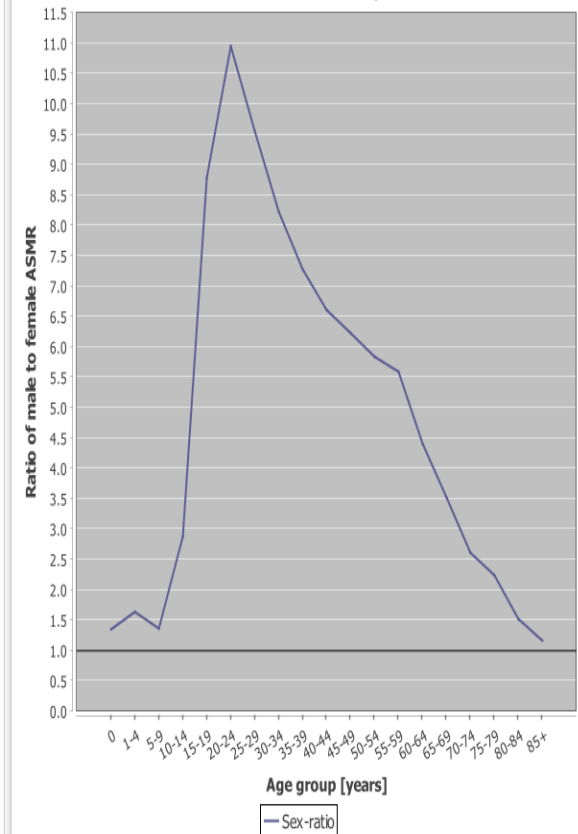
Brazil, 2015

Sex ratio of GBD Group 2



Brazil, 2015

Sex ratio of GBD Group 3



Brazil, 2015



STEP 9: Leading causes of disease

- Policy makers are generally interested in the top 10 or 20 leading causes of death
- This step identifies and ranks the 20 leading causes of death according to your data
- Assesses their reliability/usability
- Compares differences between men and women.

Leading causes of death Brazil 2015

ICD Males				ICD Females			
Rank	% of causes	ICD code	Name of category	Rank	% of causes	ICD code	Name of category
1	7.4	I21.9	Acute myocardial infarction, unspecified	1	6.6	I21.9	Acute myocardial infarction, unspecified
2	3.6	J18.9	Pneumonia, unspecified	2	4.9	J18.9	Pneumonia, unspecified
3	3.4	R99.-	Other ill-defined and unspecified causes of mortality	3	3.6	I64.-	Stroke, not specified as haemorrhage or infarction
4	2.9	I64.-	Stroke, not specified as haemorrhage or infarction	4	2.9	E14.9	Unspecified diabetes mellitus without complications
5	2.7	X95.4	Assault by other and unspecified firearm discharge, street and highway	5	2.8	R99.-	Other ill-defined and unspecified causes of mortality
6	2.1	C34.9	Bronchus or lung, unspecified	6	2.7	C50.9	Breast, unspecified
7	2.0	C61.-	Malignant neoplasm of prostate	7	2.4	I10.-	Essential (primary) hypertension
8	1.7	E14.9	Unspecified diabetes mellitus without complications	8	1.9	G30.9	Alzheimer's disease, unspecified
9	1.5	I10.-	Essential (primary) hypertension	9	1.9	C34.9	Bronchus or lung, unspecified
10	1.5	R98.-	Unattended death	10	1.7	A41.9	Septicaemia, unspecified
11	1.4	J44.9	Chronic obstructive pulmonary disease, unspecified	11	1.6	N39.0	Urinary tract infection, site not specified
12	1.4	J44.0	Chronic obstructive pulmonary disease with acute lower respiratory infection	12	1.5	R98.-	Unattended death
13	1.3	X95.9	Assault by other and unspecified firearm discharge, unspecified place	13	1.5	J44.0	Chronic obstructive pulmonary disease with acute lower respiratory infection
14	1.2	A41.9	Septicaemia, unspecified	14	1.4	I67.8	Other specified cerebrovascular diseases
15	1.2	C16.9	Stomach, unspecified	15	1.4	J44.9	Chronic obstructive pulmonary disease, unspecified
16	1.1	I69.4	Sequelae of stroke, not specified as haemorrhage or infarction	16	1.3	J18.0	Bronchopneumonia, unspecified
17	1.0	I67.8	Other specified cerebrovascular diseases	17	1.3	I69.4	Sequelae of stroke, not specified as haemorrhage or infarction
18	1.0	I61.9	Intracerebral haemorrhage, unspecified	18	1.3	I50.0	Congestive heart failure
19	1.0	J18.0	Bronchopneumonia, unspecified	19	1.2	I11.0	Hypertensive heart disease with (congestive) heart failure
20	1.0	K74.6	Other and unspecified cirrhosis of liver	20	1.2	I61.9	Intracerebral haemorrhage, unspecified



STEP 10: Vital Statistics Performance Index

- Measures the overall quality of the mortality output of the CRVS system in one composite indicator
- Provides a summary indicator you can use for monitoring of whether your CRVS data are improving

Classification of global CRVS systems based on VSPI

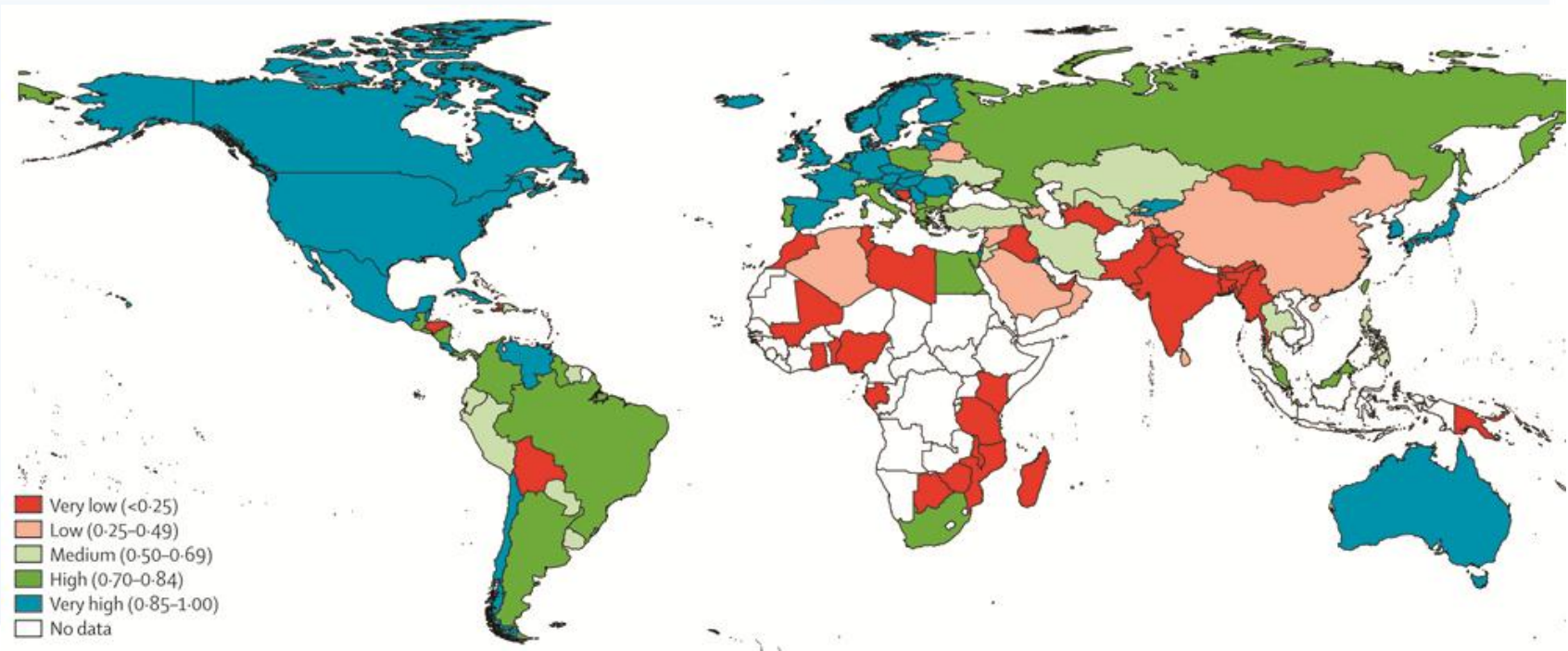


Figure 1: Typology of CRVS systems based on the VSPI scores, best possible available year, 2005 -12

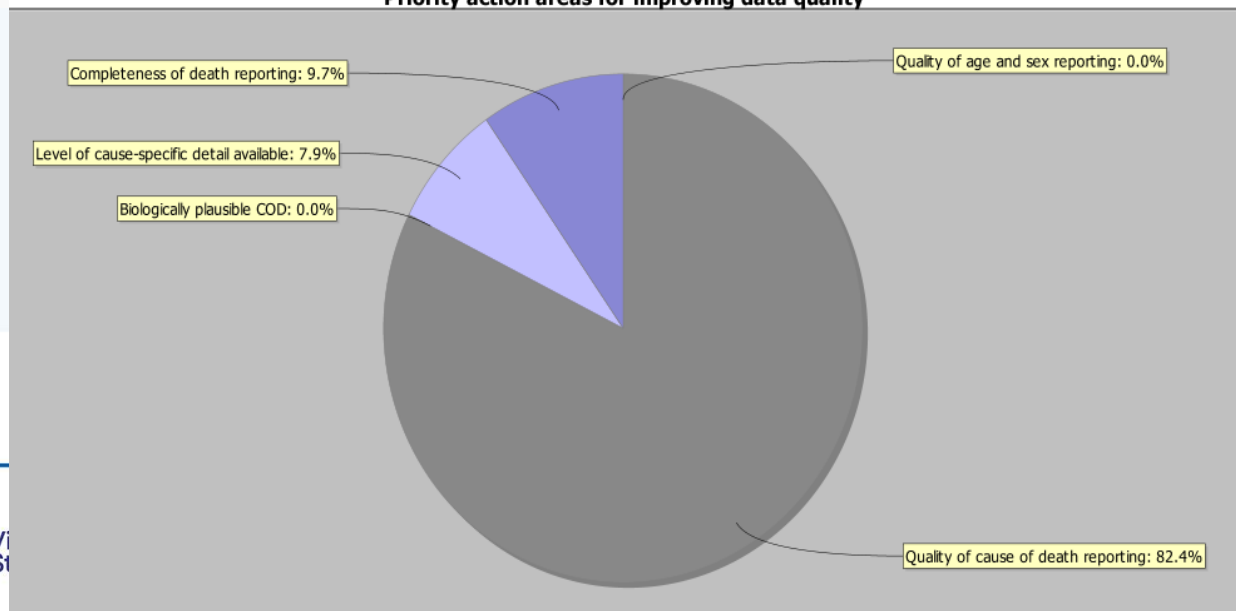
Relative contribution of each dimension to gap (vs 100%) in VSPI(Q)

Country: BRA (Brazil)
Region: 7 (Latin America/Caribbean)
Summary Score: 78.1%
Classification: High

VSPI Quality Components Score

Component	Score (weighted)
Quality of age and sex reporting	100.0
Quality of cause of death reporting	81.3
Biologically plausible COD	100.0
Level of cause-specific detail available	98.2
Completeness of death reporting	97.8

Priority action areas for improving data quality



The VSPI(Q) Mikkelsen et al, Lancet 386: 1395–406 2015

The VSPI(Q) index is classified into one of five categories):

- Very high (0.85-1.0)
- High (0.70-0.84)
- Medium (0.50-0.69)
- Low (0.25-0.49)
- Very low(<0.25)



THE UNIVERSITY OF
MELBOURNE

Bloomberg
Philanthropies  DATA FOR
HEALTH INITIATIVE

