ORIGINAL ARTICLE

Assessing the quality of cause of death data in six high-income countries: Australia, Canada, Denmark, Germany, Japan and Switzerland

Lene Mikkelsen[1] · Kim Moesgaard Iburg[2] · Tim Adair[1] · Thomas Fürst[3] · Michael Hegnauer[3] ·
Elena von der Lippe[4] · Lauren Moran[5] · Shuhei Nomura[6] · Haruka Sakamoto[6] · Kenji Shibuya[7] ·
Annelene Wengler[4] · Stephanie Willbond[8] · Patricia Wood[8] · Alan D. Lopez[1]

Abstract

Objectives To assess the policy utility of national cause of death (COD) data of six high-income countries with highly developed health information systems.

Methods National COD data sets from Australia, Canada, Denmark, Germany, Japan and Switzerland for 2015 or 2016 were assessed by applying the ANACONDA software tool. Levels, patterns and distributions of unusable and insufficiently specified "garbage" codes were analysed.

Results The average proportion of unusable COD was 18% across the six countries, ranging from 14% in Australia and Canada to 25% in Japan. Insufficiently specified codes accounted for a further 8% of deaths, on average, varying from 6% in Switzerland to 11% in Japan. The most commonly used garbage codes were Other ill-defined and unspecified deaths (R99), Heart failure (I50.9) and Senility (R54).

Conclusions COD certification errors are common, even in countries with very advanced health information systems, greatly reducing the policy value of mortality data. All countries should routinely provide certification training for hospital interns and raise awareness among doctors of their public health responsibility to certify deaths correctly and usefully for public health policy.

Keywords Causes of death · Medical certification · Data quality · Garbage codes · Assessment of data

Introduction

The Global Burden of Disease (GBD) studies have documented remarkable improvements in health that have occurred during the last decade, but also how unevenly health outcomes are distributed between and within populations. A recent study has raised concerns that adult death rates for many diseases have plateaued and, in some cases, increased, including in high-income countries (Roth et al. 2018). Subnational data have demonstrated surprising health inequality in some countries with well-developed health systems (AIHW 2018; Chammartin et al. 2016;

Mahapatra et al. 2007; Roth and Dwyer-Lindgren 2017). To be able to monitor such health trends and the impact of interventions accurately, valid, reliable, regular and up-to-date national mortality and morbidity data are essential (Lopez 2013; Shibuya 2006).

The global health goals and accountability for their achievements have led to significant interest in monitoring data quality and to the development of summary indicators such as the Vital Statistics Performance Index (VSPI) which measures the quality and timeliness of available mortality data (Mikkelsen et al. 2015; Philips et al. 2014). However, to be able to determine what actions need to be taken to improve statistical outputs, a more comprehensive review of the data is needed to better understand the main data quality issues and their origins. The usual research methodology to assess the accuracy of causes of death (COD) is to undertake an independent review of a sample of medical records and compare the records with the

---

Extended author information available on the last page of the article

cause(s) written on the death certificate (Alpérovitch et al. 2009), or to compare the clinical COD to an autopsy-based COD (Schdev 2001). Such studies, however, are complex and expensive to carry out and usually conducted in only one or a handful of hospitals. The findings have often indicated that even in countries with well-functioning civil registration and vital statistics (CRVS) systems, the quality of the medical certification is not as good as might be expected (Rampatige et al. 2014; Adair et al. 2019). Accuracy in COD certification is likely to be a growing issue in countries experiencing significant population ageing, in particular related to dementia and multiple chronic conditions that make accurate and consistent certification more challenging (Naghavi et al. 2010).

Within this context, it is crucial to be able to understand how well national mortality data systems of high-income countries are performing, given the expectations for them. To address this issue, we assessed the certification specificity and policy utility of the national COD data from six high-income countries with highly developed health information systems: Australia, Canada, Denmark, Germany, Japan and Switzerland.

Methods

Medical certification of death is a requirement in all the countries included in this study. Certifiers, mostly physicians and special health care providers, are asked to complete a medical death certificate indicating what was, in their opinion, the sequence of morbid events leading to death. Subsequently, the information provided on the death certificate is coded by trained coders, applying the International Statistical Classification of Diseases and Related Health Problems (ICD) and its rules for COD coding, and compiled by health or statistical authorities (WHO 2016).

National data on population and COD coded to ICD-10, by age group and sex, were provided by Australia, Canada and Germany for 2016 and downloaded from the World Health Organization (WHO) website for Denmark, Japan and Switzerland for 2015 (WHO 2019). To evaluate the data sets, the ANACONDA tool (see Annex 1, Mikkelsen et al. 2020) that assesses the quality of national and subnational COD data was used.

The accuracy of the COD output is dependent on physicians providing enough information on the death certificate for coders to select and code the underlying COD. When this process is not completed correctly, the COD output may contain codes labelled "garbage" codes (Murray and Lopez 1996) because they are of little or no use for policy decision-making. Historically, "garbage" codes are defined as causes that cannot or should not be an underlying cause of death (Naghavi et al. 2010). The term,

despite its inelegance, has now been an integral part of the literature for more than a quarter of a century. Since ANACONDA uses the same concept and definition for these codes, we have retained the original terminology in the paper.

To provide additional insight into the provenance and policy implications of these codes, ANACONDA classifies garbage codes into two distinct typologies. In the first typology, garbage codes are grouped into five categories based on ICD concepts:

- Category 1: Codes relating to symptoms, signs and ill-defined conditions (most drawn from ICD Chapter XVIII); e.g. R99 Other ill-defined and unspecified causes of mortality).
- Category 2: Codes that are not valid as an underlying cause of death (e.g. T12 Fracture of lower limb).
- Category 3: Codes that represent intermediate causes of death (e.g. I50 Heart failure).
- Category 4: Codes that represent immediate causes of death (e.g. I46 Cardiac arrest).
- Category 5: Codes that represent insufficiently specified causes within ICD chapters or within a larger disease category (e.g. D48.9 Neoplasm of uncertain or unknown behaviour, unspecified).

ANACONDA also includes a second typology, which focuses much more on the potential impact that garbage codes might have on misguiding policy and planning (Naghavi 2020). In this typology, garbage codes are grouped into four impact levels, from "very high" (level 1) to "low" (level 4):

- Very high (level 1): This highest level represents causes for which the true underlying cause could be a communicable or non-communicable disease, or the result of an injury (e.g. septicaemia).
- High (level 2): These are causes with substantial negative impact, but where the true cause is mostly limited to one of the three broad cause groups mentioned above, e.g. essential (primary) hypertension that can be due to different non-communicable diseases.
- Medium (level 3): CODs classified to the third level are only considered to have a medium negative impact for policy since, in this case, the underlying cause is likely to be within the same ICD chapter (e.g. unspecified cancer).
- Low (level 4): Causes classified as having low negative impact are those where the true underlying cause is likely to be confined to a single disease of injury group, such as unspecified stroke or unspecified pneumonia.

These "impact-level" categories can be further grouped into those that provide no or little useful information about the true underlying cause (levels 1–3), which we therefore

refer to as "unusable", and those in level 4 that provide sufficient information to guide public health interventions but not for research and technology development (Naghavi et al. 2010). We refer to the latter as "insufficiently specified" causes as they impair evidence-based health policy processes only to a limited extent. However, correcting these becomes increasingly important if our health information systems are to appropriately guide research, hospital financial flows, resource allocation and healthcare strategies (WHO 2019). This is likely to be particularly of relevance in countries with ageing populations where most deaths happen in hospitals, primarily from non-communicable diseases.

In countries where unusable codes are assigned to a large proportion of all deaths, the true COD distribution can be seriously distorted and thereby mislead policy dialogue. This is particularly serious when garbage codes are common among the leading causes of death. From the input data, ANACONDA automatically provides a listing of the top-20 COD for males and females and indicates those that are considered to be unusable (levels 1–3) or insufficiently specified (level 4). The higher the number of these codes and the higher their ranking, the greater their impact on misinforming policy is going to be.

The relationship between age and garbage codes is also investigated with the ANACONDA tool to verify whether they are particular to certain age groups. Furthermore, given that some differences might exist in population age structure between the six countries, we used the global proportion of deaths by age from the latest Global Burden of Disease Study as the standard (Murray et al. 2018) to age-standardize the garbage codes in the countries.

Data completeness, a key indicator of data quality, was not considered, given that all six countries have civil registration systems that register all deaths. The focus of our data quality analysis therefore was limited to the levels, patterns and distribution of garbage codes.

Results

Despite the six countries being from three different geographic regions—Europe, Asia-Pacific and North America—their health systems and socio-economic indicators are comparable (Annex 2 Table 1S). Life expectancy varies from 78 to 81 years for males and 83 to 87 years for females, with all having very low child mortality rates of 2–5 per 1000 live born. The total fertility rates and proportion of 65 years and above indicate that Australia and Canada have somewhat younger populations than Denmark, Germany, Japan and Switzerland. Switzerland, with a private health insurance system, spends significantly more money on health care per person than the other countries. The Socio-Demographic Index (SDI) (Wang et al. 2016), a measure of national development based on income, education and fertility, is high for all countries, especially Denmark, while Germany and Denmark are doing slightly less well than the others on the Health Access and Quality Index (HAQ). The VSPI(Q), a measure of the overall quality of mortality data calculated by ANACONDA, is the highest in Australia and the lowest in Japan.

In the six countries studied, the average proportion of unusable codes (levels 1–3) was 18%, being slightly lower (14%) in Australia and Canada, while higher in Japan, where one in four deaths is assigned an unusable cause (Table 1). Insufficiently specified codes (level 4), in addition, averaged 8%, varying from 6% in Switzerland to 11% in Japan. Three of the most common CODs in the insufficiently specified group are pneumonia, stroke and diabetes all unspecified. For these CODs, the certifier could have increased the utility of the information provided on the medical certificate of death by specifying whether the pneumonia was bacterial or viral, the stroke ischaemic or haemorrhagic and the diabetes type 1 or 2.

Given the highly developed status of the six countries, the distribution of deaths on the three broad GBD groups of health conditions as expected showed that communicable and maternal diseases as well as injuries are minor contributors to their disease burden (Table 1). Non-communicable diseases on average accounted for 67.8% of all causes of death. Only Japan showed an unlikely low proportion (58.5%) that points to the impact that garbage codes can have on the cause pattern of mortality.

Since most deaths in these countries occur at older ages, it might be expected that these age groups also account for most of the garbage codes. That is indeed the case; in all six countries, between 85 and 92% of the garbage codes occur at ages 65 years and over. However, garbage codes are not limited to the oldest ages; they also comprise a sizeable proportion of deaths in several other age groups, particularly in the younger adult age groups where they constitute between 20 and 30% of all deaths. Even for child deaths, we do not know the true underlying cause in 10% of cases (Fig. 1). Deaths at these ages are often entirely preventable, but to do so public policy must be guided by accurate and specific COD data and how they are changing.

The first ANACONDA typology classifies the total amount of garbage codes according to five categories of certification errors and shows the percentage of each in relation to total deaths, and as a percentage of the total number of garbage codes (Table 2). In all countries except Germany, insufficiently specified COD (Category 5) was the most common error. The reporting of intermediary instead of underlying COD (Category 3) was the second most frequent issue in Australia, Canada, Japan and

Table 1 Total number of deaths and percentage of death by three Global Burden of Disease broad cause groups[a] and unusable and insufficiently specified causes, Australia, Canada, Denmark, Germany, Japan and Switzerland, 2015–2016

Country	Total number of deaths	Group 1% communicable	Group 2% NCDs	Group 3% external causes	% Unusable causes	% Insufficiently specified causes	Total
Australia	158.504	1.3	71.1	5.0	14.1	8.2	100
Canada	267.213	1.4	72.1	4.5	14.3	7.6	100
Denmark	52.224	1.9	66.9	2.9	20.0	8.1	100
Germany	910.902	1.1	69.6	3.6	18.7	7.0	100
Japan	1290.444	1.2	58.5	4.5	25.0	10.7	100
Switzerland	67.606	1.6	68.7	5.1	18.3	6.2	100
Mean		1.4	67.8	4.3	18.4	8.0	100

[a]Group I: Infections and 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) often referred to as "external diseases" or "non-natural" cause of death

Switzerland. In Denmark, the second most frequent reporting flaw was Category 1 (the reporting of signs, symptoms or other ill-defined COD), followed by the reporting of an intermediary COD (Category 3). It is to be expected that all countries assign some deaths to ICD-10 code R99 (Other ill-defined and unspecified causes of mortality) since there always will be deaths for which the cause was unknown. However, Denmark and Japan stand out by coding more than 7% of all deaths to category 1, which is much higher than in any of the other countries. Further in-depth analysis revealed the reason in Japan being because of using Senility (R54) as a COD, while in Denmark, it is indeed the frequent use of the code R99.

On average, certifiers in the six countries reported an intermediary COD (Category 3) as being the underlying COD for 9% of deaths. This error was particularly common in Japan (13%) and Germany (11%), almost twice as high as in the other countries. Further investigation showed that "Heart failure, unspecified" (I50.9) and "Congestive heart failure" (I50.0) were used more frequently in Japan and Germany than in other countries. Regarding the two remaining categories, irrespective of country, very few doctors certified an impossible COD (Category 2) or just provided the immediate COD (Category 4).

The second typology of garbage codes provides important insight into the potential impact that garbage codes might have in guiding or misguiding public policy. This categorization showed a similar pattern for all countries with the "very high" impact category being the biggest problem for all six countries, followed by the "low" impact category, except for Australia, where the order of these top two impact categories was inverted in comparison with all other countries (Table 3). However, of note, the percentage at the "very high" impact level showed substantial differences between countries. For instance, in Japan, 21% of all deaths and 58% of all garbage codes had a very high impact for policy, while in Australia the comparable figures were only 8% and 35%, respectively. A closer investigation of the specific codes revealed that the three ICD codes that account for most of the "very high" impact garbage codes were "Other ill-defined and unspecified deaths" (R99), "Heart failure" (I50.9) and "Senility" (R54).

The high and medium levels typically only accounted each for 2–4% of all deaths in all countries. The most common misdiagnosis for the high level was Essential (primary) hypertension (I10), Unspecified external factor (X59) and Gastrointestinal bleeding (K92.2), while for the medium level it was Unspecified cancer (C80.9) for all. The low-impact garbage codes were generally between 6% (Switzerland) and 11% (Japan) of all deaths and, on average, were used for 8% of all deaths, accounting for 31% of all garbage codes. In the low-impact group, the biggest contributor was "Stroke not specified" (I64), except for Japan where certifiers seem to better distinguish between haemorrhagic or infarction stroke, but not between the different types of pneumonia).

Standardizing the age structure for each country had a minor impact on the proportions of garbage codes for five of the countries. Only in Japan, where a higher proportion of deaths occur at the very oldest ages, age standardization reduced the amount of total garbage codes by 6%. However, Japan still remained the country with the highest proportion of deaths assigned to garbage codes (Table 3).

The unusable codes among the leading causes of death are identified by red cells in Table 4. All countries, except Australia, had at least one cell with unusable codes (red) among the top 10 causes of male deaths and all had one or

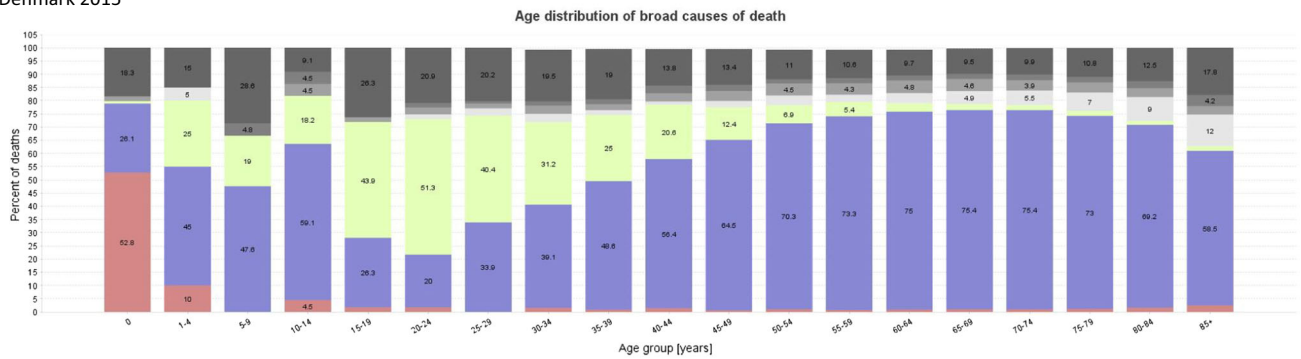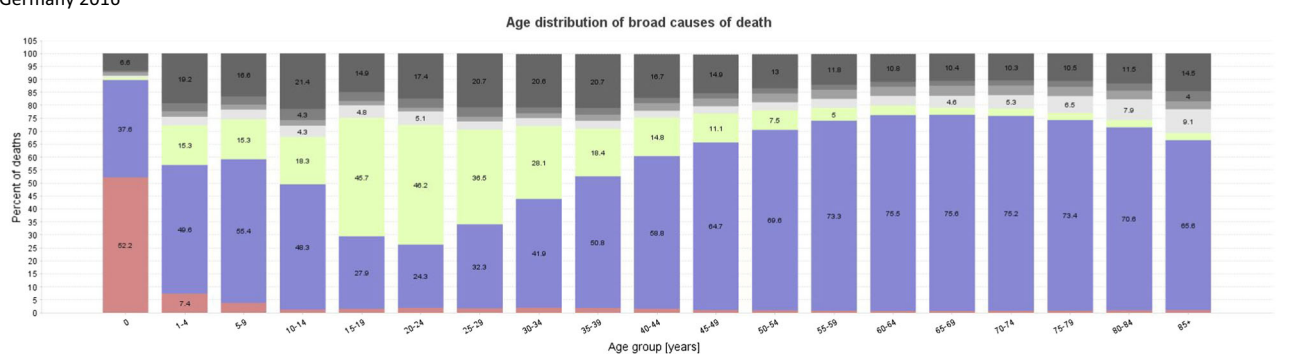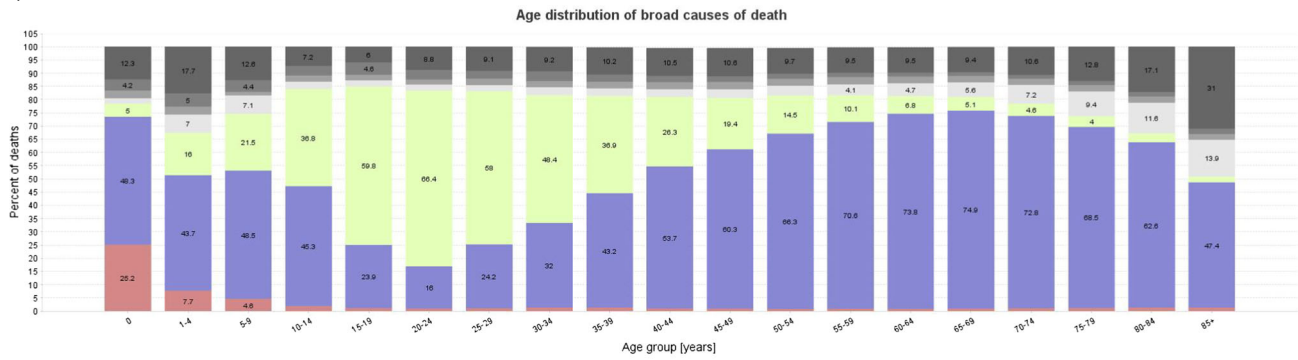Australia 2016
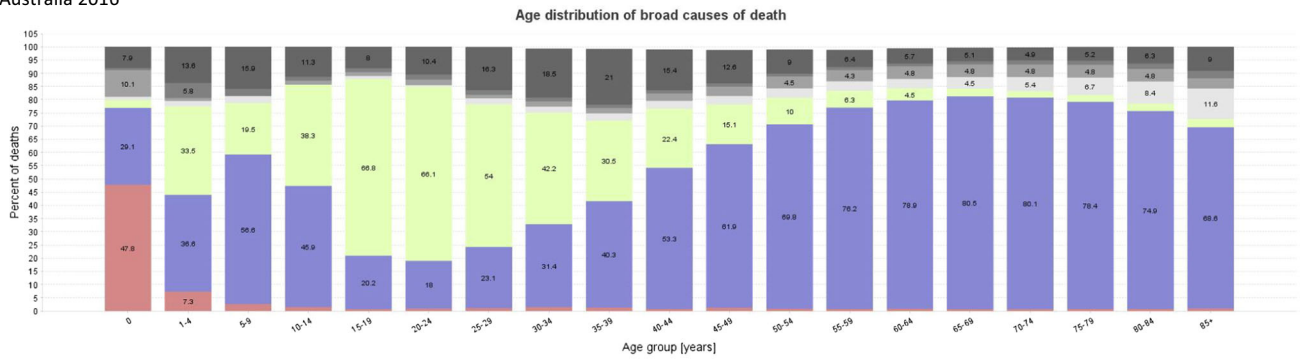
Canada 2016
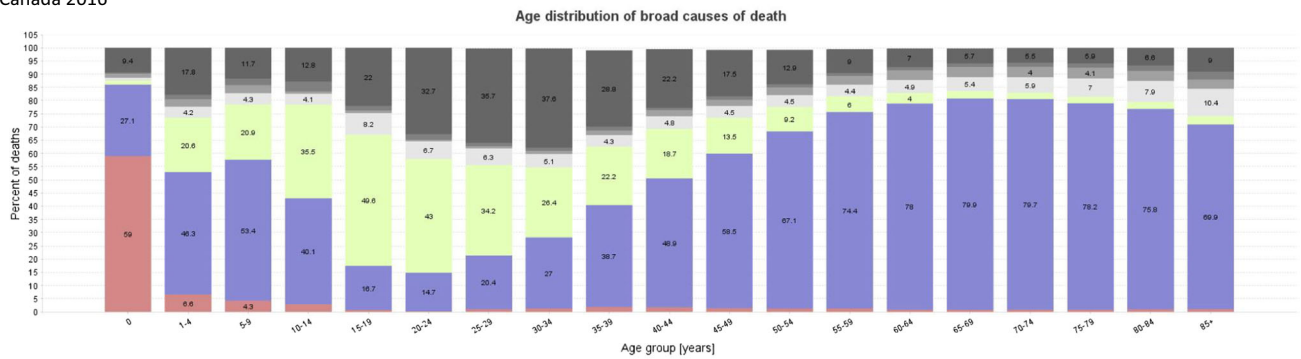
Denmark 2015

Germany 2016

Fig. 1 Age distribution of deaths on broad Global Burden of Disease groups and garbage causes for six countries: Australia, Canada, Denmark, Germany, Japan and Switzerland, 2015–2016
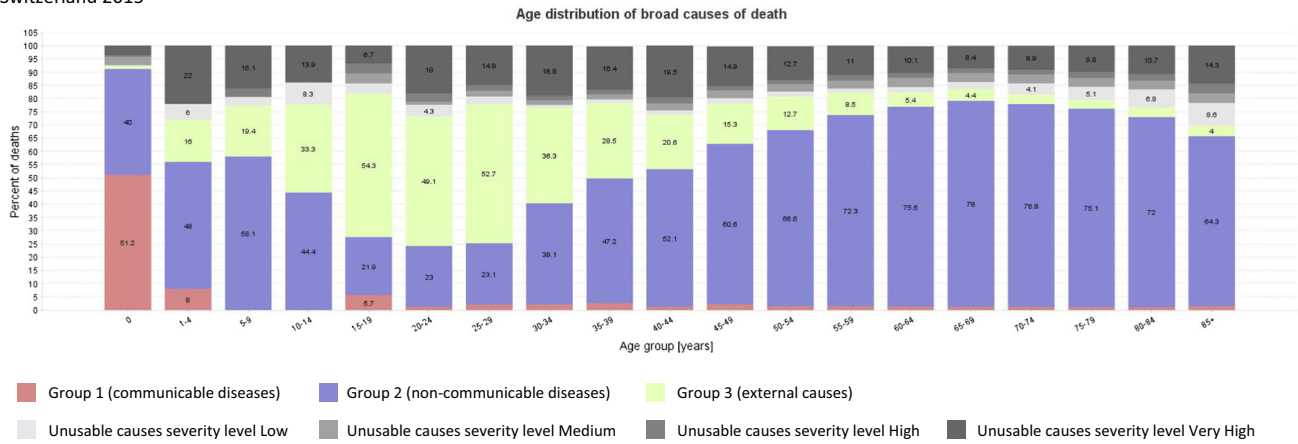
Japan 2015

Switzerland 2015

Fig. 1 continued

Table 2 Number of deaths with a garbage code and % of different garbage types as (i) % of all deaths and (ii) % of deaths with a garbage code (in brackets), Australia, Canada, Denmark, Germany, Japan and Switzerland, 2015–2016

Country	Number of deaths with a garbage code	Category 1 signs, symptoms, ill-defined COD	Category 2 impossible COD	Category 3 intermediate COD	Category 4 immediate COD	Category 5 insufficiently specified COD	Per cent of total garbage codes
Australia 2016	34.829	0.9 (4.3)	1.1 (5.2)	6.6 (30.0)	0.1 (0.7)	13.1 (59.8)	22.0 (100)
Canada 2016	58.503	1.8 (8.2)	1.2 (5.4)	6.6 (30.2)	0.2 (0.9)	12.1 (55.3)	21.9 (100)
Denmark 2015	14.672	7.6 (27.2)	1.3 (4.7)	6.3 (22.4)	0.3 (1.2)	12.5 (44.5)	28.1 (100)
Germany 2016	234.254	3.1 (12.0)	1.1 (4.1)	11.2 (43.6)	0.4 (1.6)	10.0 (38.7)	25.7 (100)
Japan 2015	459.913	7.7 (21.7)	1.1 (3.0)	12.9 (36.2)	1.0 (2.7)	13.0 (36.4)	35.6 (100)
Switzerland 2015	16.601	4.2 (16.9)	1.9 (7.6)	7.8 (31.7)	1.3 (5.1)	9.5 (38.8)	24.6 (100)
Mean		4.2 (15.1)	1.3 (5.0)	8.6 (32.4)	0.6 (2.0)	11.7 (45.6)	26.3 (100)

more among the top 20 causes (Table 4). Japan had four red cells, and three of these were among the top 10 causes. Denmark and Canada had one red cell, while Switzerland and Germany had two red cells in the top half of the ranking. The specific unusable causes were very similar across the countries and included Other ill-defined and unspecified deaths (R99), Unspecified heart failure (I50.9), Congestive heart failure (I50.0), Unspecified cardiac arrest

Table 3 Number of deaths with garbage codes, classified by severity of impact, provided as (i) % of total deaths and (ii) % of deaths with a garbage code (in brackets), Australia, Canada, Denmark, Germany, Japan and Switzerland, 2015–2016

Country	Number of deaths with a garbage code	Very high	High	Medium	Low	Total % garbage	Age-adjusted total % garbage
Australia 2016	34.829	7.7 (34.9)	2.0 (9.1)	4.3 (19.5)	8.0 (36.4)	22.0 (100)	21.0
Canada 2016	58.503	8.7 (39.6)	2. 0 (9.4)	3.5 (16.2)	7.6 (34.9)	21.9 (100)	20.2
Denmark 2015	14.672	13.5 (47.9)	2.9 (10.2)	3.7 (13.2)	8.1 (28.7)	28.1 (100)	26.1
Germany 2016	234.254	12.5 (48.5)	2.9 (11.3)	3.3 (12.9)	7.0 (27.3)	25.7 (100)	23.7
Japan 2015	459.913	20.8 (58.4)	1.8 (5.0)	2.3 (6.6)	10.7 (30.0)	35.6 (100)	29.3
Switzerland 2015	16.601	12.1 (49.3)	2.8 (11.3)	3.5 (14.1)	6.2 (25.4)	24.6 (100)	21.8
Mean		12.6 (46.4)	2.4 (9.4)	3.4 (13.8)	7.9 (30.5)	26.3 (100)	23.7

(I46.9), Unspecified malignant neoplasm (C80.9), Senility (R54), Pneumonitis (J69) and Unattended deaths (R98). With the exception of Japan, all countries also had two cells with insufficiently specified causes (orange) among the 20 leading causes of death.

For females, the 20 top disease rankings were even more saturated with unusable and poorly specified disease groups and included, apart from those mentioned for males, Essential (primary) hypertension (I10), Septicaemia (A41.9) and Malignant neoplasm of overlapping lesion of bronchus and lung (C34.8). Japan and Switzerland each had five red cells among the leading causes of female deaths, Denmark four, Canada and Germany three and Australia two.

Fortunately, a relatively small number of ICD codes are responsible for the major share of the garbage codes in these countries. In Table 5, the most common garbage codes for each country have been identified. If these relatively few codes were not used, it would lead to a 25% reduction in the total amount of garbage codes. In Japan, this could be achieved very easily by avoiding the use of two codes: Senility (R54) and Unspecified heart failure (I50.9). As shown in Table 5, Denmark, Germany and Switzerland would need to focus on three codes and Canada and Australia, respectively, on six and seven. In other words, significant reductions in garbage codes could be achieved if certifiers, instead of just certifying that patients died from old age, heart failure, hypertension, septicaemia and unspecified cancer, could more accurately report the sequence of events leading to death, including the underlying cause of that sequence. As noted above, the problem with the code R99 (ill-defined and unspecified causes of mortality) is not that it cannot be used but that it is over-used, e.g. it is unlikely that no cause could be identified in Denmark for 5% of all deaths.

Discussion

Although the countries included in the study have highly developed mortality information systems and have been producing COD data aligned with international standards for many years, the assessment of their data still revealed that there were some unexpected deficiencies in their statistics that could have significant implications for policy dialogue, monitoring health progress and evaluating intervention impact. This would appear to be due, in large part, to the lack of standardized instructions for medical certifiers about how to correctly complete the death certificate. All six countries declared that such basic information is not part of the standard training provided to young doctors. While physicians may not need to be trained to use the ICD, they should at least be taught how to properly certify the sequence of events leading to death to ensure that coders can identify correctly the underlying cause that led to the person's death. It is this information that is critical for guiding public health policies to further reduce premature mortality and address the rising costs of health care.

In systems where all deaths are registered with a cause, the bias in the data is largely determined by the level and type of garbage codes they contain. To certify that 17% of all deaths in the 70-plus age group were due to old age (Japan) is unhelpful if health authorities want to have a better understanding of disease management in later life. Most people in the considered countries die at older ages and are likely to have had frequent contact with the health system. It is reasonable therefore to assume that comprehensive medical records exist that should allow physicians to more accurately certify deaths. The tendency to assign "old age" as a COD strongly suggests that the certifier has not been trained and is unaware of the important public health use of the death certificate.

While all countries will have a small proportion of deaths where the circumstances leading to death are either not known or cannot be further specified, hence justifying

Table 4 Top-20 ICD causes of death ranked for males and females, Australia, Canada, Denmark, Germany, Japan and Switzerland, 2015–2016

MALES

Rank	Australia	Canada	Denmark	Germany	Japan	Switzerland
1	Bronchus or lung, unspecified	Bronchus or lung, unspecified	Other ill-defined and unspecified causes of mortality	Atherosclerotic heart disease	Pneumonia, unspecified	Chronic ischaemic heart disease, unspecified
2	Acute myocardial infarction, unspecified	Atherosclerotic heart disease	Bronchus or lung, unspecified	Bronchus or lung, unspecified	Bronchus or lung, unspecified	Bronchus or lung, unspecified
3	Chronic ischaemic heart disease, unspecified	Acute myocardial infarction, unspecified	Malignant neoplasm of prostate	Acute myocardial infarction, unspecified	Stomach, unspecified	Malignant neoplasm of prostate
4	Malignant neoplasm of prostate	Unspecified dementia	Stroke, not specified as haemorrhage or infarction	Malignant neoplasm of prostate	Heart failure, unspecified	Unspecified dementia
5	Unspecified dementia	Malignant neoplasm of prostate	Acute myocardial infarction, unspecified	Heart failure, unspecified	Senility	Other ill-defined and unspecified causes of mortality
6	Chronic obstructive pulmonary disease, unspecified	Chronic obstructive pulmonary disease, unspecified	Unspecified dementia	Other ill-defined and unspecified causes of mortality	Acute myocardial infarction, unspecified	Acute myocardial infarction, unspecified
7	Atherosclerotic heart disease	Stroke, not specified as haemorrhage or infarction	Chronic obstructive pulmonary disease, unspecified	Unspecified dementia	Pneumonitis due to food and vomit	Chronic obstructive pulmonary disease, unspecified
8	Stroke, not specified as haemorrhage or infarction	Colon, unspecified	Chronic obstructive pulmonary disease with acute lower respiratory infection	Chronic obstructive pulmonary disease, unspecified	Liver cell carcinoma	Heart failure, unspecified
9	Pancreas, unspecified	Other ill-defined and unspecified causes of mortality	Chronic ischaemic heart disease, unspecified	Pneumonia, unspecified	Sequelae of cerebral infarction	Unspecified fall, unspecified place
10	Intentional self-harm by hanging, strangulation and suffocation, unspecified place	Pancreas, unspecified	Overlapping lesion of bronchus and lung	Pancreas, unspecified	Cerebral infarction, unspecified	Pneumonia, unspecified
11	Alzheimer's disease, unspecified	Pneumonia, unspecified	Heart failure, unspecified	Parkinson's disease	Malignant neoplasm of prostate	Alzheimer's disease, unspecified
12	Parkinson's disease	Chronic obstructive pulmonary disease with acute lower respiratory infection	Atrial fibrillation and atrial flutter, unspecified	Hypertensive heart disease with (congestive) heart failure	Malignant neoplasm of rectum	Stroke, not specified as haemorrhage or infarction
13	Malignant neoplasm of rectosigmoid junction	Parkinson's disease	Alcoholic cirrhosis of liver	Alcoholic cirrhosis of liver	Pancreas, unspecified	Parkinson's disease
14	Chronic obstructive pulmonary disease with acute lower respiratory infection	Bladder, unspecified	Mental and behavioural disorders due to use of alcohol, dependence syndrome	Colon, unspecified	Intracerebral haemorrhage, unspecified	Pancreas, unspecified
15	Malignant melanoma of skin, unspecified	Alzheimer's disease, unspecified	Chronic obstructive pulmonary disease with acute exacerbation, unspecified	Stroke, not specified as haemorrhage or infarction	Acute ischaemic heart disease, unspecified	Chronic obstructive pulmonary disease with acute lower respiratory infection
16	Other ill-defined and unspecified causes of mortality	Oesophagus, unspecified	Malignant neoplasm, unspecified	Atrial fibrillation and atrial flutter, unspecified	Intentional self-harm by hanging, strangulation and suffocation, home	Bladder, unspecified
17	Oesophagus, unspecified	Intentional self-harm by hanging, strangulation and suffocation	Pancreas, unspecified	Unspecified fall	Oesophagus, unspecified	Cardiac arrest, unspecified
18	Unspecified fall, unspecified place	Chronic ischaemic heart disease, unspecified	Atherosclerotic heart disease	Malignant neoplasm of rectum	Interstitial pulmonary disease, unspecified	Liver cell carcinoma
19	Pneumonia, unspecified	Congestive heart failure	Malignant neoplasm of rectum	Unattended death	Chronic renal failure, unspecified	Colon, unspecified
20	Bladder, unspecified	Other interstitial pulmonary diseases with fibrosis	Bacterial pneumonia, unspecified	Chronic obstructive pulmonary disease with acute lower respiratory infection	Other ill-defined and unspecified causes of mortality	Alcoholic cirrhosis of liver

Table 4 continued

FEMALES

Rank	Australia	Canada	Denmark	Germany	Japan	Switzerland
1	Unspecified dementia	Unspecified dementia	Unspecified dementia	Atherosclerotic heart disease	Senility	Unspecified dementia
2	Acute myocardial infarction, unspecified	Bronchus or lung, unspecified	Other ill-defined and unspecified causes of mortality	Unspecified dementia	Pneumonia, unspecified	Chronic ischaemic heart disease, unspecified
3	Bronchus or lung, unspecified	Atherosclerotic heart disease	Bronchus or lung, unspecified	Heart failure, unspecified	Heart failure, unspecified	Breast, unspecified
4	Chronic ischaemic heart disease, unspecified	Acute myocardial infarction, unspecified	Stroke, not specified as haemorrhage or infarction	Breast, unspecified	Bronchus or lung, unspecified	Alzheimer's disease, unspecified
5	Stroke, not specified as haemorrhage or infarction	Breast, unspecified	Senility	Bronchus or lung, unspecified	Pneumonitis due to food and vomit	Bronchus or lung, unspecified
6	Breast, unspecified	Stroke, not specified as haemorrhage or infarction	Breast, unspecified	Acute myocardial infarction, unspecified	Sequelae of cerebral infarction	Other ill-defined and unspecified causes of mortality
7	Alzheimer's disease, unspecified	Alzheimer's disease, unspecified	Chronic obstructive pulmonary disease, unspecified	Hypertensive heart disease with (congestive) heart failure	Cerebral infarction, unspecified	Stroke, not specified as haemorrhage or infarction
8	Chronic obstructive pulmonary disease, unspecified	Chronic obstructive pulmonary disease, unspecified	Chronic obstructive pulmonary disease with acute lower respiratory infection	Atrial fibrillation and atrial flutter, unspecified	Acute myocardial infarction, unspecified	Heart failure, unspecified
9	Atrial fibrillation and atrial flutter, unspecified	Pneumonia, unspecified	Acute myocardial infarction, unspecified	Stroke, not specified as haemorrhage or infarction	Stomach, unspecified	Unspecified fall, unspecified place
10	Pancreas, unspecified	Colon, unspecified	Overlapping lesion of breast	Chronic obstructive pulmonary disease, unspecified	Breast, unspecified	Acute myocardial infarction, unspecified
11	Congestive heart failure	Pancreas, unspecified	Atrial fibrillation and atrial flutter, unspecified	Pneumonia, unspecified	Pancreas, unspecified	Pneumonia, unspecified
12	Pneumonia, unspecified	Malignant neoplasm of ovary	Chronic obstructive pulmonary disease with acute exacerbation, unspecified	Other ill-defined and unspecified causes of mortality	Congestive heart failure	Hypertensive heart disease with (congestive) heart failure
13	Unspecified fall, unspecified place	Atrial fibrillation and atrial flutter, unspecified	Malignant neoplasm of ovary	Pancreas, unspecified	Liver cell carcinoma	Essential (primary) hypertension
14	Malignant neoplasm of ovary	Congestive heart failure	Alzheimer's disease with late onset	Unspecified fall	Chronic renal failure, unspecified	Hypertensive heart disease without (congestive) heart failure
15	Chronic obstructive pulmonary disease with acute lower respiratory infection	Chronic obstructive pulmonary disease with acute lower respiratory infection	Malignant neoplasm, unspecified	Malignant neoplasm of ovary	Intracerebral haemorrhage, unspecified	Chronic obstructive pulmonary disease, unspecified
16	Malignant neoplasm of rectosigmoid junction	Other ill-defined and unspecified causes of mortality	Heart failure, unspecified	Colon, unspecified	Unspecified dementia	Pancreas, unspecified
17	Atherosclerotic heart disease	Unspecified fall	Overlapping lesion of bronchus and lung	Aortic (valve) stenosis	Acute ischaemic heart disease, unspecified	Malignant neoplasm of ovary
18	Septicaemia, unspecified	Malignant neoplasm, primary site unknown, so stated	Atherosclerotic heart disease	Essential (primary) hypertension	Atrial fibrillation and atrial flutter, unspecified	Cardiac arrest, unspecified
19	Heart failure, unspecified	Chronic ischaemic heart disease, unspecified	Alzheimer's disease, unspecified	Parkinson's disease	Colon, unspecified	Senility
20	Colon, unspecified	Parkinson's disease	Colon, unspecified	Cerebral infarction, unspecified	Septicaemia, unspecified	Endocarditis, valve unspecified

Red = Garbage codes with very high, high, medium impact and thus considered "unusable". Orange = Garbage codes with low impact and thus considered "insufficiently specified"

Table 5 Top unusable codes as % of total garbage codes

ICD code	Cause of death	No. of deaths	% of all garbage
Australia 2016			
I50.0	Congestive heart failure	1.748	5.0
A41.9	Septicaemia, unspecified	1.485	4.3
C80.9	Malignant neoplasm, unspecified	1.320	3.8
C80.0	Malignant neoplasm, primary site unknown, so stated	1.234	3.5
I50.9	Heart failure, unspecified	1.127	3.2
C26.0	Intestinal tract, part unspecified	1.049	3.0
R99.-	Other ill-defined and unspecified causes of mortality	1.043	3.0
Canada 2016			
R99.-	Other ill-defined and unspecified causes of mortality	3.969	6.8
I50.0	Congestive heart failure	3.055	5.2
C80.0	Malignant neoplasm, primary site unknown	2.251	3.8
I50.9	Heart failure, unspecified	1.941	3.3
A41.9	Septicaemia, unspecified	1.716	2.9
X42.-	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	1.462	2.5
Denmark 2015			
R99.-	Other ill-defined and unspecified causes of mortality	2.687	18.3
R54.-	Senility	835	5.7
I50.9	Heart failure, unspecified	685	4.7
Germany 2016			
I50.9	Heart failure, unspecified	33.259	14.2
R99.-	Other ill-defined and unspecified causes of mortality	18.368	7.8
I10.-	Essential (primary) hypertension	7.078	3.0
Japan 2015			
R54.-	Senility	84.810	18.4
I50.9	Heart failure, unspecified	58.240	12.7
Switzerland 2015			
R99.-	Other ill-defined and unspecified causes of mortality	2.125	12.8
I50.9	Heart failure, unspecified	1.407	8.5
I10.-	Essential (primary) hypertension	784	4.7

Cut-off for each country-specific listing was a cumulative % of total garbage codes of 25%. Australia, Canada, Denmark, Germany, Japan and Switzerland, 2015–2016

the use of the R99 code "Other ill-defined and unspecified deaths", it is of concern when this cause appears among the leading causes. The same can be said for "Unspecified heart failure", which is an intermediary COD and which accounts for between 3 and 14% of the garbage codes in the six countries examined. Rather, physicians should specify the underlying cause that led to death, which in the case of heart failure could, among others, be myocardial infarction, chronic renal failure, cerebrovascular accident, poisoning and haemorrhage.

Decreasing the proportion of deaths that are of no or little policy value should be a priority for all country health information systems. This assessment and analysis have shown that it is possible for health authorities to quickly obtain insight into the quality problems and certification

errors in the data, which can assist them to take corrective action. The results of the analysis reveal the main certification errors committed by doctors and identify the CODs that introduce bias into the leading causes of death. As demonstrated, the quantity, severity and type of garbage codes appearing in the national cause of death data vary across countries, suggesting that certification problems and coding may be somewhat culture specific. Identifying the specific garbage codes that produce the most unusable data in each country is an essential first step so that strategies can be tailored and developed to determine effective ways to train or inform certifiers. This will help certifiers to be more specific when they complete medical death certificates and to avoid committing errors that lead to garbage codes. Being aware of quality problems in the data can also

assist the authorities responsible for publishing the COD data to provide the explanations needed to correctly interpret the statistics. For instance, in Australia the government agency that publishes the COD data is careful to add the poisoning agent when it publishes data on the accidental poisoning garbage codes such as X42 and X44, thus increasing the information content of the data for policy that normally would be missing from these garbage codes.

National COD data represent a compilation of data from different geographic areas and health facilities, which may vary in their death certification practices, and hence accuracy. It is therefore advisable that countries undertake subnational assessments to verify how certification practices vary and tailor intervention strategies accordingly with more local approaches. For instance, lack of certain diagnostic imaging and analysis and under-staffing can lead to less-than-optimal medical records and make it even more challenging to correctly certify the COD.

Undertaking this type of assessment and communicating the findings to health authorities and medical associations will result in greater awareness of the need to pay more attention to the quality of medical certification. Medical schools in all six countries should give higher priority to the certification duty of their profession, and correct medical certification certainly should be included as a compulsory element of the induction programs for interns. Doctors perform a very important public health function in documenting the COD of their patients, not only for the families of the deceased, but also for society. They are generally unaware of this, or of its importance for public policy. Without reliable and detailed information on the leading CODs, and how they are changing, health planning and policy will be less cost-effective than otherwise might be the case, potentially resulting in lost opportunities to improve population health.

While ANACONDA cannot verify whether the physician diagnosed the correct COD or whether a COD was miscoded, it can detect whether the code assigned is a valid underlying cause and whether the certifier originally completed the death certificate according to ICD guidance. Adopting the ICD classification without adhering to its rules and standards for coding and guidance for certification will not provide good-quality COD information for public health use. This study has demonstrated that even for countries with very advanced health information systems, it is very informative to undertake an assessment of the COD data as the information content may be reduced by a high proportion of unusable and insufficiently specified causes.

Once the data have been evaluated and the specific problems revealed, focused action should be taken to reduce the amount of garbage codes by, for example, introducing certification training for hospital interns and awareness raising in the medical community of the important functions of the death certificate for the national health information system. The large and complex bureaucracy that all countries have established to collect these data needs to meet the demands of increasingly sophisticated and complex health systems and provide them with the detailed information required for avoiding premature deaths and keeping people alive and healthy for as long as possible.

Compliance with ethical standards

References

Adair T, Kippen R, Naghavi M, Lopez AD (2019) The setting of the rising sun? A recent comparative history of life expectancy trends in Japan and Australia. PLoS ONE 14(3):e0214578. https://doi.org/10.1371/journal.pone.0214578

AIHW (2018) Australia's health 2018, Australian Institute of Health and Welfare 2018. Australia's health series no. 16. AUS 221. AIHW, Canberra. https://www.aihw.gov.au/getmedia/7c42913d-295f-4bc9-9c24-4e44eff4a04a/aihw-aus-221.pdf.aspx?inline=true

Alpérovitch A, Bertrand M, Jougla E, Vidal J-S, Ducimetière P, Helmer C et al (2009) Do we really know the cause of death of the very old? Comparison between official mortality statistics and cohort study classification. Eur J Epidemiol 24(11):669–675

Chammartin F, Probst-Hensch N, Utzinger J, Vounatsou P (2016) Mortality atlas of the main causes of death in Switzerland, 2008–2012. Swiss Med Wkly 146:14280

Civil Registration and Vital Statistics Knowledge Gateway. https://crvsgateway.info/search-filter?s_q_fulltext=ANACONDA

ICD-11 Reference Guide (2019). https://icd.who.int/icd11refguide/en/index.html

Lopez AD (2013) Reducing risks to health: what can we learn from the Global Burden of Disease 2010 Study? Int J Public Health 58(5):645–646

Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon F, Rao C, Sretzer S (2007) Civil registration systems and vital statistics: successes and missed opportunities. Lancet 370:1653–1663

Mikkelsen et al (2015) Counting births and deaths 3—a global assessment of civil registration and vital statistics systems: monitoring data quality and progress. Lancet 386:1395–1406

Mikkelsen L, Moesgaard K, Hegnauer M, Lopez AD (2020) ANACONDA: a new tool to improve mortality and cause of death data. BMC Med

Murray CJL, Lopez AD (eds) (1996) The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. Global Burden of Disease and Injury Series, vol I. Harvard University Press on behalf of the World Health Organization and The World Bank, Cambridge, MA

Murray CJL et al (2018) Population and fertility by age and sex for 195 countries and territories 1950–2017: a systematic analysis for the Global Burden of Disease 2017. Lancet 392:1995–2051

Naghavi M (2020) Improving the quality of death data for public health policy: are all 'garbage' codes equally problematic? BMC Med

Naghavi M, Makela S, Foreman K, O'Brien J, Pourmalek F, Lozano R (2010) Algorithms for enhancing public health utility of national causes-of-death data. Popul Health Metr. 8:9

Philips DE, Lozano R, Naghavi M, Atkinson C, Gonzales-Medina D, Mikkelsen L, Murray CJL, Lopez AD (2014) A composite metric for assessing data on mortality and causes of death: the Vital Statistics Performance Index. Popul Health Metr 12:14

Rampatige R, Mikkelsen L, Hernandez B, Lopez AD (2014) Hospital cause-of-death statistics: what should we make of them? Bull World Health Organ 92:3–3A. https://doi.org/10.2471/BLT.13.134106

Roth GA, Dwyer-Lindgren L (2017) Trends and patterns of geographic variation in cardiovascular mortality among US counties, 1980–2014. JAMA 317(19):1976–1992. https://doi.org/10.1001/jama.2017.4150

Roth GA et al (2018) Global, regional, and national age-sex-specific mortality for 282 causes of death, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392:1736–1788

Schdev A et al (2001) Problems with proper completion and accuracy of the cause-of-death statement. Arch Intern Med 161:277–284

Shibuya K (2006) Counting the death is essential for health. Bull WHO 84(3):161–256

World Health Organization (2016) International statistical classification of diseases and related health problems, 10th revision, vol 2, 5th edn. World Health Organization, Geneva

WHO Mortality Database. CoDQL cause of death query online. http://apps.who.int/healthinfo/statistics/mortality/causeofdeath_query/

Wang H, Naghavi M, Allen C, Lopez AD, Murray CJL (2016) Global, regional, and national life expectancy, all-cause specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the global Burden of Disease Study 2015. Lancet 388:1459–544

Affiliations

Lene Mikkelsen[1] · Kim Moesgaard Iburg[2] · Tim Adair[1] · Thomas Fürst[3] · Michael Hegnauer[3] · Elena von der Lippe[4] · Lauren Moran[5] · Shuhei Nomura[6] · Haruka Sakamoto[6] · Kenji Shibuya[7] · Annelene Wengler[4] · Stephanie Willbond[8] · Patricia Wood[8] · Alan D. Lopez[1]

✉ Alan D. Lopez
  alan.lopez@unimelb.edu.au

  Lene Mikkelsen
  l.mikkelsen@unimelb.edu.au

  Kim Moesgaard Iburg
  kmi@ph.au.dk

  Tim Adair
  timothy.adair@unimelb.edu.au

  Thomas Fürst
  thomas.fuerst@swisstph.ch

  Michael Hegnauer
  michael.hegnauer@swisstph.ch

  Elena von der Lippe
  vonderLippeE@rki.de

  Lauren Moran
  lauren.moran@abs.gov.au

  Shuhei Nomura
  nom3.shu@gmail.com

  Haruka Sakamoto
  harukask1231@gmail.com

  Kenji Shibuya
  kenjishibuya1@gmail.com

  Annelene Wengler
  WenglerA@rki.de

  Stephanie Willbond
  stephanie.willbond@canada.ca

  Patricia Wood
  patricia.wood@canada.ca

[1] Bloomberg Data for Health Initiative, University of Melbourne, Melbourne, Australia

[2] Institute of Public Health, Aarhus University, Århus, Denmark

[3] Swiss Tropical and Public Health Institute, Basel, Switzerland

[4] Department of Epidemiology and Health Monitoring, Robert Koch Institute, Berlin, Germany

[5] Australian Bureau of Statistics, Canberra, Australia

[6] Department of Global Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[7] University Institute of Population Health, King's College London, London, UK

[8] Statistics Canada, Ottawa, Canada