

## ANACoD version 1.0

# Analysing mortality level and cause-of-death data

### About this mortality tool

The production of the version 1 of the electronic tool has been made possible with the collaboration of the University of Queensland, World Health Organization and Health Metrics Network.

This tool is freely available for public use. However requests for permission to translate or draw material from the tool should be addressed to WHO, Department of Health Statistics and information Systems, 20 Ave Appia, 1211 Geneva 27, Switzerland. When using the tool, the responsibility for the interpretation of the results lies with the reader.

### Objectives of this tool

This tool is intended to help build analytical capacity to assess the quality of mortality statistics in order to enhance their value in informing health policies and programmes.

Countries routinely invest significant resources into collecting mortality data from a variety of sources, including civil registration systems, health care facilities, ongoing longitudinal demographic and health surveillance, and from other data sources such as the census or household surveys. The primary purpose is, of course, to generate critical information to guide public health decision-making. However, data cannot be used appropriately or with any confidence if insufficient attention is paid to their quality. In the absence of systematic data quality assessment, and adjustment where necessary, the data that have been collected – often at great expense – cannot be used to their full potential to guide decision-making.

This tool provides relatively simple ways of analysing the *internal* validity and coherence of mortality data and shows how comparisons with other, *external*, sources of mortality data can be used to assess data consistency and plausibility. By carrying out these simple checks, data collectors and practitioners will be able to diagnose weaknesses in their data.

The tool has 3 parts:

- I. Input data (Steps 0-1)
- II. Mortality Levels analysis (Steps 2-5)
- III. Causes of deaths analysis (Steps 6-10)

### I. Steps 0-1                    Input data and check

Steps 0 is used to input the data: it is the basic sheet where users will paste the mortality and population data, as well as select the relevant characteristics of the data: country, year, ICD-10 used. The tool will then calculate and produce all the tables and graphs. Before any analysis, Step 1 will first make calculations that will allow to identify basic errors.

### II. Steps 2-5                    Mortality Levels

Steps 2-5 focus on simple steps to assess the plausibility of the mortality levels. The tool compiles and formats the raw data to enable the calculation of crude death rates, age-specific mortality rates, life expectancy at birth and child mortality. These simple measures of mortality can provide insights into the health status of a population over time. In addition the crude death rates provide a useful indicator of

possible problems with the completeness of mortality data. The tool allows also the comparison of the results obtained with regional estimates.

## II. Steps 6-10 Causes of death

Steps 6-10 focus on simple steps to assess the plausibility of data on causes of death. Information on the levels and patterns of mortality among different population groups is essential for public health authorities and for the effective allocation of resources to health care. However, a fully functioning civil registration and vital statistics system should not only register deaths by age and sex, but should also collaborate with the health sector to implement mechanisms for assigning the cause of death according to international standards as expressed in the International Statistical Classification of Diseases and Related Health Conditions (ICD) 10<sup>th</sup> revision. Only a medically qualified doctor should determine the cause of death, and a coding expert trained in the ICD rules and principles should determine the underlying cause of death, from a death certificate properly filled out by a physician, as defined in the ICD. Note that this coding expert should NOT be a medical doctor as this is NOT the best use of his/her time.

The objectives of steps 6-10 are to enable users to:

- Calculate broad patterns of causes of death using available data on mortality by age, sex and cause
- Critically analyse and interpret cause of death data
- Assess the plausibility of the cause of death patterns emerging from the data

### Definition of the underlying cause of death

The quality of cause of death data depends on the reliability of **death certification** and the accuracy of **coding**. These are two separate, but related functions. Death certification, which should only be done by a qualified medical practitioner, involves correctly completing an international form (Medical Certificate of Death). This information is then translated into a code (alpha-numeric digital code) from among the 10 000 codes that can be used as the underlying causes of death in the ICD by a qualified and trained coder (**not** the physician who certified the death as he/she is unlikely to have been formally trained in the coding of information given on a death certificate).

There are well-established rules for assigning the cause of death and it is essential that deaths be classified not by the **immediate** cause of death but by the **underlying** cause, that is, the cause that initiated the sequence of events leading to the death. It is the underlying cause of death that generates information that is useful for public health purposes. The underlying cause of death is defined by the World Health Organization (WHO) as the disease or injury that initiated the train of events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury. Under international rules for selecting (i.e. coding) the underlying cause from the reported conditions, every death is attributed to one (and only one) underlying cause based on information reported on the death certificate. The International Form of Medical Certificate of Cause of Death was specially designed to facilitate the selection of the underlying cause of death when two or more causes are recorded on the death certificate. This certificate is shown in Box 1 and should be filled in only by a trained medical practitioner. Moreover, all countries are strongly urged to use this certificate to certify death, and NOT some other adaptation of it which will be of limited public health value.

**Box 1: International Form of Medical Certificate of cause of death**

INTERNATIONAL FORM OF MEDICAL CERTIFICATE OF CAUSE OF DEATH

Cause of death		Approximate interval between onset and death
<p><b>I</b></p> <p>Disease or condition directly leading to death*</p>	<p>(a) .....</p> <p>due to (or as a consequence of)</p>	<p>.....</p>
<p><b>Antecedent causes</b></p> <p>Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last</p>	<p>(b) .....</p> <p>due to (or as a consequence of)</p> <p>(c) .....</p> <p>due to (or as a consequence of)</p> <p>(d) .....</p>	<p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>
<p><b>II</b></p> <p>Other significant conditions contributing to the death, but not related to the disease or condition causing it</p>	<p>.....</p> <p>.....</p>	<p>.....</p> <p>.....</p>

\*This does not mean the mode of dying, e.g. heart failure, respiratory failure. It means the disease, injury, or complication that caused death.

Currently only about 70 WHO member countries produce good quality cause-of-death data from their civil registration and vital statistics systems.<sup>1</sup> Although a further 50 countries produce some cause-of-death data, the quality of the information is problematic because of poor certification and coding practices.

Even where medical certification of the cause of death is common practice, it does not necessarily mean that the correct cause of death is written on the death certificate in the correct way. Doctors may lack diagnostic facilities, are inexperienced in filling the death certificate and unaware of the importance of cause-of-death data. All these factors contribute to poor diagnostic accuracy. In addition, there may be financial or social consequences for the family that deter the doctor from reporting the true cause of death.

For all these reasons, any data set with information on causes of death by age and sex should be carefully reviewed and assessed to identify and correct potential quality problems. Unless this is done as a matter of course, public health authorities using the data risk diverting resources away from those conditions that are causing the most serious problems of suffering and death in their communities.

**Step 0: Input of deaths by age, sex and cause of death**

The first step is to aggregate the individual death records and tabulate the available data on deaths by age, sex and causes (using ICD codes).

As a minimum, the tabulations should include:

- Numbers of deaths for a specified year
- By sex (i.e. for males and females separately)
- By age at death<sup>2</sup> using the following age groupings:
  - Within the first year after birth;

<sup>1</sup> Mathers C, Ma Fat DM, Inoue M et al. (2005) Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bulletin of the World Health Organization*. 2005 Mar;83(3):171-7. Epub 2005 Mar 16.

<sup>2</sup> 'Completed years of age' is a demographic concept that defines the exact age at death. For example, a person who dies *after* their 34<sup>th</sup> birthday, but *before* their 35<sup>th</sup> birthday, will have completed 34, but NOT 35 years of age. The age of death in completed years is thus 34.

27 June 2011

- Between completed years 1-4
- Completed years 5-9
- Completed years 10-14 etc, by 5-year age groups, up to completed years 80-84 at the minimum up to 90-94
- Completed years 85 and over up to 95 year and over
- By ICD-10 list of causes<sup>3</sup>

In addition, the tabulations should include the mid-year population for the same year, sex and age group. Population estimates are generally available from the decennial census and intercensal projections produced by the National Statistics Office. These data will be used for the calculation of rates and ratios that will be explained in the subsequent steps.

It is usual practice to use five year age groups except for deaths occurring in children under 5 which are subdivided into those occurring within the first year of life and those occurring between the ages of 1 to 4 years. Precision is also important at older ages which should continue to be grouped into five year categories at least up to the age of 85 years<sup>4</sup>.

It is poor practice to only tabulate age of death to some relatively low terminal age such as 55+ or 65+. Increasingly, more and more deaths are occurring in populations after age 50 or so, and it is extremely important for preventive efforts to distinguish between a death at age 80-84 for example, and an adult death at a much younger age, say 60-64 or 65-69 years. The use of these standard five-year age groupings is important because the same age groups are used to compile census data on population size and distribution that are used as denominators for the calculation of rates and ratios.

## Step 1: Basic checks

The purpose of this initial tabulation sheet is to identify gross errors in the data set. The WHO mortality tool will automatically make graphs that will illustrate population and mortality age-patterns, in order to depict errors that can be identified visually. Tabulations will also identify clearly incorrect causes of death such as male maternal deaths, suicides among young children or prostate cancer deaths in females. It will also draw the attention of users to invalid use of certain codes as underlying causes of death, thus alerting data managers to potential quality problems in coding causes of death from death certificates, or in the certification process (e.g. implausible sequence of morbid conditions reported on the death certificate).

**If potential errors are identified, users should go back to Step 0 and corrected the data as appropriate.**

### Dealing with fluctuations

In countries and settings with small population numbers, the annual number of deaths at specific ages may be very small. As a result, the data derived from it would tend to fluctuate and be too unstable for analysis. In order to overcome this problem, it is usual to calculate an average of number of death and population over a 3-5 year period in order to average out annual fluctuations. This is illustrated in the example given below in Figure 1 for a small Pacific island population which shows the large fluctuations in annual age specific mortality rate (see step 1 - 2.1 in the tool and step 3 for explanation on ASMR) and the smoothed trend data produced

<sup>3</sup> <http://apps.who.int/classifications/apps/icd/icd10online/>

<sup>4</sup> If possible, age at death should be tabulated in 5 year age groups, up to age 95-99 and 100+ as the terminal age group for statistical purposes. If this is not possible, 85+ should be chosen as the terminal age group.

by using a 3-year moving average. Alternatively, it is possible to expand the age-group or area to be studied, thus increasing the numbers of deaths in the calculation of ASMR.

**Figure 1: Annual age-specific mortality rates for selected age groups (males), and smoothed trends using a 3-year moving average**



Once the data have been entered, and the gross errors identified and corrected, the tool will automatically calculate totals and distributions of deaths by sex, age-group and cause. The subsequent steps outlined below involve the calculation and analysis of key indicators that can alert users to possible weaknesses in their mortality data set.

### Summary of step 1

- The tools identify gross errors in the input data set. If potential errors are identified, users should go back to Step 0 and corrected the input data as appropriate.

## Step 2: Crude death rates

The second step in assessing the quality of a set of mortality data is to review the calculated level of the crude death rate (CDR). This is done for two reasons. First, the CDR is the simplest measure of mortality that can provide insights into the health status of a population over time. Second, the CDR provides a useful indicator of possible problems with the completeness of mortality data.

The objectives of step 2 are to enable users to:

- Define and calculate the CDR
- Understand the public health relevance of the CDR
- Interpret the CDR and judge its limitations
- Use the CDR as an approximate indicator of completeness of death registration.
- Use the CDR as the first step to analyse the quality of mortality data

### Definition and calculation of the crude death rate

The Crude Death Rate (CDR) is a measure of the number of deaths in a population, relative to the size of that population over a given period of time. The CDR is typically expressed in units of deaths per 1000 individuals

27 June 2011

per year; thus, a crude death rate of 9.5/1000 in a population of 500,000 indicates there were 4,750 deaths per year in the total population (9.5/1000 X 500,000). The CDR is defined and calculated as follows:

$$\text{Crude death rate} = \frac{\text{Number of deaths in the usual resident population in a given year}}{\text{Size of the midyear resident population in that year}} \times 1000$$

Because mortality rates for males and females differ across all ages, it is useful to calculate the crude death rate separately for males, females and both sexes combined.

It is important that both numerator and denominator refer to the same population in terms of geography and time. It is standard practice to take the size of the population at midyear as the denominator because population size may vary during the year (due to migration, births and deaths) and the midyear population serves as an estimate of the average population exposed to the risk of dying over the course of the year.

### Interpreting the crude death rate

The CDR is called a “crude” rate because it does not take into consideration the age and sex structure of the population. In practice, the risk of death in a given population group varies according to age and sex as well as patterns of socio-economic status, environmental and other factors. For example, populations with a large proportion of young children or a high proportion of elderly people will, other things being equal, have relatively higher crude death rates. This is because mortality risks are highest at very young and the oldest ages. In general, mortality rates are higher among males than females.

### Crude death rate and population structures

In order to interpret the CDR, it is helpful to refer to the population age-sex pyramid, a graphical illustration of the distribution of the population by standard (usually 5-year) age groups. The population pyramid typically consists of two back-to-back bar graphs, with age groups on the vertical axis and population size in each age group on the horizontal axis. Males are conventionally shown on the left and females on the right. The bars can represent either the absolute numbers (more common) or percentages of the total (male or female) population in each 5-year age group.

In most developed countries, the age-sex pyramid is constructed on the basis of annual birth and death data from the civil registration system and the decennial census. In countries where civil registration systems are weak, age-sex population pyramids can only be reliably estimated from the census. Intercensal estimates of population size by age and sex generally need to be estimated from mortality rates derived from model life tables, which are inherently uncertain. The United Nations Population Division generates regular updates on national population sex and age structures which should be used where there is doubt about the reliability of country population data.<sup>5 6</sup>

The use of population pyramids in helping to interpret Crude Death Rates is illustrated in **Figure 2**. The CDR for Sudan in 2005 is estimated at 13 per 1000 population compared with 9 per 1000 population in Japan. This difference reflects the fact that Sudan has a high proportion of children aged below 4 years and this is precisely the age group where mortality rates are highest. By contrast, Japan has a much smaller percentage of population in this age group, although it has a large proportion of older people aged 60+, when death rates

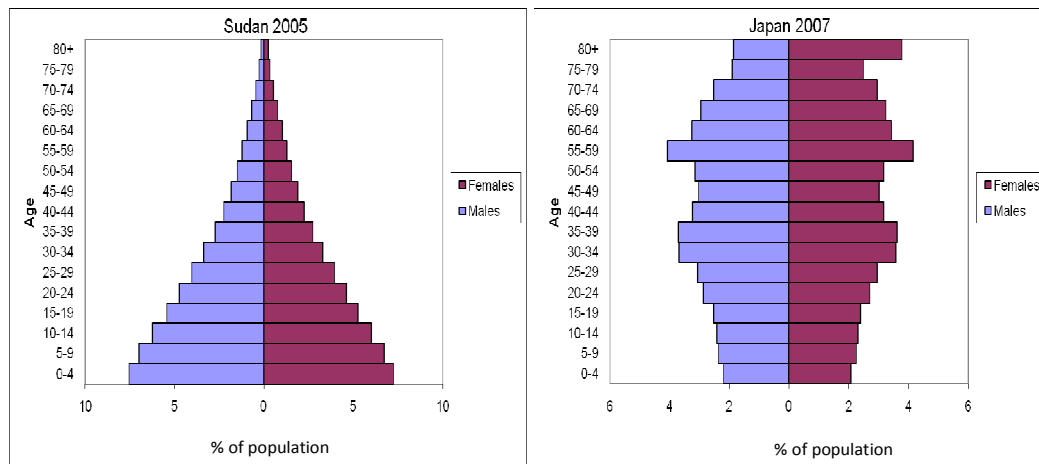
---

<sup>5</sup> <http://esa.un.org/unpd/wpp>

<sup>6</sup> For more details on drawing a population pyramid, go to <http://www.census.gov/ipc/www/idb/#>

are also high. However, this is insufficient to counteract the effect of a large population of children in Sudan among whom death rates are comparatively high.

**Figure 2: Population pyramids for Sudan 2005 and Japan 2007**



Source: Calculated from UN Population Division estimates <http://esa.un.org/unpp/index.asp>.<sup>1</sup>

### Lower limits for the crude death rate

Based on many decades of experience in calculating crude death rates, demographers have demonstrated that there is generally a lower limit for the CDR of around 5 per 1000. For example, over the past 20-30 years, Japan has consistently registered the lowest age-specific mortality rates in the world. Yet throughout this period, the CDR in Japan NEVER fell below 5 per 1000.

In the Step 2 sheet, the table "Expected crude death rates at different levels of life expectancy and population growth (based on Coale-Demeny West model)" shows the combinations of life expectancy and population growth rates that are associated with different levels of the CDR. They have been estimated from the Coale-Demeny 'west' family regional model life tables.<sup>7</sup> In many parts of the developing world, population growth rates are typically around 2% per year. In such populations, the CDR can never get below 5/1000, and even for the CDR to fall below 7/1000, life expectancy would need to be 75 years or more. This is relatively uncommon in developing countries and hence **low** CDR's should be treated with great suspicion.

**Any CDR below 5 per 1000 should be treated with extreme caution as such a figure is strongly suggestive of INCOMPLETE death registration.**

However, there are exceptional populations which have both high growth rates – because of natural increase (excess of births over deaths), immigration, or both – and low age-specific mortality rates, including low child death rates, implying comparatively high life expectancy at birth. Several of the Gulf States do in fact have a CDR below 5 because of this particular demographic configuration. In the vast majority of countries however, this does not apply and low CDRs below 5/1000 are typically a sign of under-reporting of deaths.

### Summary of step 2

- The tool calculates the CDR. A level below 5 per 1000 is strongly indicative of incomplete registration of deaths.

<sup>7</sup> Coale AJ, and Demeny P. (1966) *Regional Model Life Tables and Stable Populations* (Princeton, N. J., Princeton University Press, 1966).

- It calculates a population pyramid for your country. If the proportion of young children in the population is high, you should expect the CDR to be relatively high. The same is true when the proportion of older people in the population rises.
- It examines the CDR for males and females separately. You should generally expect the CDR for males to be higher than for females. Deviations from this pattern could indicate that women and girls face severe disadvantages in terms of health and nutrition. Alternatively, there may be problems with data completeness and quality with systematic underreporting of female deaths.
- It compares the calculated CDR with the expected CRD based on the combinations of life expectancy and population growth rates that are associated with different levels of the CDR.

### Step 3 Age- and sex-specific death rates

In Step 2 we analyzed a mortality data set by calculating the CDR for the population. However, the CDR is a “crude” rate because it does not take into consideration the age and sex structure of the population. As we saw from step 2, populations with a large proportion of young children or a high proportion of elderly people will, other things being equal, have relatively higher crude death rates because mortality risks are highest at very young and oldest ages. Moreover, mortality rates are generally higher among males than females across all age groups. Therefore, when comparing the mortality of populations across countries, geographic areas or over time, it is important to use both age-specific and sex-specific mortality rates alongside the crude death rate, and to examine these detailed age- and sex-specific rates for possible age misreporting of deaths.

The objectives of step 3 are to enable users to:

- Define and calculate the mortality rate specific to a population age group (usually a five-year grouping), known as the **age-specific mortality rate** (ASMR)
- Understand the public health relevance of the ASMR
- Interpret the ASMR and understand its limitations
- Use the ASMR to assess the quality of mortality data

#### Definition and calculation of age-specific mortality rates

The age-specific mortality rate is calculated as the total number of deaths, occurring at a specified age or in a specified age group, in a defined geographic area (country, state, county, etc.) divided by the mid-year population of the same age in the same geographic area. By contrast to the CDR, which is expressed per 1000 population, the ASMR is generally expressed as a rate per 100,000 population. This is because there are many fewer deaths within each age group compared with the numbers occurring in the total population. The standard demographic practice is to calculate the ASMR for 5-year age-groups, namely 0-4, 5-9, 10-14, ..., 80-84, 85+. The ASMR is calculated as follows:

$$\text{ASMR} = \frac{\text{Deaths in a specific age group in a population over a specified time period}}{\text{Total mid-year population in the same age group, population and time period}} \times 100,000$$

#### Disaggregation of age-specific mortality rates by sex

As noted in step 2, there are important differences in patterns and levels of mortality between males and females across all age groups. Therefore, it is standard practice to calculate age-specific mortality rates separately for males and females within each age group.



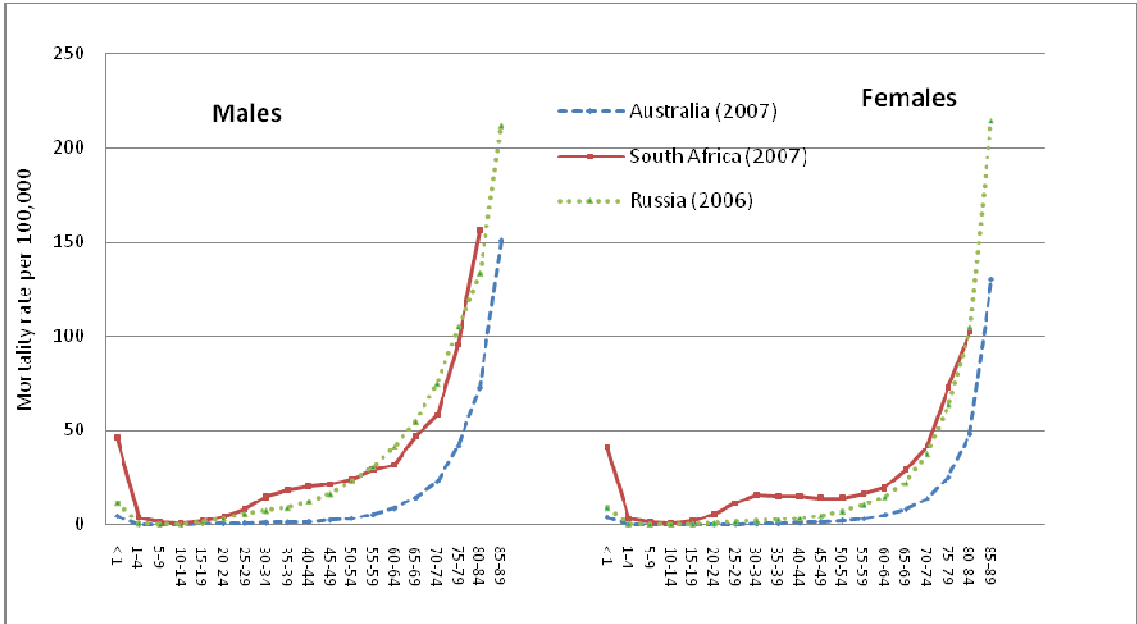
## Interpreting age and sex patterns of mortality

### Age-specific mortality rates

Once the ASMR has been calculated for each age group and sex, the next step is to examine the pattern of the data by age to assess plausibility. In order to do this, it is important to have an independent source of comparative data on ASMR, for example the census. If there is no independent source within a country, it is possible to compare the ASMR with figures from similar countries and settings. The following examples can help in improving the understanding and interpretation of age and sex patterns of mortality in a given country. They also show how this analysis can assist in determining the quality and completeness of the mortality data within specific age groups.

As a general rule, in all settings, mortality rates are high during infancy and early childhood and fall to their lowest levels between the ages of 5 to 14 years. Subsequently, mortality rates start to rise with increasing age and increase exponentially beyond age 35 or so. Figure 3 shows patterns of mortality across age for Australia – where death registration is complete – compared with Russia and South Africa – where death registration is less complete and/or essential information about the death is missing, such as unknown age or sex. In Australia, mortality rates are very low up to the age of 15 or so, and although there is a small increase for males during the ages of 15-34 years due to accidents and other injuries, death rates only really begin to rise sharply after about age 55 years or so. This pattern is typical of most low-mortality populations.

Figure 3: ASMR for Australia, Russia and South Africa, males and females, 2000



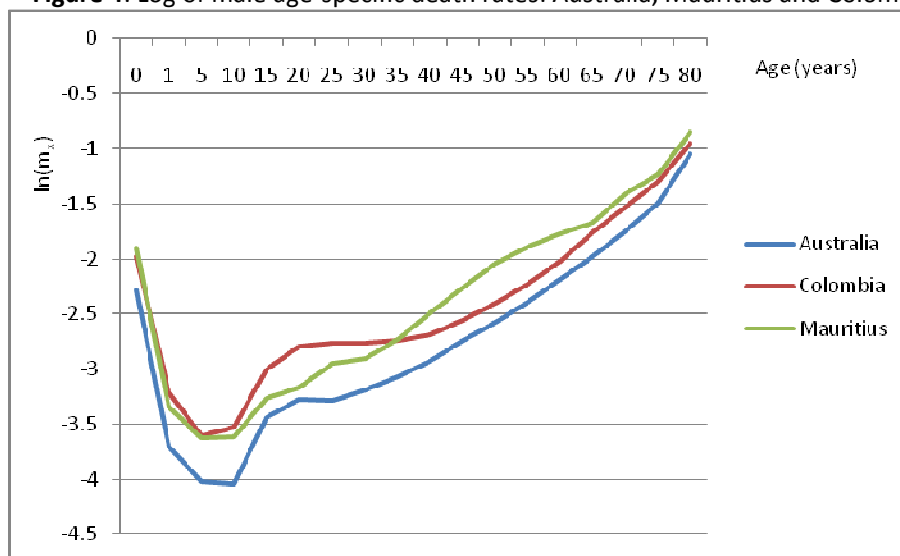
In Russia and South Africa, mortality in infants is relatively high (this is particularly marked in South Africa) but declines during childhood. In South Africa, there is a “bump” in mortality during reproductive ages in both sexes, reflecting premature mortality due to HIV/AIDS. A similar bump may occur in females of reproductive ages in settings where maternal mortality is very high.

Comparing your data with this pattern can provide a simple check on the quality of the mortality data and indicate possible under registration of deaths at certain ages. It is not the *level* of mortality that matters in this comparison but the *relative age pattern* of the ASMR between different age groups.

As noted above, beyond about 35 years of age, death rates rise exponentially with age. Therefore, the *natural logarithm* of the age-specific death rate ( $m_x$ ), written as  $\ln(m_x)$  should be **a straight line** as age ( $x$ ) increases.<sup>8</sup> Figure 4 shows examples of  $\ln(m_x)$  for three countries – Australia, Colombia and Mauritius – with very different patterns of mortality and variable quality of mortality data.

The primary purpose of preparing a graph of the log of the death rate at each age is to examine the data for irregular or implausible changes in  $\ln(m_x)$  from age to age. In countries with high maternal or injury mortality in young adults (especially males), death rates will rise steeply (i.e.  $\ln(m_x)$  will rise) around age 15 years, peak at age 25, and decline to a new low around age 35 years or so. Subsequently the ASMR will rise linearly with age. Any other departure from this linear pattern in adult death rates suggests that deaths are being selectively (by age) underreported, or that there is mis-reporting of the correct age of death. This is particularly common at older ages.

**Figure 4:** Log of male age-specific death rates: Australia, Mauritius and Colombia



With this in mind, we can make the following observations from Figure 4 showing age-specific death rates for males.

**Australia:** all deaths are registered and hence the  $\ln(m_x)$  increases smoothly in a straight line with increasing age( $x$ ), as would be expected. Note the slight bump around ages 15-25, indicating an excess in injury-related deaths in this age group.

**Mauritius:** notice that in this case the  $\ln(m_x)$  does NOT increase linearly with age after about age 65, suggesting underreporting of deaths, particularly at the oldest ages;

**Colombia:** note the large bump in mortality at ages 15-34 due to accidents and other violent deaths. One would expect to see a similar large bump in the  $\ln(m_x)$  graph at these ages in countries with high HIV mortality.

Thus, plotting the  $\ln(m_x)$  will help to identify if there are any age groups where deaths are being selectively underreported (e.g. older ages in Mauritius). In addition, by comparing the graph of  $\ln(m_x)$  for your population with a neighbouring country with good quality mortality data, it will be possible to assess whether and to what

<sup>8</sup>  $m_x$  is the standard demographic notation to indicate the level of the ASMR (written as “ $m$ ”) in any age group “ $x$ ”.

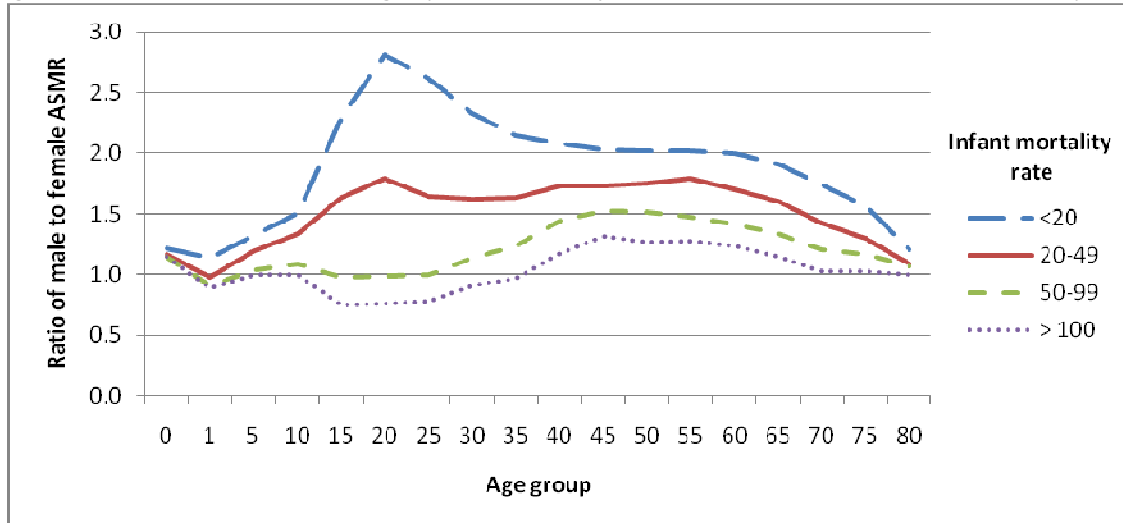
extent, deaths are being systematically underreported at all ages. This will be the case if the graph for  $\ln(m_x)$  for your population is systematically lower than the graph for a neighbouring population.

**Ratio of male to female mortality rates**

As already observed, mortality rates tend to be higher at all ages for males than females. To better understand these male-female differences, it is useful to calculate the ratio of male to female mortality rates by age group. If the ASMR was the same for both sexes, the ratio would be 1 (i.e. a straight line) for all ages. In practice, the male to female ASMR shows considerable variation over different age-groups and at different period of time. Figure 5 shows typical patterns of the male to female ratio in settings with different overall mortality levels, as reflected by levels of infant mortality.

Male death rates are higher than female death rates everywhere except in societies with very low female status. As the status of women in society improves, and discriminatory practices against females disappear, female death rates should be lower than male rates at all ages.<sup>9</sup> As Figure 5 shows, in settings with high levels of infant mortality (100 per 1,000 live births and higher), the male mortality excess is relatively small because of high female mortality in reproductive ages. As overall mortality declines, this pattern changes and male mortality is higher than female mortality across all age groups. As already noted, death rates among males aged 15-29 tend to be higher largely because of mortality due to accidents and other external causes among young men. A secondary peak in the male to female ratio of mortality rates typically occurs around age 55-64 years because males tend to die at higher rates from chronic disease than females, due primarily to greater exposure to risk factors such as tobacco, poor diet and overweight/obesity.

**Figure 5:** Ratio of male to female age-specific mortality rates at different levels of infant mortality



The tool will prepare a similar chart showing the male to female ratio of age-specific death rates based on the mortality data; for reference of the overall pattern across ages, the graph shows also the ratio for the world. You can also compare your pattern with one of the curves shown in Figure 5. If your pattern of male to female ratio of age-specific death rates is very different from what would be expected given your level of infant mortality, there are good reasons for questioning the quality – that is, the completeness of death registration – of the reported data, particularly for females.

<sup>9</sup> Waldron, D (1982) The sex differential in mortality rates: Demographic and behavioral factors. *American Journal of Epidemiology*, 115(2)

**Note that in comparing your age patterns of the sex mortality ratio to one of those from Figure 5, it is important to use an independent value of the infant mortality rate derived from censuses or surveys, or estimated by the United Nations, WHO or other sources. Do NOT use the value from your vital registration data which could be underestimated.**

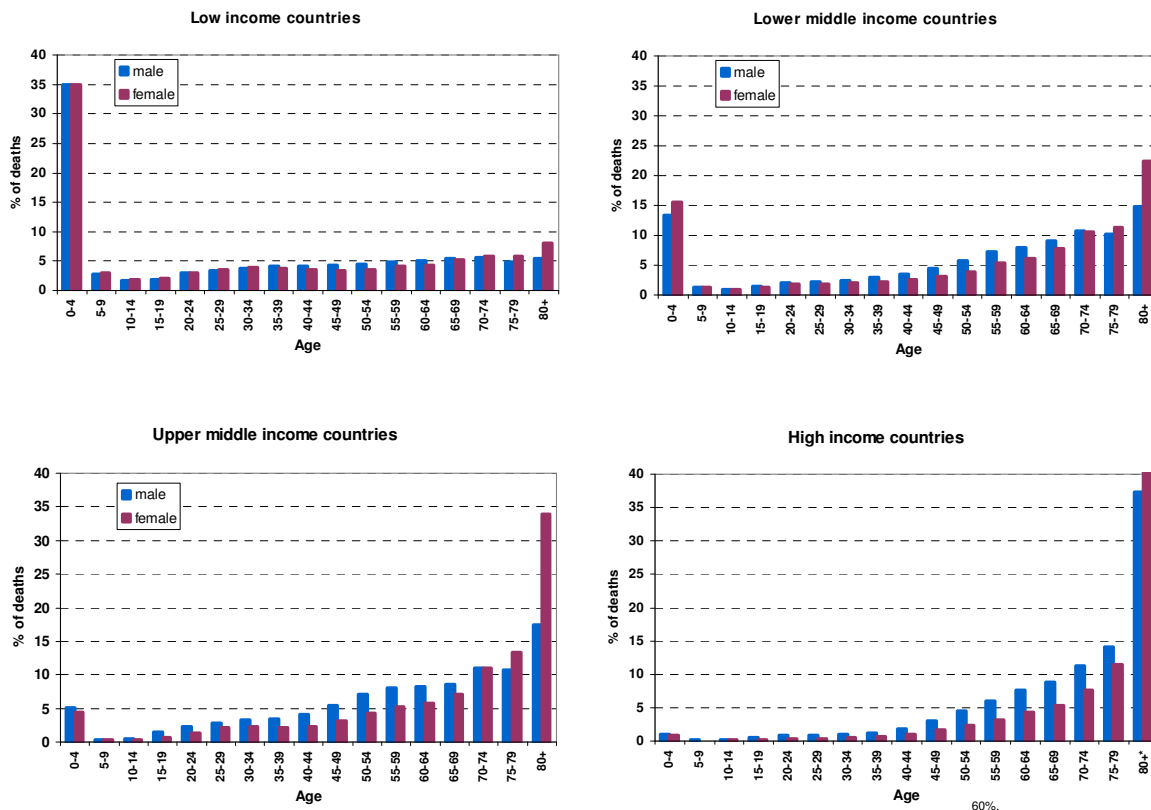
### Summary of step 3

- The tool calculates age and sex-specific mortality rates (ASMR)
- It examines ASMR across all age groups for each sex separately. You should find a pattern of relatively high mortality in the 0-4 years age group, very low mortality in the age groups 5-14, and an exponentially increasing mortality rate after the age of 35 or thereabouts.
- It plots the logarithm of the death rate at each age. It should increase smoothly and linearly with age after about age 35.
- It examines the ratio of male to female age-specific mortality rates across all ages. In general, you would expect male mortality rates to be higher than rates for females, especially in the age groups 15-35 or so as young males are more likely to die as a result of violence, road traffic accidents and other external causes of death. High mortality rates in young adults may also be due to HIV. In some cases, female deaths are less likely to be recorded than male deaths, leading to higher than expected ratios of male to female death rates.

## Step 4 Review the age distribution of deaths

The objective of step 4 is to examine the **age distribution** of reported deaths. This age-distribution should look quite different depending on the overall level of mortality in a population. The tool prepares a chart showing the distribution of deaths by age group. The calculated distribution of deaths should then be compared with one of the expected distributions shown in Figure 6 that corresponds to the income group the country belongs to.

**Figure 6: Age distributions by income grouping**

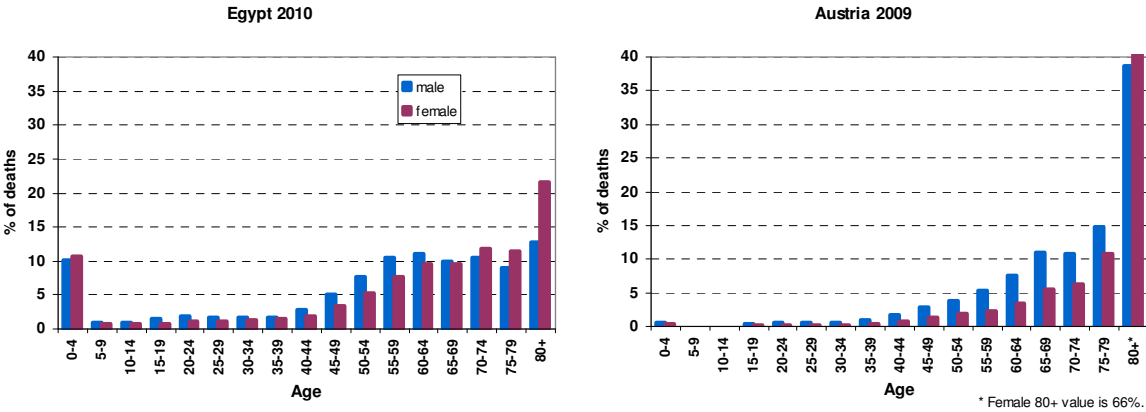


Source: Disease and injury regional estimates - Cause-specific mortality: regional estimates for 2008. Geneva, WHO Department of Health statistics and informatics, 2011 ([www.who.int/healthinfo/global\\_burden\\_disease/estimates\\_regional](http://www.who.int/healthinfo/global_burden_disease/estimates_regional)).

Significant departures from these model age-distributions of deaths suggest that the reporting of deaths by age is selectively biased. One reason for such bias may be the way age at death is reported. For example, people tend to have a strong preference to report age at death as a number ending in 0 or 5 (e.g. 45, 50, 55). This is commonly known as *digit preference* or *age-heaping*. In other instances, the age of the deceased person may be misreported; it is common for families to report that the deceased person was older than was actually the case. This highlights the importance of checking the plausibility of age-patterns of mortality, and to test for underreporting of deaths in certain age groups by plotting the graph of  $\ln(m_x)$  versus age ( $x$ ), as described above.

An example of the application of this check on data quality is shown in Figure 7 which gives the reported age distributions of deaths calculated from civil registration data for Egypt, and for Austria 2009. Egypt is a lower middle income country and Austria a high income country. The comparison shows that the age distribution of deaths in Egypt and Austria is very similar to what was expected.

**Figure 7: Age distribution of reported deaths in Egypt 2010 and Austria 2009**



Any distortions in mortality patterns compared with the corresponding income group chart may be due to poor recording of age at death, and should be investigated.

**Summary of step 4**

- The tool compares the age and sex distribution of your reported deaths with expected age-sex distributions based on the distribution of estimated deaths by income group as shown in Figure 7. Departures from these expected patterns can be indicative of underreporting of deaths at certain ages for males or females.

**Step 5 Child mortality rates**

Mortality among children under five years old, more than any other age group, reflects a range of economic, social and health conditions that affect the health of populations. This indicator thus serves as a key indicator for public health monitoring. Mortality in children under five can be divided into several components.

- Infant mortality – mortality among infants aged less than one year
- Under-5 mortality – mortality among children aged less than 5 years old.

The objectives of step 5 are to enable users to:

- Define and calculate indicators of under-five mortality
- Understand the public health relevance of measures of under -five mortality
- Interpret the indicators of under five-mortality and understand their limitations
- Use under-five mortality indicators from various sources to analyze the quality of mortality data

## Definition and calculation of under five mortality indicators

### Infant mortality rate (IMR) and Under-5 mortality rate (U5MR)

The infant mortality rate (IMR) is the probability (expressed as a rate per 1,000 live births) of a child born in a specified year dying before reaching the age of one if subject to current age-specific mortality rates. Likewise, the under-five mortality rate (U5MR) is the probability (expressed as a rate per 1,000 live births) of a child born in a specified year dying before reaching the age of five if subject to current age-specific mortality rates. Thus, the IMR and U5MR are, strictly speaking, not rates (i.e. the number of deaths divided by the number of population at risk during a certain period of time) but a probability of death, expressed as a rate per 1,000 live births.

There is a method for calculating the probability of a child dying between birth and age 1 (written as  ${}_1q_0$ ) and child dying between birth and age 5 (written as  ${}_5q_0$ ), from data on the ASMR at age 0 (defined as deaths at age 0 divided by mid-year population at age 0, and written  ${}_1m_0$ ) and at age 1-4 (defined as deaths at age 1-4 divided by mid-year population at ages 1-4, and written  ${}_4m_1$ ). Specifically,

$${}_1q_0 = {}_1m_0 / [1 + (1 - a) {}_1m_0] \text{ where } a = \text{Fraction of last year of life}$$

$$= 0.1 \text{ for high income countries and } a = 0.3 \text{ for other income group}$$

$${}_5q_0 = 1 - (1 - {}_1q_0)(1 - {}_4q_1) \text{ where } {}_4q_1 = 4 * {}_4m_1 / [1 + 4 (1 - 0.4) {}_4m_1]$$

An example of the calculation of the U5MR, IMR from civil registration data is given below:

**Table 1: Child deaths by age and calculation of mortality indicators: Norway 2009**

$x$ beginning of the age interval	$n$ number of years in the interval	Population	Deaths	${}_n m_x$	${}_n q_x$
0	1	110892	278	0.00251	0.00250
1	4	431711	86	0.00020	0.00080

$$\Rightarrow \text{IMR} = 1000 * 0.00250 = 2.5$$

$$\Rightarrow \text{U5MR} = 1000 * 1 - (1 - 0.00250)(1 - 0.00080) = 3.3$$

These calculations are performed automatically in the accompanying electronic tool.

## Sources of data on under-five mortality

In principle, the civil registration system can generate annual data on under-five mortality at both national and subnational levels, on a continuous basis. Where civil registration systems are complete, age-specific mortality rates among children and infants can be calculated directly from the number of deaths by age registered and population. However, the coverage and quality of civil registration systems is often questionable in developing countries and the resulting vital statistics may be incomplete and biased.

There are particular reasons why deaths occurring in young children are less likely to be registered than deaths in adulthood. In settings where civil registration is not universal, deaths are generally only registered when

there are some benefits attached to doing so, for example to claim land ownership and inheritance or to claim compensation by the dependants. Registering the death of a child is not usually linked to such a benefit and as a result many such deaths remain unregistered. In such settings, data on infant and child mortality estimated from censuses and surveys tend to be more reliable.

In countries with incomplete registration systems the decennial census can be used to generate estimates of child mortality using direct or indirect techniques.<sup>10</sup> The direct method involves questions to respondents about deaths in the household over a specified period of time. More commonly, an indirect method is used based on questions to female respondents on children ever born and children that are still alive. Brass-type methods and model life tables are then used to obtain an estimate of under-five mortality.<sup>11</sup> However, the census is, by definition, an infrequent occurrence (usually every 10 years) so that it is not a good source of data for ongoing monitoring. It does, however, serve a very useful function of providing an alternate source that can be used to validate data from vital registration on the number of child deaths registered and hence the level of child mortality.

In most developing countries, household surveys provide the most common source of data on child mortality using both direct and indirect methods. The indirect method asks questions about children ever born and children still alive, as for the census. The direct method involves taking a detailed birth history for each birth that a woman has had during her lifetime. These birth histories are then converted to rates of child mortality corresponding to a particular period in time.

### **Interpreting different estimates of the under-five mortality rate**

Most countries have data on child mortality from multiple sources, including the civil registration system, censuses, household surveys and the routine health information system. In this section, we show how information from censuses and surveys can be used to assess the completeness of child mortality reporting by the civil registration system.

Large differences between the levels of under-five mortality rates calculated from the reported data and other sources, such as censuses, household surveys or estimates developed by UN international agencies are likely to be due to underreporting of child deaths in the country.

Figure 8 shows data on under-five mortality for Kazakhstan. The data are derived from various data sources, including censuses, surveys and the civil registration system. This visual display of data from different sources clearly shows the extent to which the under-five mortality rates derived from civil registration appear to be in general relatively lower than those derived from the census or household surveys, especially during the earlier periods. This is indicative of substantial under-reporting of deaths of children under five in the civil registration system. By comparing the line of best “fit” for the estimated under-five mortality rate derived from censuses and surveys with observed values calculated from the civil registration system for the same year(s) (shown as “diamonds” in Figure 8), it is possible to estimate the completeness of civil registration of child deaths by comparing the distance of the vital registration estimate (“diamonds”) from the solid line, year by year.

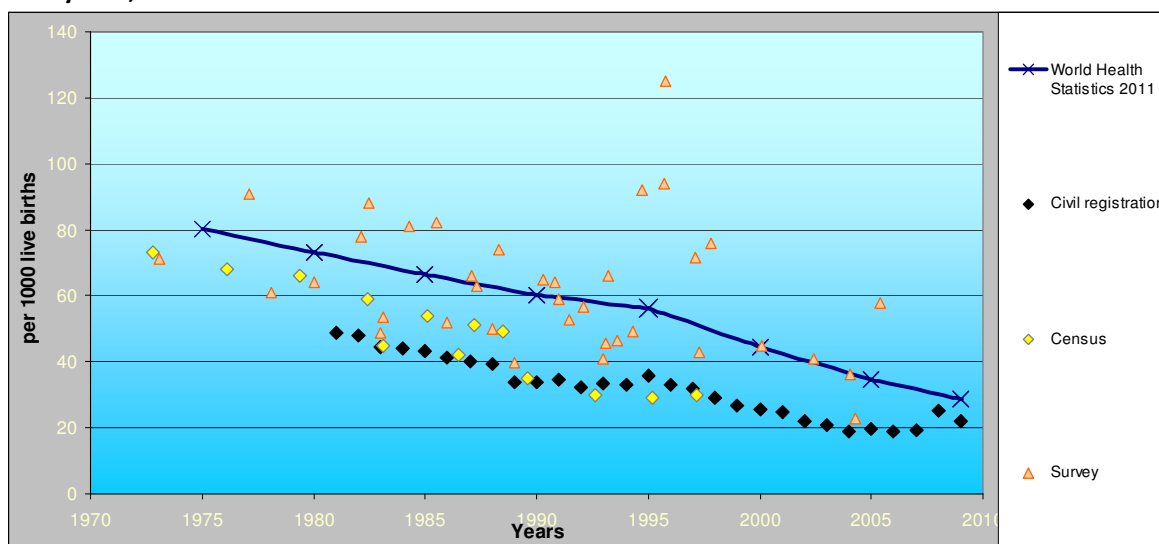
---

<sup>10</sup> United Nations Population Division (2001) Principles and recommendations for a vital statistics system, revision 2. Department of Economic and Social Affairs, Statistics Division, ST/ESA/STAT/SER.M/19/Rev2  
[http://unstats.un.org/unsd/publication/SeriesM/SeriesM\\_19rev2E.pdf](http://unstats.un.org/unsd/publication/SeriesM/SeriesM_19rev2E.pdf)

<sup>11</sup> United Nations Population Division. Manual X: indirect demographic estimation. New York: United Nations, 1983.

From this analysis it can be concluded that under-five deaths in Kazakhstan and were grossly under-reported in the national civil registration system during the 1970s and 1980s. However, levels of reporting appear to have improved dramatically in the most recent decade (the trend in the “diamonds” for Kazakhstan is getting closer and closer to the solid line of best fit for the true level of the child mortality rate). Underreporting of under-five mortality in Kazakhstan appears to have diminished significantly in recent years.

**Figure 8: Observed (from civil registration, census, surveys) and estimated levels of the under-five mortality rate, Kazakhstan 1970-2010**



Users should produce similar figures for their country or populations with death registration, bringing together on one graph estimates of under-five mortality derived from difference sources including civil registration, to help interpret the multiple data points and diagnose possible incompleteness levels in death registration. To facilitate this, users can refer to the UNICEF/WHO Child Mortality Estimation database that brings together available data sets from different sources on a country-by-country basis and presents the information in tables and figures.<sup>12</sup>

### Direct measures of incompleteness of death reporting

Special studies can also be carried out to determine the extent of under-reporting of deaths. The most widely used of these so called “direct” methods are “capture-recapture” studies<sup>13</sup> where deaths reported in the civil registration system for a sample of the population are compared (on a case-by-case basis) with deaths “captured” in an independent<sup>14</sup> survey of the same population.

While not all countries will have the technical and financial resources to carry out capture-recapture studies, we wanted here to highlight the fact that underreporting of deaths is likely to be much higher among children

<sup>12</sup> <http://www.childmortality.org/cmeMain.html>

<sup>13</sup> Sekara C, Deming W (1949) On a method of estimating birth and death rates and the extent of registration. *Journal of the American Statistical Association*, 44(245): 101-15.

<sup>14</sup> “Independence” as applied to capture-recapture studies means that the probability of a death not being reported under the civil registration system is not related to – is independent of – the probability that the same death will not be reported in another system or survey. In practice, this is very difficult to achieve.



than adults, and hence special attention should be paid to evaluating probable levels of underreporting of child deaths using the methods proposed in this section.

### Summary of step 5

- The tool calculates under-five and infant mortality rates.
- The tool shows in one chart, estimates over the past 20 to 30 years of the probability of dying before age 5 ( ${}_5q_0$ ) from different sources, including civil registration, the census, household surveys and other studies, as shown in Figure 8. Use the results to estimate the likely degree of underreporting of deaths in children under five in the civil registration system by comparing levels with those estimated from censuses or surveys.

## Step 6 Distribution of major causes of death according to the Global Burden of Disease Cause list<sup>15</sup>

A first step in any quality assessment of cause of death data is to calculate the percentage distribution of deaths by broad disease groups and compare the results with what would be expected given the level of life expectancy for the population. These expected patterns have been developed by demographers and epidemiologists on the basis of many years of data and observations on patterns of causes of death in different settings. Any significant deviation from the expected pattern that cannot be explained by some local, external factor should be viewed as a potential problem with the quality of the cause of death data.

The ICD contains over 10,000 codes that are possible causes of death. All of these causes can be further condensed into three very broad groups of causes of death as per the Global Burden of Disease cause list:

- Group I<sup>16</sup>: Communicable diseases (e.g. TB, pneumonia, diarrhoea, malaria, measles), maternal and perinatal causes (e.g. maternal haemorrhage, birth trauma) and nutritional conditions (e.g. protein-energy malnutrition)
- Group II<sup>17</sup>: Non-communicable diseases (e.g. cancer, diabetes, heart disease, stroke)
- Group III<sup>18</sup>: External causes of mortality (e.g. accidents, homicide, suicide)

The expected percentage distribution of causes of death into these three broad groups varies in different countries according to where they stand in relation to the so called “health transition” – an interrelated set of changes in demographic structures, patterns of disease and risk factors. Demographic changes include lower mortality rates among children under five years and declining fertility rates, which result in an ageing population. Epidemiological changes include a shift in the main causes of death and disease away from infectious diseases, such as diarrhoea and pneumonia (diseases traditionally associated with poorer countries), towards non-communicable diseases such as cardiovascular disease, stroke and cancers. Changes in patterns of risk include declines in risk factors for infectious diseases (undernutrition, unsafe water and poor

---

<sup>15</sup> The Global Burden of Disease: 2004 Update, World Health Organization 2008.

<sup>16</sup> ICD-10:A00-B99, G00-G04, N70-N73, J00-J06, J10-J18, J20-J22, H65-H66, O00-O99, P00-P96, E00-E02, E40-E46, E50, D50-D53, D64.9, E51-64

<sup>17</sup> ICD-10:C00-C97, D00-D48, D55-D64 (minus D 64.9) D65-D89, E03-E07, E10-E16, E20-E34, E65-E88, F01-F99, G06-G98, H00-H61, H68-H93, I00—I99, J30—J98, K00-K92, N00-N64, N75-N98, L00-L98, M00-M99, Q00-Q99

<sup>18</sup> ICD-10: V01-Y89

sanitation, for example) and increases in risk factors for chronic diseases (such as overweight, and use of alcohol and tobacco). Thus, a simple but effective way of checking the plausibility of mortality data is to compare the observed patterns of causes of death with what would be expected given the local levels of life expectancy. As a general rule, countries with low life expectancy are characterised by high levels of mortality due to infectious and parasitic diseases especially in childhood, along with high maternal mortality (i.e. Group I causes). As life expectancy rises, the pattern of mortality changes, with more deaths occurring in older age groups due to non-communicable conditions such as cardiovascular diseases and cancers (i.e. Group II causes).

Table 2 shows how the percentage of deaths assigned to various causes in each of Groups I, II and III is expected to change as life expectancy increases. Thus, a country with an average life expectancy of 55 years would typically have about 22 % of deaths due to Group I causes of death, and 65 % due to Group II causes. A country with higher life expectancy of 65 years, would typically have a smaller percentage of deaths due to Group I conditions (around 13%) and correspondingly more deaths due to Group II conditions (74%).

**Table 2: Expected distribution of causes of death according to life expectancy by broad groups**

Life Expectancy	55 years	60 years	65 years	70 years
<b>Group I causes of death</b>	22%	16%	13%	11%
<b>Group II causes of death</b>	65%	70%	74%	78%
<b>Group III causes of death</b>	13%	14%	13%	11%

Note that these are model-based percentage distributions derived from the large database held by WHO on causes of death and mortality rates. It is unlikely that any country would fit exactly these proportions, but significant departures from them, suggests potential problems with the certification and/or coding of causes of deaths.

Users should review their most recent available data on causes of death data and calculate the distribution by broad groups of causes . They should also calculate the number of deaths coded as due to ill-defined causes (ICD-10 Chapter XVIII Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified, codes R00-R99). When the proportion of those ill-defined causes to total deaths is substantially high (> 20%) , care should be taken in using such data. If proportion of the the ill-defined causes is low, users should redistribute them proportionately to causes in Group I and II.

The findings can then be compared with the expected distribution in Table 2 according to the average life expectancy in the country. However, in doing this comparison it is important to use an independent source of life expectancy data (e.g. WHO, or the United Nations, or from your census), not the life expectancy calculated from the civil registration data as this may be unreliable if the system is incomplete.

### Summary of step 6

- The spreadsheet tabulates your data on cause of death by age, sex, and broad cause of death (Groups I, II and III).
- It also calculates the percentage distribution of deaths by broad cause groups (Groups I, II and III) and the proportion of the ill-defined causes of death. It redistributes the number of deaths with unknown age proportionately over the defined age-groups and number of deaths due to ill-defined causes proportionately to causes in Group I and II. .

## Step 7 Age pattern of broad groups of causes of death

All leading causes of death in a population follow a predictable age pattern that has been identified from decades of epidemiological research. The next step is to check whether the age-pattern of deaths from broad causes is consistent with what one would expect from epidemiological research and modelling. These age-patterns do not change very much with increasing life expectancy. The Figure 9 shows a typical distribution of deaths across Groups I, II & III at different ages for a country (Venezuela) with a life expectancy of around 70 years.<sup>19</sup> At each age, the graph shows the expected proportion (fraction) of deaths at that age that are likely to occur on average; the three fractions, at any age, sum to 100%.

**Figure 9: Distribution of Broad Cause of Death Groups (I, II and III) by age**

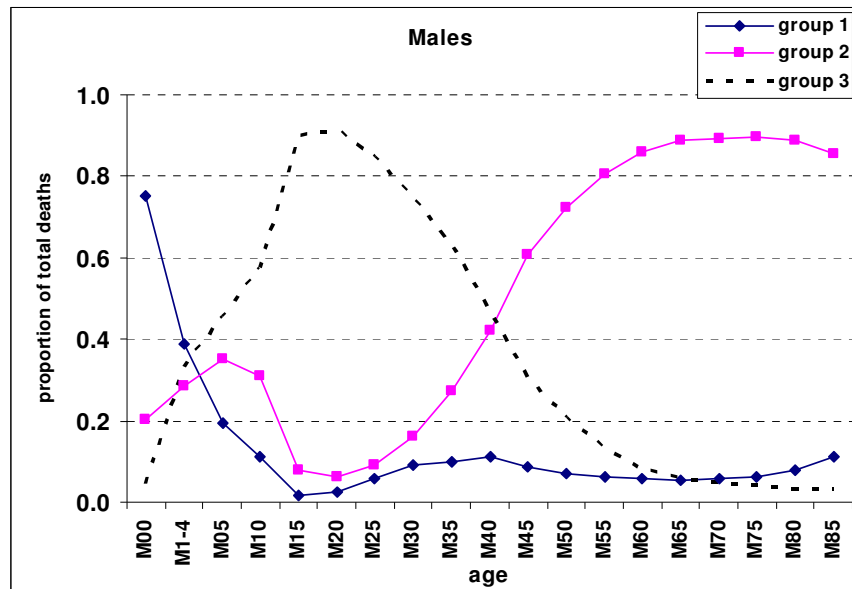


Figure 9 shows a commonly found pattern of distribution of causes of death by age in settings with relatively high life expectancy..

The proportion of deaths due to **Group I** causes (communicable diseases, maternal, perinatal and nutritional conditions) is high among children but declines thereafter to very low levels although it may rise again at older ages (above 80 years or thereabouts) due to pneumonia.

The proportion of deaths due to **Group II** causes is relatively high in children (due to congenital anomalies for e.g.), declines in adulthood, but rises significantly at older ages due to the increasing incidence of cancers and cardiovascular diseases and stroke.

The proportion of deaths due to **Group III** causes, i.e. external causes of death including accidents and violence, is generally highest in young adulthood. This pattern is especially marked among males.

This is a typical cause of death pattern by age and would not be replicated exactly in every country. However, significant departures from this pattern should be closely investigated as they are suggestive of problems such as poor medical certification of the cause of death and coding practices and age-misreporting of deaths.

<sup>19</sup> World Health Organization mortality database, <http://www.who.int/healthinfo/morttables/en/index.html>

27 June 2011

In general, the charts for males and females should be broadly similar although there is often higher mortality due to external causes among young males, while young women may have high mortality due to maternal causes (which would increase the fraction from Group I causes).

The principal reason for carrying out this step is to identify any serious biases in the data. Depending on the data source or socio-cultural context, there are strong tendencies to avoid certifying deaths to infectious diseases for e.g.. This check will help to identify the extent of these biases in your data.

### **Summary of step 7**

- The spreadsheet plots the cause of death patterns by sex and age group and allows you to compare your findings with the typical patterns according to the country income-level.

## **Step 8      Leading causes of death**

An analysis of the leading causes of death can also indicate certain biases in the certification and coding of causes of death thus indicating the degree of reliability of the cause of death data from the civil registration system. The percentage distribution of leading causes (by specific disease groups) globally, and in low-income, middle-income and high-income countries (using World Bank definitions) are shown in the spreadsheet. These charts can assist countries to ascertain divergences in their reported leading causes of death compared with leading causes of death estimated by WHO, These global estimates refer to the average experience of all countries in each of the country groups and hence it is unlikely that the percentage distribution of deaths in any one country would exactly match them. However, significant departures from these average rankings of leading causes of death are suggestive of problems with the quality of cause of death data.

Another group of causes which is suggestive of the problems with the quality of data is the category of ill-defined causes (ICD-10 Chapter XVIII). Countries should include this category in their rankings in order to see how frequently these causes are coded. In many cases, ill-defined causes may be in the top 3 or 4 leading causes of death. This suggests serious problems with certification and/or coding in the country. These ill-defined causes, while unfortunately commonly reported, are of absolutely NO value for informing public health policies and debates in countries.

### **Summary of step 8**

- The spreadsheet calculates the leading causes of death from your data and allows you to compare the findings with the typical patterns for all ages, both sexes as shown in the same spreadsheet.

## **Step 9      Ratio of non-communicable to communicable causes of death**

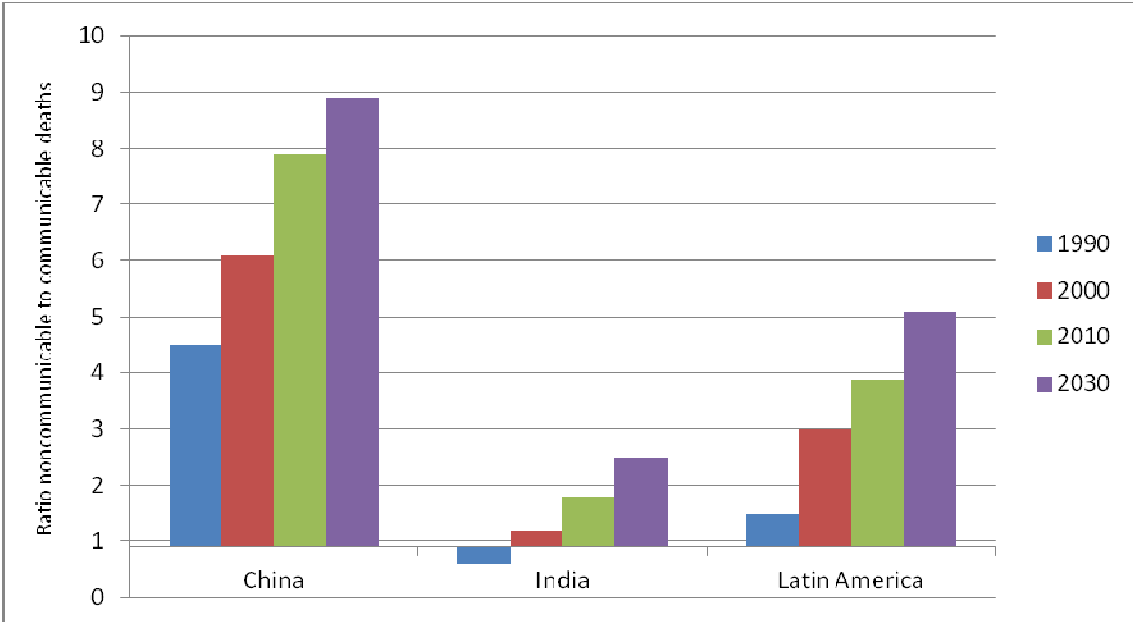
As countries develop their health systems, communicable disease such as diarrhoea and pneumonia, as well as maternal, perinatal and nutritional risks will be increasingly brought under control. As a result, more and more people will survive to adulthood where chronic diseases such as ischaemic heart disease, stroke, cancer and chronic obstructive pulmonary diseases, claim more and more lives. Hence, the simple ratio of Group II deaths/Group I deaths should progressively increase as a country moves through the epidemiological transition (i.e. as life expectancy increases).

This is illustrated in the spreadsheet where the chart shows the ratio of deaths from non-communicable diseases (Group II) to communicable diseases (Group 1) in selected World Bank income groupings (both sexes combined).<sup>20</sup> If there were the same numbers of deaths in each broad disease group, the ratio would be 1.

As is clear from the chart in the spreadsheet, globally, there are over twice as many deaths due to Group II causes as Group I causes. In high-income countries, non-communicable diseases account for 13 times as many deaths as communicable diseases. By contrast, in low-income countries, there are more deaths due to communicable diseases as the ratio is less than 1. In the upper middle-income countries, there are about five times as many deaths due to non-communicable diseases compared with communicable diseases. This reflects the fact that in high-income and upper middle-income countries, most deaths occur late in life, due to chronic conditions such as cancers and cardiovascular diseases. In low-income countries by contrast, most deaths occur in childhood, due to infectious diseases conditions such pneumonia, diarrhoea, and vaccine preventable conditions as well as perinatal causes.

Over time, as child mortality decreases and life expectancy increases, the pattern in low-income countries will start to look more like that observed in middle-income and high-income countries. This is illustrated in the Figure 10 which shows estimated trends in the ratio of non-communicable to communicable conditions in China, India and Latin America. In India in 1990, there were more deaths due to communicable diseases than to non-communicable diseases, hence the ratio is less than 1. Since 2000, however, deaths due to non-communicable diseases have exceeded those due to communicable diseases.

**Figure 10: Estimated trends in ratio of non-communicable to communicable deaths, selected regions, 1990-2020**



Departures from this overall picture are suggestive of errors in cause of death data.

<sup>20</sup> Disease and injury regional estimates - Cause-specific mortality: regional estimates for 2008. Geneva, WHO Department of Health statistics and informatics, 2011 ([http://www.who.int/healthinfo/global\\_burden\\_disease/estimates\\_regional](http://www.who.int/healthinfo/global_burden_disease/estimates_regional))

### Summary of step 9:

- The spreadsheet calculates the ratio of deaths from non-communicable diseases to communicable diseases (Group II to Group I deaths) and allows you to compare the findings to those of the most appropriate comparator group.

## Step 10 Ill-defined causes of death

When a death occurs and is medically certified, every effort should be made to correctly ascertain the underlying cause of death in order to be able to draw conclusions about the leading causes and about the need for priority public health interventions. Classification of deaths to ill-defined conditions does not generate information of public health value. Where a high proportion of all deaths is classified as being due to ill-defined causes, the cause-of-death distribution will be biased and unreliable.

At the end of this section, users should be able to:

- Define and calculate the proportion of deaths attributed to ill-defined causes of death
- Understand the implications for the overall quality of mortality statistics of a high proportion of ill-defined causes of death.
- Understand the definition and calculation of Ill-defined categories in cause of death data

When certifying causes of death, these “Ill-defined” causes should be avoided or not used as the underlying cause of death as they are vague diagnoses. These ill-defined codes arise from two sources:

- i. Deaths classified as ill-defined (Chapter XVIII of ICD-10)
- ii. Deaths classified to any one of the following vague or unspecific diagnoses:
  - A40-A41 Streptococcal and other septicaemia
  - C76, C80, C97 Ill-defined cancer sites
  - D65 Disseminated intravascular coagulation [defibrination syndrome]
  - E86 Volume depletion
  - I10 Essential (primary) hypertension
  - I269 Pulmonary embolism without mention of acute cor pulmonale
  - I46 Cardiac arrest
  - I472 Ventricular tachycardia
  - I490 Ventricular fibrillation and flutter
  - I50 Heart failure
  - I514 Myocarditis, unspecified
  - I515 Myocardial degeneration
  - I516 Cardiovascular disease, unspecified
  - I519 Heart disease, unspecified
  - I709 Generalized and unspecified atherosclerosis
  - I99 Other and unspecified disorders of circulatory system
  - J81 Pulmonary oedema
  - J96 Respiratory failure, not elsewhere classified
  - K72 Hepatic failure, not elsewhere classified
  - N17 Acute renal failure
  - N18 Chronic renal failure
  - N19 Unspecified renal failure
  - P285 Respiratory failure of newborn
  - Y10-Y34, Y872 External cause of death not specified as accidentally or purposely inflicted

Deaths classified to either of these two categories of *ill-defined diagnoses* are insufficiently detailed to be of value for public health purposes, although in the majority of cases they help to describe the overall mortality due to broad disease (e.g. CVD, respiratory disease) or injury groups. Separately identifying their frequency in cause of death tabulations is essential to decide upon remedial action to reduce their usage. This could involve interventions to improve certification practices, or coding practices, or both.

While there will always be individual cases where it is not possible to classify the cause to a specific ICD category due to lack of appropriate information, such cases should be relatively infrequent. As a general principle, the proportion of ill-defined deaths coded to either category i) or category ii) above should collectively not exceed 10% for deaths at ages 65 years and over, and should be less than 5% for deaths at ages below 65 years.

When reviewing a data series of cause of death information, it is important to study how the proportion of ill-defined causes of death has changed over time. Large fluctuations may be indicative of changes in certification and coding practices rather than real changes in patterns of mortality.

The proportion of deaths assigned to ill-defined causes tends to be higher for deaths occurring at older ages. There are many possible explanations, including the fact that many such deaths occur outside health care facilities and also because of the existence of multiple comorbidities which renders such deaths harder to correctly diagnose. Nonetheless, with good certification and coding practices, it should be possible to reduce this proportion to less than 10% of deaths among the elderly.

### **Summary of step 10**

- The spreadsheet calculates the proportion of category i) and category ii) ill-defined causes in your cause of death data by sex and age-group. . The total should not exceed 5% of deaths at ages below 65, and 10% of deaths at age 65+.

## Global Burden of Disease cause categories and ICD-10 codes

GBD cause name	ICD-10 code
<b>I. Communicable, maternal, perinatal and nutritional conditions</b>	A00-B99, G00-G04, N70-N73, J00-J06, J10-J18, J20-J22, H65-H66, O00-O99, P00-P96, E00-E02, E40-E46, E50, D50-D53, D64.9, E51-E64
<b>A. Infectious and parasitic diseases</b>	A00-B99, G00, G03-G04, N70-N73
1. <b>Tuberculosis</b>	A15-A19, B90
2. <b>Sexually transmitted diseases excluding HIV</b>	A50-A64, N70-N73
a. Syphilis	A50-A53
b. Chlamydia	A55-A56
c. Gonorrhoea	A54
Other STDs	A57-A64, N70-N73
3. <b>HIV/AIDS</b>	B20-B24
4. <b>Diarrhoeal diseases</b>	A00, A01, A03, A04, A06-A09
5. <b>Childhood-cluster diseases</b>	A33-A37, A80, B05, B91
a. Pertussis	A37
b. Poliomyelitis	A80, B91
c. Diphtheria	A36
d. Measles	B05
e. Tetanus	A33-A35
6. <b>Meningitis</b>	A39, G00, G03
7. <b>Hepatitis B</b>	B16-B19 (minus B17.1, B18.2)
<b>Hepatitis C</b>	B17.1, B18.2
8. <b>Malaria</b>	B50-B54
9. <b>Tropical-cluster diseases</b>	B55-B57, B65, B73, B74.0-B74.2
a. Trypanosomiasis	B56
b. Chagas disease	B57
c. Schistosomiasis	B65
d. Leishmaniasis	B55
e. Lymphatic filariasis	B74.0-B74.2
f. Onchocerciasis	B73
10. <b>Leprosy</b>	A30
11. <b>Dengue</b>	A90-A91
12. <b>Japanese encephalitis</b>	A83.0
13. <b>Trachoma</b>	A71
14. <b>Intestinal nematode infections</b>	B76-B81
a. Ascariasis	B77
b. Trichuriasis	B79
c. Hookworm disease (ancylostomiasis and necatoriasis)	B76
Other intestinal infections	B78, B80, B81
Other infectious diseases	A02, A05, A20-A28, A31, A32, A38, A40-A49, A65-A70, A74-A79, A81, A82, A83.1-A83.9, A84-A89, A92-A99, B00-B04, B06-B15, B25-B49, B58-B60, B64, B66-B72, B74.3-B74.9, B75, B82-B89, B92-B99, G04



<b>B. Respiratory infections</b>	J00-J06, J10-J18, J20-J22, H65-H66
1. Lower respiratory infections	J10-J18, J20-J22
2. Upper respiratory infections	J00-J06
3. Otitis media	H65-H66
<b>C. Maternal conditions</b>	O00-O99
1. Maternal haemorrhage	O44-O46, O67, O72
2. Maternal sepsis	O85-O86
3. Hypertensive disorders of pregnancy	O10-O16
4. Obstructed labour	O64-O66
5. Abortion	O00-O07
Other maternal conditions	O20-O43, O47-O63, O68-O71, O73-O75, O87-O99
<b>D. Conditions arising during the perinatal period</b>	P00-P96
1. Prematurity and low birth weight	P05, P07, P22, P27-P28
2. Birth asphyxia and birth trauma	P03, P10-P15, P20-P21, P24-P26, P29
Neonatal infections and other conditions	P00-P02, P04, P08, P23, P35-P96
<b>E. Nutritional deficiencies</b>	E00-E02, E40-E46, E50, D50-D53, D64.9, E51-E64
1. Protein-energy malnutrition	E40-E46
2. Iodine deficiency	E00-E02
3. Vitamin A deficiency	E50
4. Iron-deficiency anaemia	D50, D64.9
Other nutritional disorders	D51-D53, E51-E64
<b>II. Noncommunicable diseases</b>	C00-C97, D00-D48, D55-D64 (minus D 64.9), D65-D89, E03-E07, E10-E16, E20-E34, E65-E88, F01-F99, G06-G98, H00-H61, H68-H93, I00-I99, J30-J98, K00-K92, N00-N64, N75-N98, L00-L98, M00-M99, Q00-Q99
<b>A. Malignant neoplasms</b>	C00-C97
1. Mouth and oropharynx cancers	C00-C14
2. Oesophagus cancer	C15
3. Stomach cancer	C16
4. Colon and rectum cancers	C18-C21
5. Liver cancer	C22
6. Pancreas cancer	C25
7. Trachea, bronchus and lung cancers	C33-C34
8. Melanoma and other skin cancers	C43-C44
9. Breast cancer	C50
10. Cervix uteri cancer	C53
11. Corpus uteri cancer	C54-C55
12. Ovary cancer	C56
13. Prostate cancer	C61
14. Bladder cancer	C67
15. Lymphomas and multiple myeloma	C81-C90, C96
16. Leukaemia	C91-C95
Other malignant neoplasms	C17, C23, C24, C26-C32, C37-C41, C45-C49, C51, C52, C57-C60, C62-C66, C68-C80, C97

<b>B. Other neoplasms</b>	D00-D48
<b>C. Diabetes mellitus</b>	E10-E14
<b>D. Endocrine disorders</b>	D55-D64 (minus D64.9), D65-D89, E03-E07, E15-E16, E20-E34, E65-E88
<b>E. Neuropsychiatric conditions</b>	F01-F99, G06-G98
1. Unipolar depressive disorders	F32-F33
2. Bipolar affective disorder	F30-F31
3. Schizophrenia	F20-F29
4. Epilepsy	G40-G41
5. Alcohol use disorders	F10
6. Alzheimer and other dementias	F01, F03, G30-G31
7. Parkinson disease	G20-G21
8. Multiple sclerosis	G35
9. Drug use disorders	F11-F16, F18-F19
10. Post-traumatic stress disorder	<i>F43.1</i>
11. Obsessive-compulsive disorder	F42
12. Panic disorder	F40.0, F41.0
13. Insomnia (primary)	F51
14. Migraine	G43
Mental retardation attributable to lead exposure	F70-F79
Other neuropsychiatric disorders	F04-F09, F17, F34-F39, F401-F409, F411-F419, F43(minus F43.1), F44-F50, F52-F69, F80-F99, G06-G12, G23-G25, G36, G37, G44-G98
<b>F. Sense organ diseases</b>	H00-H61, H68-H93
1. Glaucoma	H40
2. Cataracts	H25-H26
3. Refractive errors	H524
4. Hearing loss, adult onset	H90-H91
Macular degeneration and other	H00-H21, H27-H35, H43-H61(minus H524), H68-H83, H92-H93
<b>G. Cardiovascular diseases</b>	I00-I99
1. Rheumatic heart disease	I01-I09
2. Hypertensive heart disease	I10-I13
3. Ischaemic heart disease	I20-I25
4. Cerebrovascular disease	I60-I69
5. Inflammatory heart diseases	I30-I33, I38, I40, I42
Other cardiovascular diseases	I00, I26-I28, I34-I37, I44-I51, I70-I99
<b>H. Respiratory diseases</b>	J30-J98
1. Chronic obstructive pulmonary disease	J40-J44
2. Asthma	J45-J46
Other respiratory diseases	J30-J39, J47-J98
<b>I. Digestive diseases</b>	K20-K92
1. Peptic ulcer disease	K25-K27
2. Cirrhosis of the liver	K70, K74
3. Appendicitis	K35-K37
Other digestive diseases	K20-K22, K28-K31, K38, K40-K66, K71-K73, K75-K92
<b>J. Genitourinary diseases</b>	N00-N64, N75-N98

1. Nephritis and nephrosis	N00-N19
2. Benign prostatic hypertrophy	N40
Other genitourinary system diseases	N20-N39, N41-N64, N75-N98
<b>K. Skin diseases</b>	L00-L98
<b>L. Musculoskeletal diseases</b>	M00-M99
1. Rheumatoid arthritis	M05-M06
2. Osteoarthritis	M15-M19
3. Gout	M10
4. Low back pain	M45-M48, M54 (minus M54.2)
Other musculoskeletal disorders	M00-M02, M08, M11-M13, M20-M43, M50-M53, M54.2, M55-M99
<b>M. Congenital anomalies</b>	Q00-Q99
1. Abdominal wall defect	Q79.2-Q79.5
2. Anencephaly	Q00
3. Anorectal atresia	Q42
4. Cleft lip	Q36
5. Cleft palate	Q35, Q37
6. Oesophageal atresia	Q39.0-Q39.1
7. Renal agenesis	Q60
8. Down syndrome	Q90
9. Congenital heart anomalies	Q20-Q28
10. Spina bifida	Q05
Other congenital anomalies	Q01-Q04, Q06-Q18, Q30-Q34, Q38, Q392-Q399, Q40-Q41, Q43-Q56, Q61-Q78, Q790, Q791, Q796, Q798, Q799, Q80-Q89, Q91-Q99
<b>N. Oral conditions</b>	K00-K14
1. Dental caries	K02
2. Periodontal disease	K05
3. Edentulism	—
Other oral diseases	K00, K01, K03, K04, K06-K14
<b>III. Injuries</b>	V01-Y89
<b>A. Unintentional injuries</b>	V01-X59, Y40-Y86, Y88, Y89
1. Road traffic accidents <sup>a</sup>	
2. Poisonings	X40-X49
3. Falls	W00-W19
4. Fires	X00-X09
5. Drownings	W65-W74
6. Other unintentional injuries	Rest of V, W20-W64, W75-W99, X10-X39, X50-X59, Y40-Y86, Y88, Y89
<b>B. Intentional injuries</b>	X60-Y09, Y35-Y36, Y870, Y871
1. Self-inflicted injuries	X60-X84, Y870
2. Violence	X85-Y09, Y871
3. War and conflict	Y36
Other intentional injuries	Y35

—, not available  
 STD: sexually transmitted diseases.

<sup>a</sup> For countries with 3-digit ICD10 data, for "Road traffic accidents" use: V01-V04, V06, V09-V80, V87, V89 and V99. For countries with 4-digit ICD10 data, for "Road traffic accidents" use:

27 June 2011

V01.1-V01.9, V02.1-V02.9, V03.1-V03.9, V04.1-V04.9, V06.1-V06.9, V09.2, V09.3, V10.3-V10.9, V11.3-V11.9, V12.3-V12.9, V13.3-V13.9, V14.3-V14.9, V15.4-V15.9, V16.4-V16.9, V17.4-V17.9, V18.4-V18.9, V19.4-V19.9, V20.3-V20.9, V21.3-V21.9, V22.3-V22.9, V23.3-V23.9, V24.3-V24.9, V25.3-V25.9, V26.3-V26.9, V27.3-V27.9, V28.3-V28.9, V29.4-V29.9, V30.4-V30.9, V31.4-V31.9, V32.4-V32.9, V33.4-V33.9, V34.4-V34.9, V35.4-V35.9, V36.4-V36.9, V37.4-V37.9, V38.4-V38.9, V39.4-V39.9, V40.4-V40.9, V41.4-V41.9, V42.4-V42.9, V43.4-V43.9, V44.4-V44.9, V45.4-V45.9, V46.4-V46.9, V47.4-V47.9, V48.4-V48.9, V49.4-V49.9, V50.4-V50.9, V51.4-V51.9, V52.4-V52.9, V53.4-V53.9, V54.4-V54.9, V55.4-V55.9, V56.4-V56.9, V57.4-V57.9, V58.4-V58.9, V59.4-V59.9, V60.4-V60.9, V61.4-V61.9, V62.4-V62.9, V63.4-V63.9, V64.4-V64.9, V65.4-V65.9, V66.4-V66.9, V67.4-V67.9, V68.4-V68.9, V69.4-V69.9, V70.4-V70.9, V71.4-V71.9, V72.4-V72.9, V73.4-V73.9, V74.4-V74.9, V75.4-V75.9, V76.4-V76.9, V77.4-V77.9, V78.4-V78.9, V79.4-V79.9, V80.3-V80.5, V81.1, V82.1, V82.8-V82.9, V83.0-V83.3, V84.0-V84.3, V85.0-V85.3, V86.0-V86.3, V87.0-V87.9, V89.2-V89.3, V89.9, V99 and Y850.