CRVS best-practice and advocacy

Estimating excess deaths due to COVID-19 in countries with incomplete death registration: A quick reference guide

This quick reference guide summarises approaches and methods for researchers and statisticians to estimate excess mortality due to COVID-19, in countries with incomplete death registration. For more detailed guidance on estimating excess deaths, including calculations for countries with complete death registration, see the full guide, *Approaches and methods for estimating excess deaths due to COVID-19*, on the CRVS Knowledge Gateway: https://crvsgateway.info/file/16908/4128

Introduction to excess mortality

Excess mortality, as it pertains to COVID-19, is defined as the number of deaths reported (or estimated) to have occurred during the COVID-19 pandemic, in excess of the number of deaths that would have been expected to have occurred during the same period, based on past trends. It is important to note that excess deaths are not precisely defined by the number of deaths where COVID-19 is the cause, other factors influence this. Analysis of excess mortality will provide a more comprehensive picture of the impact of COVID-19 than just relying on the reporting of COVID-19 as a cause of death (**Box 1**).

Excess mortality = reported (or estimated) deaths – expected deaths

Box 1: Cause specific or all-cause mortality as a measure for COVID-19?

Typically, public health intervention strategies require the input of accurate, cause-specific mortality data for major diseases or injuries (such as lung cancer or road traffic accidents) to assess and monitor impact. Obtaining cause-specific data to measure mortality due to COVID-19, however, has been challenging for several key reasons:

- Diagnostic uncertainty (stemming from limited experience certifying and coding these deaths, and the presence of co-morbidities at older ages creating difficulty understanding the causal chain leading to death) could lead to a miscount of COVID-19 deaths
- Categorial attribution of a COVID-19 infection does not necessarily tell a complete public health story. COVID-19 may be impacting mortality in both positive (e.g. a reduction in road traffic accidents) and, likely predominantly, negative ways (e.g. infection increasing mortality risk from other conditions), with the net contribution of this yet to be understood
- The increased strain on health care services during the pandemic may lead to increased mortality from other conditions or injuries
- The difficulty of diagnosing COVID-19 deaths outside of hospitals presents a significant barrier to obtaining accurate cause-specific data.

Due to the complexities inherent in obtaining accurate cause-specific data, all-cause mortality is a much better way of assessing the impact of COVID-19, as it captures the net effect of all factors that may increase or decrease mortality, and is readily measurable from existing CRVS systems. In order to measure all-cause mortality, an estimation of excess deaths within a population is required.

A timely and complete civil registration and vital statistics (CRVS) system is the gold standard for measurement of mortality in a population. The advantage of a CRVS system to measure excess mortality is that, when complete, it will capture all deaths in a population, irrespective of whether the deaths occurred in hospital or in the community. However, death registration in many countries is incomplete, which presents challenges when measuring excess mortality. **While measuring excess deaths is easier in countries where the CRVS system has captured most (or all) deaths in a timely and accurate manner, it is still possible to produce reliable estimates even with incomplete registration.**

Measuring excess deaths in countries with incomplete CRVS data

The measurement of excess mortality, for countries with death registration completeness of less than 90 per cent,[1] relies upon the following data:

1. Deaths *expected* to have occurred in the absence of the pandemic. These are calculated based on historical data.

AND

2. Deaths *estimated* to have occurred (the best estimate of the actual number of deaths that occurred, adjusting for incomplete death registration).

The calculations to measure excess mortality are presented below, for each of the two levels of incomplete death registration (medium and high). However, the use of incomplete death registration data to estimate excess deaths also presents a number of challenges (**Box 2**).

Box 2: Key considerations when using incomplete CRVS data to estimate completeness

Before calculating excess deaths, there are some considerations which will affect the accuracy of the estimation of excess deaths:

- *How complete are the CRVS data?* Use of incomplete death registration data to calculate excess deaths may under-estimate the true number of excess deaths.
- *Is the completeness of CRVS data during the COVID-19 pandemic different from the normal level of completeness?* Registration activities may be adversely affected because of the pandemic, or there may have been recent investments to increase registration, which may thus support increases in registration completeness.
- *Are the excess deaths due to the COVID-19 pandemic more or less likely to be captured by the death registration system?* This consideration focuses on whether the additional deaths due to the pandemic have characteristics that mean they are more or less likely to be registered. For example, if the excess deaths are disproportionately occurring in hospitals, and hospital deaths are more likely to be registered than community deaths, then it is important to account for this.

Measuring completeness of registration

Before calculating excess deaths, it is important measure the completeness of your country's death registration, because the completeness level will inform the most appropriate calculation to use and the accuracy of the subsequent measurement.

Empirical completeness method

The empirical completeness method measures completeness of death registration (or death notification/reporting) using the following data inputs, which are readily available at the national and subnational level:

- Registered crude death rate: This is the number of deaths divided by the population multiplied by 1000
- Under-five mortality rate (*5q0*): This is the best estimate of the true 5q0 and represents the level of mortality
- Percentage of the population aged 65 and over: This represents the population age structure
- Completeness of under-five death registration: This is the 5q0 calculated from the registration data, divided by the true *5q0*.

Below is an example of the empirical completeness method for a hypothetical country in 2014. The input data are shown in **Table 1**.

1 90 per cent completeness is used as a cut-off for complete registration in the context of COVID-19 because this will be higher for adults and the elderly, ages where almost all excess mortality occurs.

Table 1: Input data to estimate completeness of registration

Registered CDR	% population 65+	5q0	Under-5 completeness	Year
3.077	6.60%	0.0181	45%	2014

The equation used in the empirical completeness method to calculate completeness for both sexes, based on the input data (**Table 1**), follows. Bolded figures are the model coefficients, which are the same for all populations.

logit (c_{jk}^{All}) = (**-0.0177** * 3.0772) + (**0.6375** * 3.077) + (**-13.8914** * 0.0660) + (**-1.1136** * ln(0.0181)) + (**2.2063** * 0.45) + (**-0.0174** * 2014) + **29.3677** = 0.6617

$\frac{e^{0.6617}}{e^{0.6617}+1}$ = 66.0%

That is, completeness in 2014 is estimated to be 66 per cent.

A simpler way of measuring completeness

A more straightforward method of estimating death registration completeness to obtain an estimate of total deaths from a reliable source (such as those produced by the Global Burden of Disease [GBD] study, United Nations World Population Prospects [UN WPP], or your country's national statistical office estimates), uses this equation:

Completeness = registered deaths ÷ estimated total deaths × 100

It should be kept in mind that not all other reliable sources of mortality data will provide estimates of deaths for subnational areas; the national CDR (crude death rate) may not be appropriate to use for a subnational area.

Measuring excess deaths from a system with medium registration completeness (50 to 90 per cent)

Calculating expected deaths

The calculation of expected deaths should be based on recent historical data; where possible this should be for the last five years. The following steps are applied:

Step 1
Before estimating expected deaths, it is important to check each year of data for any mortality "shocks", such as other epidemics or natural disasters, which might not represent long-term mortality trends. If these exist, they should be removed from the data.

Step 2
Use the trend line of the deaths in the historical data to estimate expected deaths in the period of interest. You can use the FORECAST function in Microsoft Excel, or your own statistical time series model. Note that the use of average number of deaths in the last five years may bias the number of expected deaths because it does not account for the trend in these indicators. An alternative is to use the CDR (deaths divided by population multiplied by 1000) as the basis for calculating expected deaths, if the rate of change in population size has varied in the last five years.

Specific to countries with incomplete registration, some additional steps need to be undertaken:

- Calculate completeness of registration for each year
- If you are using the empirical completeness method, estimated deaths for each year need to be calculated:
  Estimated deaths = registered deaths ÷ completeness (as a fraction)
- If you have estimated total deaths for each year from a reliable source, use that in your historical data. Completeness will still need to be calculated (i.e. registered deaths divided by estimated total deaths) to adjust deaths at a later stage. Constant completeness over time cannot always be assumed.

Once the estimated deaths for each year from historical data have been calculated, deaths in the year and period of interest can be calculated. **Table 2** presents an example of estimating excess deaths where completeness of registration is 70 per cent. In this example, completeness has remained steady at 70 per cent during the 2015 to 2019 period. In 2020, 18 141 deaths are expected to occur.

Table 2: Example of estimating expected deaths, medium completeness of death registration

Year	Registered deaths (1)	Completeness (2)	Estimated total deaths (1 ÷ 2)	Population
2015	11,513	70%	16,447	2,349,592
2016	11,744	70%	16,777	2,382,486
2017	11,919	70%	17,027	2,415,841
2018	12,201	70%	17,430	2,449,663
2019	12,493	70%	17,847	2,483,958
			Expected deaths	
2020 (projected)			**18,141**	**2,518,734**

Expected deaths for a period of time within a calendar year can be calculated based on the annual estimates. This is straightforward once the period of interest is determined.

From historical data, calculate the proportion of registered deaths that occur in the period of interest, and multiply this proportion by the expected deaths for 2020. Using the data from **Table 2**, if 17 per cent of deaths each year occur in March and April, then:

Expected deaths = 17% × 18,141

= **3,084 expected deaths in March to April 2020**

Calculating excess deaths

Excess deaths = estimated deaths – expected deaths

Table 3 presents an example of estimating the number of deaths to have occurred during March to April 2020. Because completeness was steady at 70 per cent from 2015 to 2019, this is assumed to be the level of completeness in 2020. However, this figure may be adjusted if there is evidence that registration completeness may be higher or lower during the pandemic for the reasons discussed earlier. In the following example, 2768 deaths were registered in March to April 2020.

Table 3: Example 1 of estimating deaths to have occurred during March-April 2020, medium completeness of death registration

	Registered deaths (1)	Completeness (2)	Estimated Total Deaths (1 ÷ 2)	Expected Deaths (from above)
Total	2,768	70%	3,954	3,084

Using the data from **Table 3**:

Excess deaths = 3,954 – 3,084

= **870 excess deaths, or 28 per cent higher than expected**

Measuring excess deaths from a system with low registration completeness (less than 50 per cent)

Calculating expected deaths

In populations with low death registration completeness, it is not recommended to estimate total deaths from the estimate of completeness using the empirical completeness method. It is instead recommended to use estimated total deaths from a reliable data source, such as the GBD, UN WPP or national statistics office. Completeness should still be calculated from the estimated total deaths (registered deaths divided by estimated total deaths).

The process for estimating expected deaths is otherwise the same as for countries with medium completeness.

Calculating excess deaths

Excess deaths = estimated deaths – expected deaths

Table 5 presents an example of estimating deaths to have occurred during March to April 2020 where completeness of death registration is 35 per cent. This assumes that there were 7250 expected deaths estimated for 2020 (based on historical data), which results in 1233 expected deaths in March to April (multiplying 7250 [expected deaths] by 17 per cent [proportion of deaths occurring in period of interest]). Where there are 561 registered deaths, the estimated total deaths during March to April 2020 is 1603.

Table 5: Example of estimating deaths to have occurred during March-April 2020, low completeness of death registration

	Registered deaths (1)	Completeness (2)	Estimated Total Deaths (1 ÷ 2)	Expected Deaths (from above)
Total	561	35%	1,603	1,233

Using the data from **Table 5**:

Excess deaths = 1,603 – 1,233

= **380, or 30 per cent higher than expected**

Understanding the characteristics of deaths during the pandemic and excess deaths

It is good practice to compare the characteristics of registered deaths during the COVID-19 pandemic with your historical data. This may tell you much about the excess mortality occurring during the pandemic, especially if registration completeness is high. Characteristics of the deaths and decedents can include causes of death, age group of death, and place of death, among others. To do this, you can either:

1. Compare the percentage of all registered deaths during the pandemic and the historical data that are due to each characteristic; or,

2. a) Calculate the number of additional registered deaths in the pandemic (which is excess deaths where registration is complete) compared with the average number for the same time period in the historical data.

   b) For each characteristic, calculate the difference between registered deaths in the pandemic and average registered deaths in the same time period in the historical data.

   c) For each characteristic, calculate b) divided by a). This will provide an estimate of the percentage of additional registered deaths in each characteristic, whether it be cause of death, age group, or place of death.

Conclusion

Reliably measuring excess deaths can provide a comprehensive and timely estimate of the true overall impact of the COVID-19 pandemic on a population. While measurement is easier when a CRVS system has captured most (or all) deaths in a timely and accurate manner, it is still possible to estimate excess deaths even in countries with incomplete registration.

Bloomberg Philanthropies DATA FOR HEALTH INITIATIVE

Australian Government
Department of Foreign Affairs and Trade

The program partners on this initiative include: The University of Melbourne, Australia; CDC Foundation, USA; Vital Strategies, USA; Johns Hopkins Bloomberg School of Public Health, USA; World Health Organization, Switzerland.

Civil Registration and Vital Statistics partners:

Vital Strategies

CDC Foundation
Together our impact is greater

THE UNIVERSITY OF MELBOURNE

For more information contact:

CRVS-info@unimelb.edu.au
crvsgateway.info

CRICOS Provider Code: 00116K

Version: 0920-01