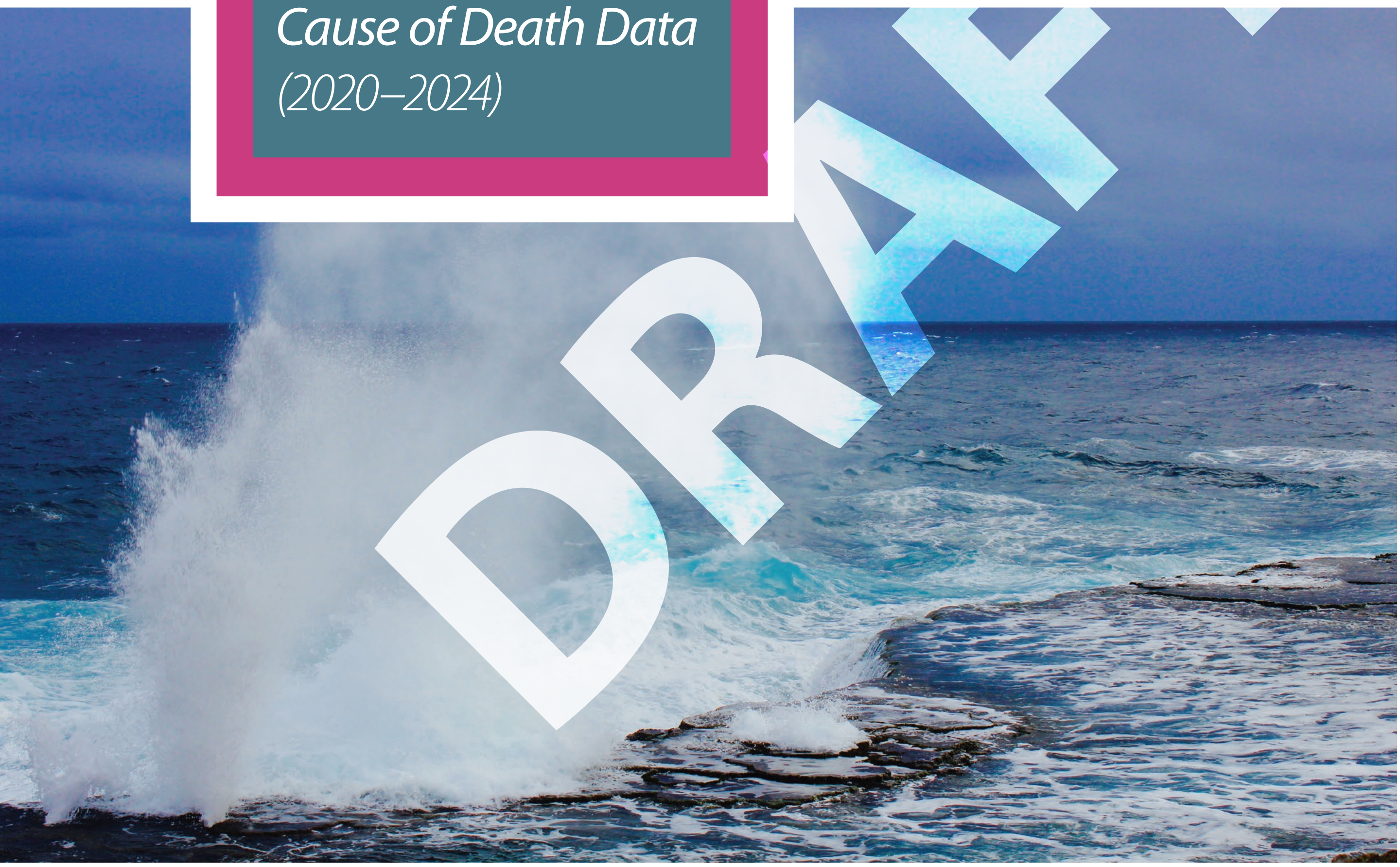
Pacific Community
Communauté du Pacifique

Making Death Count

Tonga: *Assessment of the Quality of Cause of Death Data (2020–2024)*



Making Death Count

Tonga: *Assessment of the Quality of Cause of Death Data (2020–2024)*

Pacific Community
Communauté du Pacifique

Noumea, New Caledonia
February 2026

Original text: English

Pacific Community Cataloguing-in-publication data

Design and layout by Gaëlle Le Gall-Queguineur
Cover picture credit: Flickr, MW

Prepared for publication at SPC's headquarters, B.P. D5, 98848
Noumea Cedex, New Caledonia, 2025
www.spc.int | spc@spc.int

CONTENTS

FIGURES

TABLES

ACRONYMS AND ABBREVIATIONS

ANACoD3	Analysis of Mortality and Cause of Death, 3rd version (electronic tool)
ASMR	age-specific mortality rate
CDR	crude death rate
CRVS	Civil Registration and Vital Statistics system
GBD	Global Burden of Disease
ICD-10	International Statistical Classification of Diseases and Related Health Problems, 10th revision
IGME	Interagency Group for Child Mortality Estimation
IHME	Institute of Health Metrics and Evaluation
IMR	infant mortality rate
MOH	Ministry of Health
NCD	non-communicable disease
NEC	not elsewhere classified
TSD	Tonga Statistics Department
U5MR	under-five mortality rate
UN	United Nations
WHO	World Health Organization

ACKNOWLEDGMENT

This report was prepared by Nicola Richards (Independent Consultant, Pacific Community), in collaboration with Sioape Kupu, Tifa Halafihi, Eniselina Noema, and Mioko Vailofa (Ministry of Health).

Authors are grateful to national CRVS partners, particularly Mele Manitisa (Tonga Statistics Department) and Liniola Manu (Ministry of Justice), and to regional and national development partners, particularly the Pacific Community (SPC), whose contributions during two national capacity building workshops informed the content of this report.

This work was supported by SPC as part of the Bloomberg Philanthropies Data for Health Initiative.

……
…………

The report is issued without formal editing.

KEY TERMS AND DEFINITIONS

Age-specific mortality rate: A mortality rate limited to a particular age group. The numerator is the number of deaths in that age group; the denominator is the number of persons in that age group in the population.

ANACoD3: An electronic tool that assesses the accuracy and completeness of mortality and cause of death data. It checks for potential errors and inconsistencies in the data and provides users with an understanding of basic epidemiological and demographic concepts to interpret their data.

Cause of death: Refers to 'all those diseases, morbid conditions or injuries which either resulted in or contributed to death and the circumstance of the accident or violence which produced any such injuries'.[1]

Completeness (of death reporting): the number of reported deaths divided by an estimate of the total number of actual deaths that occurred in the population during a specified period (usually one year).

$$Completeness\ \ of\ \ death\ \ registration\ \left(\%\right) = \frac{Number\ \ of\ \ registered\ \ deaths}{Expected\ \ number\ \ of\ \ deaths} \times 100$$

Crude death rate: The number of deaths relative to the size of that population during a given period, usually one year. It is expressed in units of deaths per 1000 population per year.

Garbage code: Any code that cannot or should not be used for the underlying cause of death. For instance, a "mode of death" such as, heart failure, kidney failure, etc. or symptoms such as backpain, depression, or risk factors such as high blood pressure, are garbage codes.

Ill-defined cause of death: Deaths classified as Symptoms, signs or clinical findings, not elsewhere classified (ICD-10 chapter XVIII excluding R95 Sudden infant death syndrome, ICD-11 chapter 21, excluding MH11 Sudden infant death syndrome), and vague or unspecified causes of death in other ICD-10 chapters.[2]

Redistribution: the process of reassigning ill-defined causes of death to plausible underlying causes of death. There are several methods used to determine a set of plausible underlying causes and proportions for a given group of ill-defined causes including multiple cause analysis, statistical (regression) models, impairment models (based on years of life lost), and proportional redistribution within age-sex groups.

Underlying cause of death: 'The disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury'.[3]

Unusable or insufficiently specified code: Codes that are of no utility for policy and should not be used as an underlying cause of death, such as septicaemia, heart failure, senility or headache. A cause that is insufficiently specified may still have some information content for policy, i.e., cancer, stroke etc., but could be improved (such as by stating the site of the cancer).

Usability index: Assesses the overall quality of cause of death data. The higher the usability index, the better quality the data and the more accurate the indicators. However, the usability index will realistically never be 100% because it is next to impossible to have perfect cause of death data.

$$Usability\ index\ (\%) = Proportion\ completeness * (1 - proportion\ ill\text{-}defined\ causes) * 100$$

1 *Twentieth World Health Assembly, 1967*
2 *ANACoD User Guidance, Version 3.0, October 2023*
3 *World Health Organization, 1994*

SUMMARY

The analyses presented in this report are based on a review conducted on the mortality and cause of death data collected by the civil registration and vital statistics system of Tonga between 2020–2024, representing 2,414 deaths with cause of death data. The review was performed using the ANACoD3 software tool, focussing on Steps 1–10.[4]

The report provides an overall assessment of the quality of the data for use in policy, as summarised in 22 data quality checks spread out over three main components (***Table 1***):

1. **Input data quality checks**. The number of death records with cause of death data is decreasing over time. Due to the high level of uncertainty being introduced in the resulting mortality indicators, this is limiting the usability of the information for policy and planning. Conversely, death distributions and age-specific mortality rates showed no major data quality concerns, and the proportion of invalid and inconsistent ICD codes was very low. The number of deaths assigned an unusable code also decreased between 2020 and 2024, indicating improvements in medical certification and coding practices.
2. **Mortality level analyses**. Reporting completeness was estimated at 65% during 2020–2024, and this is having a negative impact on key mortality indicators including the crude death rate and life expectancy at birth. Age- and sex-specific mortality rates indicate that deaths among male adolescents through to young adults are particularly under-reported. An examination of the under-five mortality rate highlights substantial under-reporting of these deaths.
3. **Cause of death analyses**. Age patterns of deaths by broad group differ to comparator data, however this may be due to poor comparator selection, with Tonga's mortality pattern closer reflecting that of low-income countries. There has been a decrease in the proportion of deaths assigned an ill-defined cause, from 57–45% between 2020 and 2024. Of these, unusable causes (the most severe type of ill-defined) also decreased from 18–12%. Despite these improvements, leading cause of death lists for males and females are heavily distorted due to ill-defined causes, and the proportion of unusable causes of death remains very high for older adults (aged 65+), at 21.1% for males and 24.6% for females. Overall, Tonga's cause of death data "usability index" shows little change over the past five years, primarily due to decreases in reporting completeness.

Table 1. *ANACoD3 data quality overview, Tonga, 2020–2024*

1 No action needed	2 Follow-up investigation recommended	3 Data improvement required

Review component		Comments	Status
Input data quality checks			
1.1	Population profile	Population data as provided by the UN Population Division are likely over-counts. Tonga is recommended to use their own population projections from the census when using ANACoD routinely.	**2**
1.2	Number of deaths	The number of death records with cause of death information appears to be decreasing over time, reducing the number of records available for cause-specific analyses.	**3**
1.3	Distribution of deaths by age and sex	Single-year data displayed high random variability, emphasizing the need to aggregate over time prior to conducing mortality analyses for planning or policy purposes.	**2**
1.4	Age-specific mortality rates	Single-year data displayed high random variability, emphasizing the need to aggregate over time prior to conducing mortality analyses for planning or policy purposes.	**2**

4 *Analysing Mortality and Cause of Death (Version 3.1). Geneva: World Health Organization; 2021 (https://www.who.int/standards/classifications/classification-of-diseases/services/analysing-mortality-levels-and-causes-of-death).*

Review component		Comments	Status
1.5	Invalid ICD codes	A small number of deaths were assigned an invalid code over the five-years of data analysed, with a clear improvement over the period.	1
1.6	Inconsistent ICD codes	Very few (0.4%) deaths were assigned an inconsistent cause of death over the five-years of data analysed.	1
1.7	ICD codes not to be used for underlying cause of death	The number of deaths assigned an unusable code decreased from 16–6% between 2020 and 2024. Almost half (42%) of such deaths were due to "unspecified cardiac arrest".	2
Mortality level analyses			
2.1	Age and sex distribution of deaths	The distribution of deaths by age and sex follows a mostly typical pattern based on global trends. Comparator data suggest that deaths among young children (0–4 years), deaths among male young adults (20–34), and deaths among older adults (65+ for males and 75+ for females) may be under-reported.	2
3.1	Completeness of death reporting	Reporting completeness was estimated at 65% during 2020–2024, with male deaths more likely to be reported than female deaths. Completeness appears to be decreasing over time.	3
3.2	Crude death rate	The CDR for males was 5.1 per 1000 population over the five-years of data analysed, while the CDR for females was lower, at 4 per 1000 population. This likely reflects under-reporting of deaths.	3
3.3	Life expectancy at birth	As with the CDR, the incomplete reporting of deaths is artificially increasing life expectancy at birth in Tonga, by as much as seven years for males, and six years for females.	3
4.1	Age- and sex-specific mortality rates	ASMRs highlight the under-reporting of deaths among young children and adolescents for males and females. For males, deaths among adolescents through to young adults appear particularly under-reported.	3
4.2	Ratio of male to female deaths	The ratio of male to female deaths shows noticeable peaks among ages 10–14 and 15–19, with over two males dying for every one female death. The "missing" deaths among males aged 20–24 years is apparent here, being the only age group for which the ratio is below one, which is unexpected.	3
5.1	Child mortality rates	The under-five mortality rate (U5MR) calculated from input data is lower than comparator data, and substantially lower than data as reported in the 2013–2018 Vital Statistics Report, indicating under-reporting.	3
Cause of death analyses			
6.1	Number of deaths by GBD list	The total number of ill-defined causes of death (all "R codes") is 403, or 16.9% of all deaths.	2
6.2	Number and proportion of deaths by GBD Group	There has been a noticeable decrease in the number of deaths assigned an ill-defined cause ("R codes"), from 18% in 2020 to 12% in 2024 – indicating improvements in medical certification and coding.	2
7	Distribution of major causes of death	Age patterns of deaths by GBD Group differ to comparator data. This may suggest under-reporting of specific causes of death, or poor comparator selection.	2
8	Leading causes of death	In looking at the leading causes of death, seven of the top 10 causes for males, and eight of the top 10 causes for females, were due to ill-defined causes, limiting the potential policy value of the data.	3
9	Ratio of non-communicable to communicable causes of death	The ratio has remained relatively stable over the five-years analysed, at around 3.1 deaths due to NCDs for every death due to communicable diseases. This is much lower than the comparator (10.3), and closer to low-income countries, than upper-middle income. This indicates Tonga's mortality pattern may still reflect that of low-income countries	2
10.1	Cause of death data usability index	Tonga's index shows little change over the past five years; however, this is primarily due to decreases in reporting completeness, as opposed to increases in ill-defined causes of death.	2

Review component		Comments	Status
10.2	Ill-defined causes of death	Along with a decrease in the number of deaths assigned an ill-defined cause (from 57–45%), there has been a noticeable decrease in the proportion of deaths assigned an unusable cause: from 17.8% in 2020 to 12.2% in 2024.	**1**
		Despite improvements, the proportion of unusable causes of death remains very high for older adults (aged 65+), at 21.1% for males and 24.6% for females.	**3**
10.3	Ill-defined codes	The most common ill-defined code is "septicaemia, unspecified", followed by "cardiac arrest, unspecified" and "senility".	**3**

BACKGROUND

Data quality assessment: ANACoD's main steps

All countries need accurate and up-to-date mortality statistics for a variety of purposes, including:

- informing health and social policy debates,
- monitoring progress relative to national and global development goals,
- monitoring trends in diseases and injuries, and
- evaluating policies designed to improve health outcomes.

The optimal source of cause of death data for a population is a functioning civil registration and vital statistics (CRVS) system, which registers all deaths and assigns a medically certified underlying cause of death. This individual cause of death data, once aggregated at the national level, forms the basis of mortality statistics for a country. It is therefore very important to ensure that the collected data are as accurate as possible and hence, assessing the quality of the data before dissemination is crucial to ensure users and policymakers can have confidence in the resulting statistics.

As a first step to any data improvement, gaining a detailed understanding of the types of problems with the data, particularly regarding completeness and diagnostic accuracy, is critical. A common concern with any mortality statistics produced from civil registration systems is how reliable they are in describing the actual mortality patterns in the population to which they refer.

The **Analysis of Mortality and Causes of Death, or ANACoD**, software tool used in this analysis was developed by the World Health Organization. The objective of the tool (in its third version) is to help users to perform a comprehensive and systematic analysis of mortality and cause of death data. The intended result is to enhance the value and usability of these data for informing health policies and programmes. ANACoD is specifically designed to be applied to large datasets such as those from civil registries, or any other sources that routinely collects and generates cause of death data coded to the International Standard Classification of Diseases and Related Health Problems (ICD), classified by age and sex.

The tool automatically tabulates data and presents basic mortality measures in the form of tables and figures. It highlights potential inconsistencies and errors in the data and estimates the completeness of reporting deaths with information on the cause of death. ANACoD3 generates indicators that reveal potential data quality issues, as well as an array of demographic and epidemiological indicators including sex- and age-specific mortality rates, crude death rates, life expectancy at birth, causes of death distributed by global burden of disease categories, leading causes of death, and the percentage of ill-defined causes of death. For ICD-10 data, the tool conducts an in-depth analysis of external causes of death.

ANACoD is designed to *identify problems* that need to be addressed to improve the value of mortality data for guiding health policies and practices. When applied on an annual basis it can be used to monitor the impact of improvement actions undertaken. Knowing the quality of the data from the ANACoD analysis will allow analysts and policy makers to make greater use of existing (potentially flawed) data by understanding the probable biases. It also provides critically important intelligence to guide strategies and interventions designed to improve the collection and analysis of cause of death data.

ANACoD3 is built around **12 core steps**, with two additional parts for special analyses:

1. data input checks (**step 1**),
2. mortality level analyses (**steps 2–5**),
3. cause of death data, basic (**steps 6–10**),
4. **special analyses:**
   a. cause of death data, special (external causes), and
   b. cause of death data, special (early childhood).

To diagnose possible problems in the mortality input dataset, ANACoD3 performs the following operations and analyses:

- Tabulates and/or graphs the input data in different ways to assess the plausibility of the data based on fundamental demographic and epidemiological relationships.
- Calculates the proportion and type of ill-defined codes that are of limited or no value for public health analysis.
- Compares the input data to a global source or estimate for the country or geographic region, to assess consistency.

To assess plausibility, national mortality data for a country are compared with the most recent estimates for that country or a neighbouring region. ANACoD also calculates standard mortality indicators and produces charts and tables that describe the mortality situation in the country or subregion and so can be used for reporting on leading causes of death by age and sex, along with an assessment of how reliable that information is.

Mortality and cause of death data that have been collected – often at great expense – should be of sufficient quality and representativeness to be used to their full potential. ANACoD was specifically designed to help producers and users of mortality datasets to assess whether they are 'fit for purpose', and to monitor the impact of activities implemented to improve data quality.

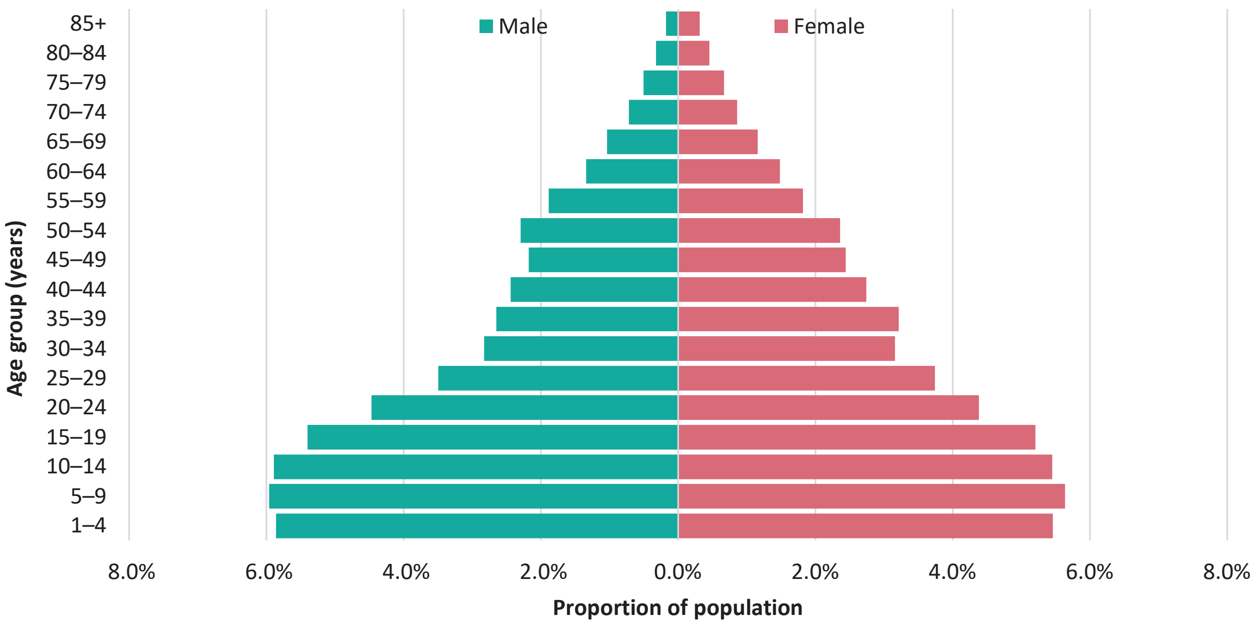
STEP 1: INPUT DATA

1.1 Population profile

Understanding the local demographic situation is important for assessing the plausibility of the tabulations, graphs, and mortality measures produced by ANACoD. Population data used in ANACoD analyses were sourced from the United Nations World Population Prospects, as provided in the tool. For multi-year analyses that combined all deaths over the five-year period (2020–2024), the 2022 mid-point population was used.

As shown in ***Figure 1***, Tonga has a relatively young population, characterised by high birth rates and few older people. From approximately age 35 years, there is a higher proportion of females than males alive in each age group, and this progressively increases with age.

Figure 1. *Population by age and sex, 2022 (mid-point population)*



1.2–1.4 Mortality profile

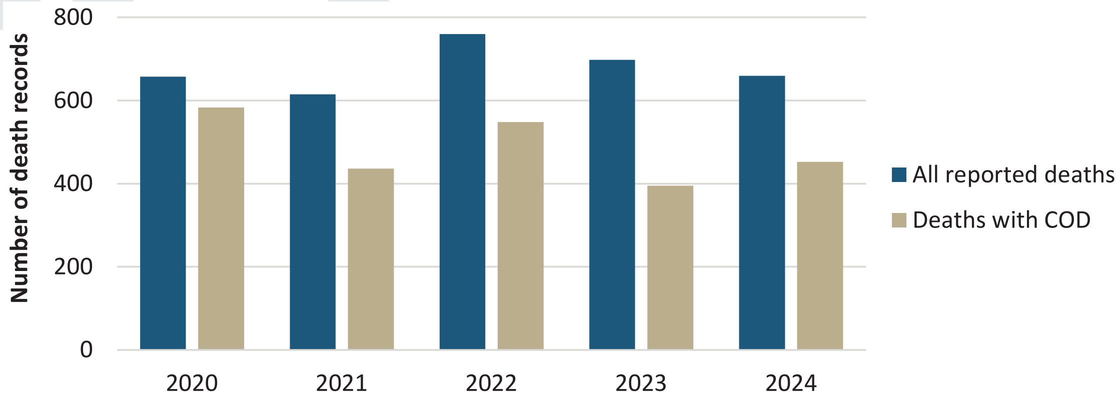
1.2 Number of deaths

A total of 3,389 deaths were reported during the five-year period analysed (***Table 2***). Of these, while 100% had usable age and sex information, the number of death records with cause of death data appears to be decreasing over time, from a high of 89% in 2020, to 67% in 2024 (***Figure 2***). In practice, this meant only 2,414 deaths were able to be included in cause-specific analyses.

Table 2. *Number of reported deaths by sex, with and without cause of death data, 2020–2024*

Year	All reported deaths			Death records with cause of death (COD)		
	Female	Male	Total	Female	Male	Total
2020	279	378	657	247	336	583
2021	270	345	615	202	234	436
2022	325	435	760	235	313	548
2023	320	378	698	194	201	395
2024	308	351	659	203	249	452
Total	**1,502**	**1,887**	**3,389**	**1,081**	**1,333**	**2,414**

Figure 2. *Number of death records with and without cause of death data, 2020–2024*



Notes: *COD = cause of death*

1.3 Distribution of deaths by sex and age group

This is discussed in more detail in Step 2.

The distribution of deaths by age and sex showed no significant data quality concerns, however single-year data highlighted considerable random variability in the data, emphasising the need to aggregate over time when conducting mortality analyses.

1.4 Age-specific mortality rates

This is discussed in more detail in Step 4.1.

As with the distribution of deaths by sex and age group, single-year data demonstrate high variability. Given the small number of deaths each year, aggregation over time is needed to strengthen the validity of any mortality measures (rates, ratios, etc.) calculated.

1.5–1.7 Cause of death profile

1.5 Invalid codes

In ANACoD3, an invalid code is one that does not exist in the ICD format selected, often due to data entry and transcription errors. Over the five-years of data analysed, 30 invalid codes were identified, with a noticable decrease over the period, indicating improvements in mortality coding and data entry (***Table 3***). Of these invalid codes, all were able to be corrected before moving on to ***Step 2*** of the tool.

Table 3. *Invalid ICD codes by sex, 2020–2024 (all deaths combined)*

Year	Female	Male	Total
2020	3	5	8
2021	5	5	10
2022	6	3	9
2023	0	1	1
2024	2	0	2
Total	**16**	**14**	**30**

1.6 ICD codes inconsistent with age and sex

The tool enumerates codes that are unlikely to cause death for a certain sex and/or in a certain age group, for example, male maternal deaths, prostate cancer deaths among females, or suicides among infants. Out of the five-years of data analysed, two deaths among males were assigned an unlikely cause of death (***Table 4***), while seven diseases unlikely to cause death at a certain age were recorded (***Table 5***).

The overall proportion of death records with an inconsistent cause of death – at 0.4% – is very low, indicating good data quality. To further reduce the occurrence of these types of inconsistencies, data quality checks using ANACoD should be performed as close to the time of reporting as possible, for instance, every week, fortnight, or month. Once these inconsistent codes have been identified, verification with original medical records will need to be conducted to assess if more information is available to "correct" the death record (either the sex, age, and/or cause of death) – if needed.

***Table 4.** Diseases unlikely to cause death in a certain sex, 2020–2024 (all deaths combined)*

ICD code	Disease	Sex recorded	Number of deaths
C571	Broad ligament	M	1
N80	Endometriosis	M	1
Total			**2**

***Table 5.** Diseases unlikely to cause death at a certain age, 2020–2024 (all deaths combined)*

ICD code	Disease	Age recorded	Number of deaths
N926	Irregular menstruation, unspecified	<12 or >49 years	1
O348	Maternal care for other abnormalities of pelvic organs	<12 or >49 years	1
P220	Respiratory distress syndrome of newborn	>1 years	2
P240	Neonatal aspiration of meconium	>1 years	1
P369	Bacterial sepsis of newborn, unspecified	>1 years	1
R54	Senility	<65 years	1
Total			**7**

1.7 ICD codes not to be used for underlying cause of death

This is discussed in more detail in Steps 6 and 10.

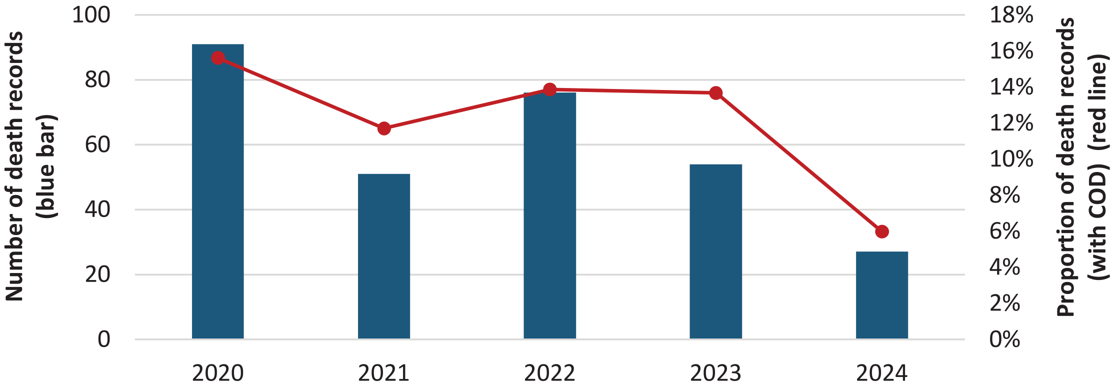
Understanding the number of deaths that have been coded to an **ill-defined code i**s of critical importance when assessing data quality. Also referred to as "unusable", "insufficiently specified", or "garbage codes", these codes contain no or very little useful information about the probable cause of death for that individual, and hence are of dubious value for guiding public health policy. When the proportion of ill-defined codes is large (above 10–15%)[5] it will bias the cause-specific distribution of deaths, as the data will not represent the true health status of the population.

5 *Mikkelsen L, Lopez AD. Guidance for assessing and interpreting the quality of mortality data using ANACONDA. CRVS Resources and tools. Melbourne, Australia; Bloomberg Philanthropies Data for Health Initiative, Civil Registration and Vital Statistics Improvement, University of Melbourne; 2017.*

© Flickr, Soniahhhh

As shown in ***Figure 3***, the number of death records assigned an ill-defined cause of death decreased noticably during the five-years analysed, from a high of 16% in 2020, to the current low of 6% in 2024. This indicates improvements in the quality of medical certification and mortality coding practices.

Figure 3. *Number and proportion of death records with an ill-defined cause of death, 2020–2024*



Despite these improvements, 299 deaths were assigned an unusable code between 2020–2024, with almost half of these (42%) due to “unspecified cardiac arrest”, followed by “palliative care” (***Table 6***).

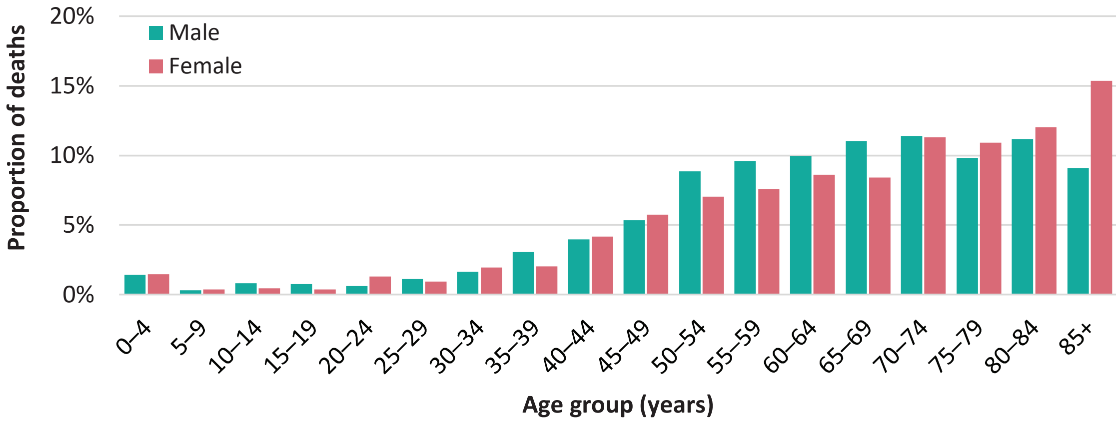
Table 6. *Top five ICD codes not to be used for underlying cause of death, 2020–2024 (all deaths combined)*

Rank	ICD code	Disease	Number of deaths
1	I469	Cardiac arrest, unspecified	124
2	Z515	Palliative care	69
3	T751	Drowning and nonfatal submersion	13
4	R69	Unknown and unspecified causes of morbidity	11
5	S099	Unspecified injury of head	9
	T71	Asphyxiation	9
Total (top five)			**235**
—	—	All other ill-defined codes	64
Total			**299**

2. Age and sex distribution of deaths

The age and sex distribution of deaths varies considerably depending on the overall level of mortality in a country, which determines the risk of dying at each age, and the size of the population currently alive at each age. Irrespective the level of mortality, the number of deaths should gradually increase from the age of five years onwards. As shown in ***Figure 4***, the distribution of deaths by age and sex in Tonga follows a mostly typical pattern based on global epidemiological and demographic trends, with relatively higher mortality among the youngest age group (0–4 years), which then decreases, only to progressively increase with age.

Figure 4. *Age distribution of deaths by sex, 2020–2024 (all deaths combined)*



Adolescent males account for a slightly higher proportion of deaths than females of the same age, and this is generally linked with increased risk-taking among males during these ages. The higher proportion of deaths among females aged 20–24 years (and minimally among those aged 25–29 and 30–34 years) is not expected and may be due to several factors, including under-reporting of male deaths in these age groups, and random variation. It could also reflect higher pregnancy- and maternal-related health risks among this age group, with any resulting deaths more likely to be recorded by the Ministry of Health, given the high proportion of women giving birth in health facilities.

The higher proportion of deaths among males from age 50–69 is to be expected, given their higher age-specific mortality rates. From age 70, there is a higher proportion of deaths among females than males, which also is to be expected given their comparatively lower age-specific mortality rates and higher life expectancy.

Comparator data for Tonga are shown for males and females separately in ***Figures 5 & 6*** respectively. For males, WHO estimates have more deaths among young children (0–4 years), young adults (20–34) and among older adults (65+) than what is currently being recorded. For females, WHO estimates have more deaths among young children (0–4 years), and among older adults (age 75+) than recorded.

Figure 5. *Age distribution of male deaths [with comparator], 2020–2024 (all deaths combined)*

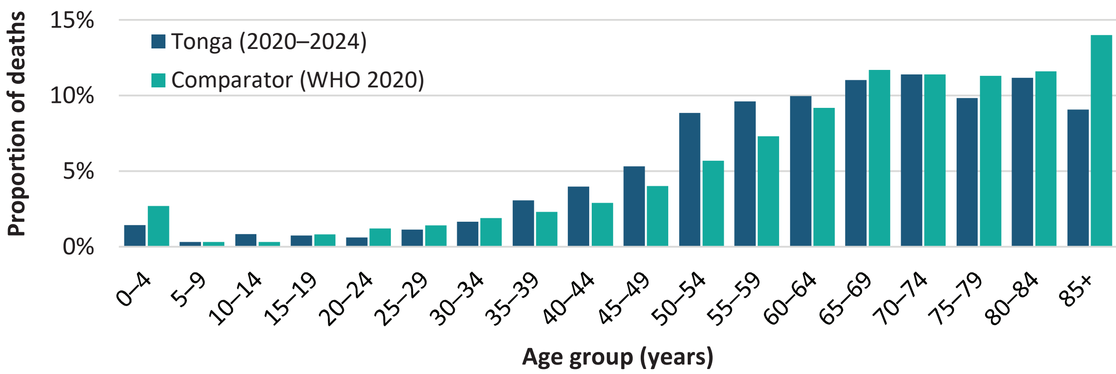
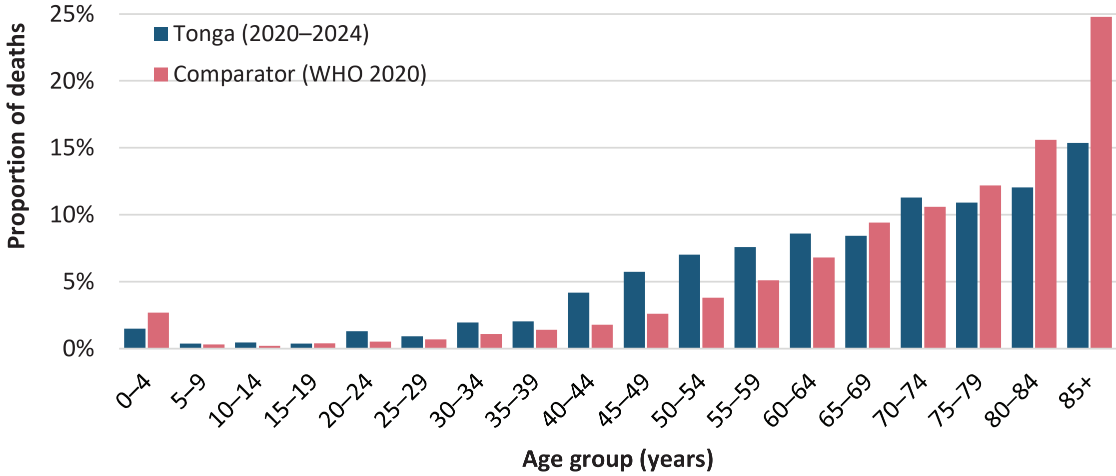


Figure 6. *Age distribution of female deaths [with comparator], 2020–2024 (all deaths combined)*



These differences between what is estimated and observed may simply reflect the limitations of using income level as a predictor of mortality, especially for countries like Tonga, who is still in the early stages of the demographic transition.

Differences may also be due to low reporting/registration completeness, particularly for child deaths – a known data quality issue for many countries. Further, as many deaths reported by the Ministry of Health are those that have occurred in health facilities, they generally do not include accidents (a major cause of death for males) and deaths among the elderly (particularly for females). As such, these differences (to the comparator) may indicate that child deaths are being under reported/registered for males and females, along with deaths among young adult males, and deaths among older adults (particularly females).

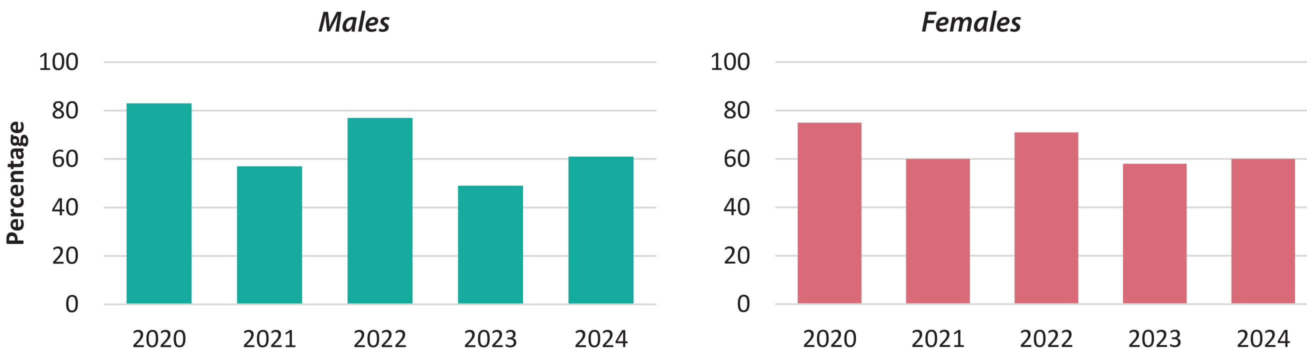
3. Summary indicators of mortality

3.1 Completeness of death reporting

Understanding the completeness of death reporting can be used to monitor the performance of the CRVS system in capturing deaths, and to allow for adjustment of incomplete data. Completeness is defined as the proportion of actual deaths in a population that are reported, divided by the estimated ("true" or "actual") number of deaths that occurred in the same year. The estimated number of deaths in ANACoD3 is sourced from the United Nations.

When applied to the number of deaths reported each year in Tonga, the completeness of death reporting during 2020–2024 was 65% for males and females, and this appears to be decreasing over time (***Figure 7***). Completeness of reporting for female deaths is lower than for male deaths, and while the gap is decreasing, this appears to be due to the decrease in reporting completeness for males, rather than an improvement in reporting completeness for females.

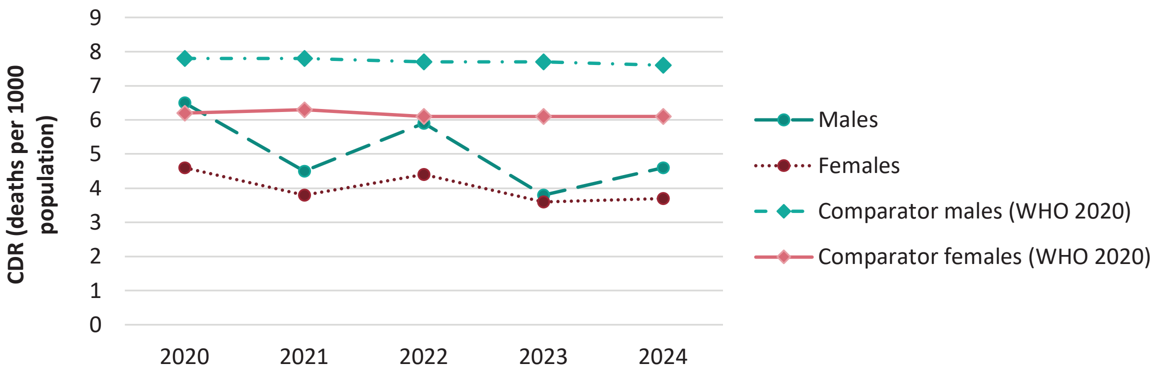
Figure 7. *Completeness of death reporting by sex, 2020–2024*



3.2 Crude death rate

As a mortality indicator, the crude death rate (CDR) is the simplest measure of population health status. It is a measure of the number of deaths relative to the size of that population during a given period, usually one year. It is expressed in units of deaths per 1000 population per year. The CDR is also an important measure of data quality, as in a system where not all deaths are registered, the CDR will underestimate the true level of mortality (i.e., the calculated CDR will be lower than the "true" CDR). Based on many decades of observing the CDR, demographers have concluded there is generally a lower limit of around 5 deaths per 1000 population.[6] Thus, any CDR below 5 per 1000 should be viewed with caution, since it is likely to reflect an under-reporting of deaths rather than a true reflection of low overall mortality. As shown in ***Figure 8***, based on input data, the CDR for males was 5.1 per 1000 population over the five-years of data analysed, while the CDR for females was lower, at 4.0. This is lower than comparator data from WHO (at 7.7 and 6.2 respectively), indicating low reporting completeness.

Figure 8. *Crude death rate by sex [with comparators], 2020–2024*



3.3 Life expectancy at birth

As with the CDR, the incomplete reporting of deaths is artificially increasing life expectancy at birth in Tonga, by as much as seven years for males, and six years for females (***Table 7***).

Table 7. *Life expectancy at birth by sex [with comparators], 2020–2024*

Year	Reported deaths		WHO estimates (2020)	
	Males	Females	Males	Females
2020	71	78	68	74
2021	76	81	68	74
2022	72	79	69	74
2023	79	83	69	74
2024	76	81	69	74

4. Age- and sex-specific mortality patterns

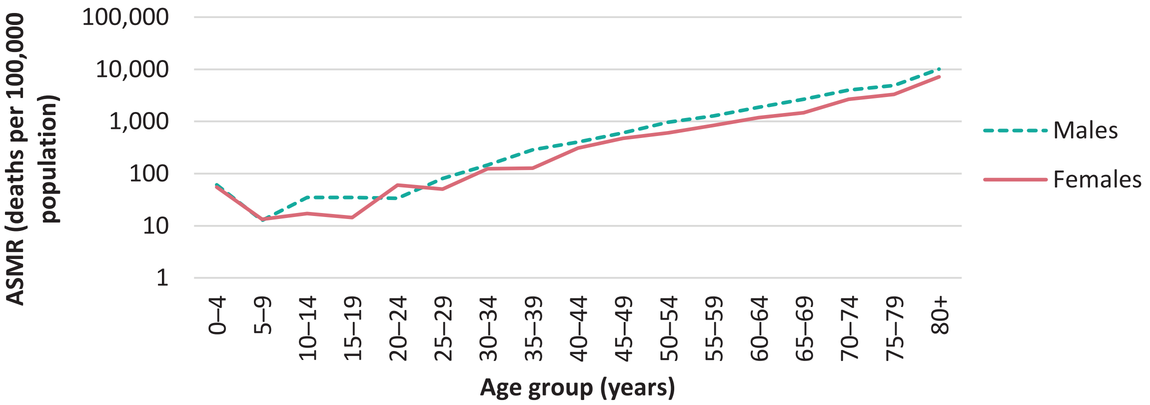
4.1 Age- and sex-specific mortality rates

Age-specific mortality rates (ASMRs) can be used to assess the quality of mortality data by comparing the rates calculated from the input data with expected age patterns of mortality risk. Generally, mortality rates are high during infancy and early childhood and fall to their lowest levels between 5–14 years. After this, mortality rates rise with increasing age. Male mortality is usually higher than female at all ages, with a peak of excess mortality between ages 15–24 years. From around age 35 years, mortality rates for both males and females generally increase on an almost straight line.

6 *Mikkelsen L, Lopez AD. Guidance for assessing and interpreting the quality of mortality data using ANACONDA. CRVS Resources and tools. Melbourne, Australia; Bloomberg Philanthropies Data for Health Initiative, Civil Registration and Vital Statistics Improvement, University of Melbourne; 2017.*

As shown in ***Figure 9***, Tonga's ASMRs demonstrate the comparatively higher levels of mortality among children aged 0–4 years. From about age 20 years, mortality rates begin to increase exponentially, and while expected, this does indicate higher levels of mortality at younger ages. There is a noticeable "bump" in excess male mortality between the ages of 10–19 years, and while expected, this is also occurring at younger ages than what is seen globally. As mentioned previously, the higher number of female deaths in the 20–24 age group is not expected and may reflect random variation, or the noted under-reporting of male deaths at this age.

Figure 9. *Age-specific mortality rates by sex, 2020–2024 (all deaths combined)*



Comparator data from WHO are shown for males and females separately below. While the general trends are similar, they highlight the under-reporting of deaths among young children and adolescents for males and females. For males, deaths among adolescents through to young adults (approximately aged 10–24 years) appear to be particularly under-reported.

Figure 10. *Age-specific mortality rates for males [with comparator], 2020–2024 (all deaths combined)*

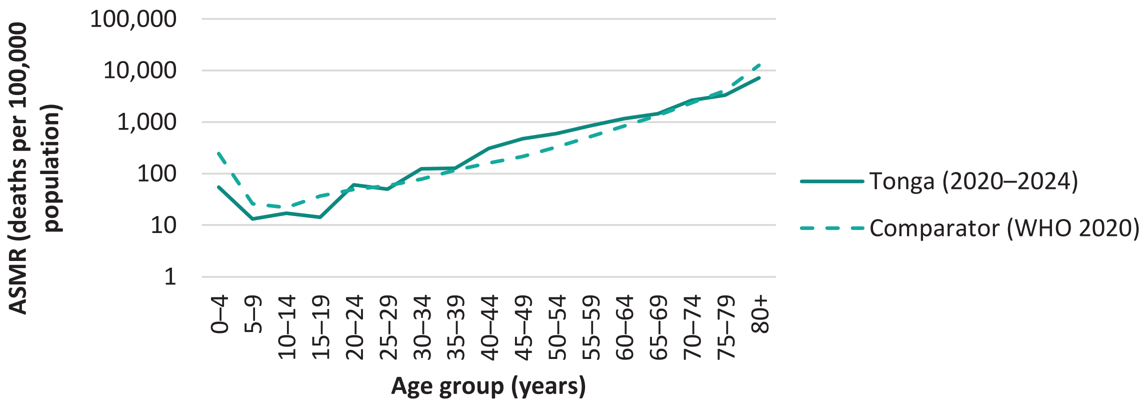
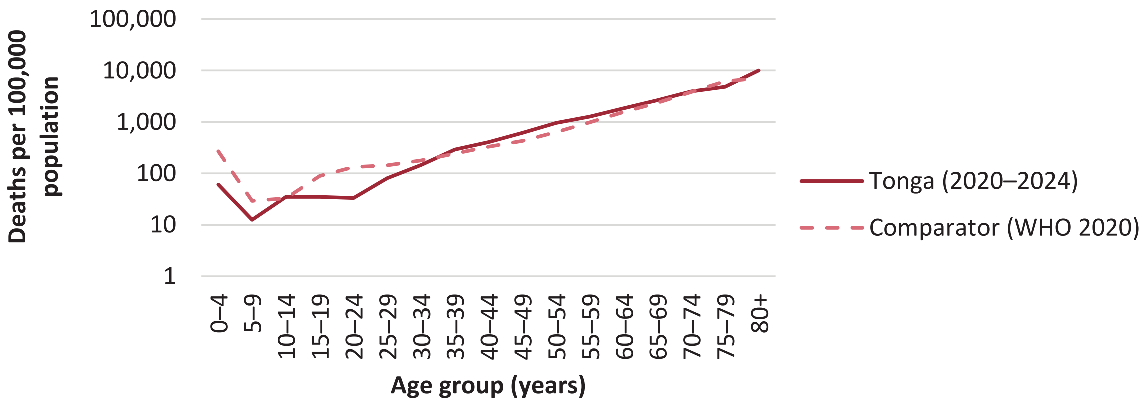


Figure 11. *Age-specific mortality rates for females [with comparator], 2020–2024*

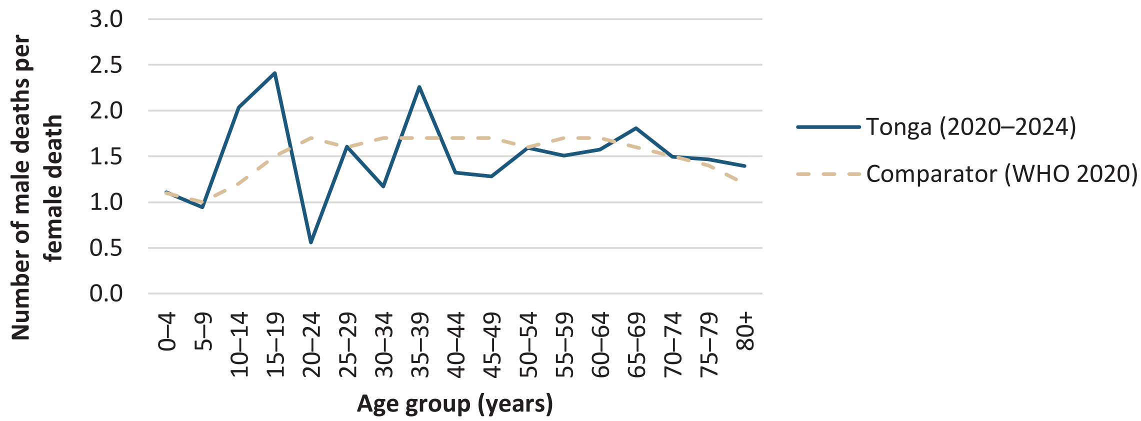


4.2 Sex ratio

If males and females died at the same rate at each age group, the ratio between male and female mortality rates would equal one. As discussed previously, male death rates are generally higher than female at all ages, and so the ratio should be above one. The ratio is typically 2-3:1 between ages 15–34 due to much higher male mortality associated with accidents, suicides, and violence (e.g. for every one female death at these ages, there are two to three male deaths). Sometimes there is also a secondary, lower peak in the ratio between ages 50–64 years due to the higher number of men dying prematurely from non-communicable diseases (NCDs). At the very highest age groups, the female rates might be identical or slightly higher than the male rates so that the ratio becomes negative and dips below one.

As shown in ***Figure 12***, the ratio of male to female deaths (using the ASMR) shows noticeable peaks among ages 10–14 and 15–19, with over two males dying for every one female death. There is a secondary peak in the 34–39 age group, which, if due to premature mortality from NCDs, is occurring at much younger ages than what is expected based on global trends. As previously highlighted, the "missing" deaths among males aged 20–24 years is apparent here, being the only age group for which the ratio is below one, which is unexpected.

Figure 12. *Ratio of male to female deaths by age group [with comparator], 2020–2024*



5 Child mortality rates

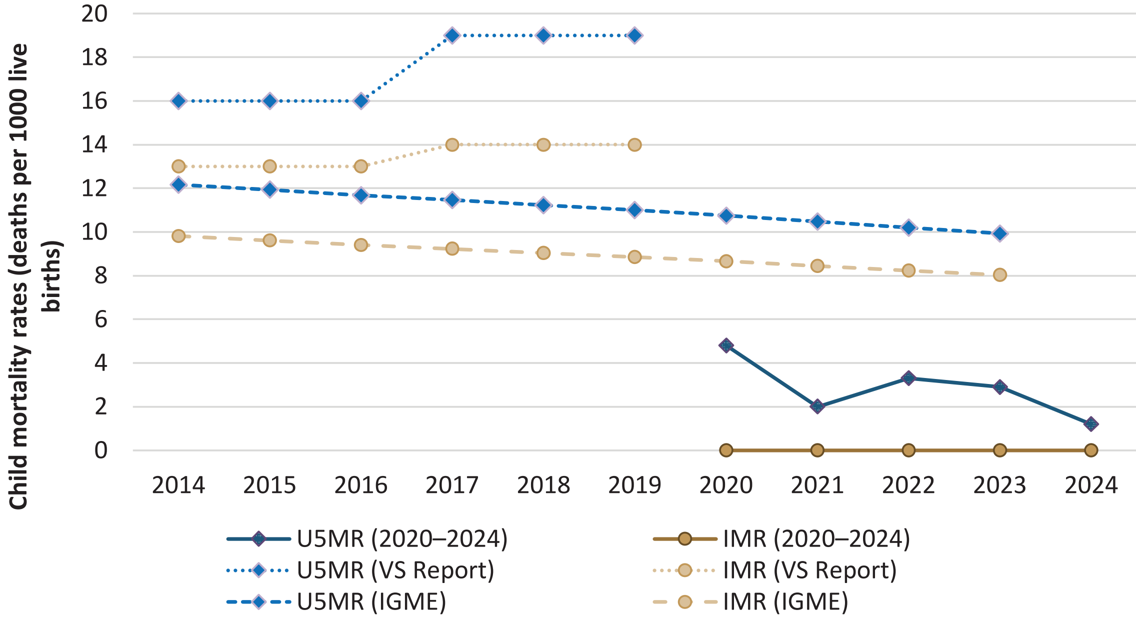
More than any other age group, the level of mortality among children under five years old reflects a country's economic, social, and health conditions. However, research also shows that this is the age group for which deaths are most likely to go unreported to health authorities and the civil registration system. To assess the potential level of under-registration of child mortality in civil registration data, the Interagency Group for Child Mortality Estimation (IGME) compares reported child deaths with survey- and census-derived estimated of child mortality, which are more likely to be accurate.[7] Data from the IGME are used as comparators in the ANACoD3 tool.

As ***Figure 13*** highlights, the under-five mortality rate (U5MR) calculated from input data ranged from 4.8–1.2 deaths per 1000 live births between 2020–2024, lower than comparator data provided by the IGME (range 12.2–9.9 between 2014–2023), and substantially lower than data as reported in the 2013–2018 Vital Statistics Report (16–19 deaths per 1000 population).

Due to the format of input data provided for this analysis, which did not disaggregate age below one-year, the infant mortality rate (IMR) was unable to be calculated.

7 *For more information on the United Nations Inter-agency Group for Child Mortality Estimation, see* https://resourcecentre.savethechildren.net/publishers/un-igme-united-nations-inter-agency-group-child-mortality-estimation

Figure 13. *Child mortality rates [with comparators], 2014–2024 (all deaths combined)*



6. Distribution of deaths by Global Burden of Disease

6.1 Number of deaths by Global Burden of Disease list

This step provides a summary tabulation of deaths according to the Global Burden of Disease (GBD) list, by sex and age group, which can be used to create customised leading cause of death lists.

As highlighted in ***Table 8***, the number of "ill-defined diseases" in this tabulation has increased from 299 deaths (as shown in ***Step 1.7***) to 403 deaths. This increase in the proportion of ill-defined causes (from 12.4–16.9%) is due to ANACoD3's classification rules, which have included additional poorly specified deaths from Chapter XVIII, Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (also referred to as "R codes"), which were not included with the initial tabulation during ***Step 1.7***.

Table 8. *Number of deaths by GBD cause, by sex, 2020–2024 (all deaths combined)*

GBD cause	Reported deaths		
	Males	Females	Total
Communicable, maternal, perinatal, and nutritional conditions	239	240	479
Non-communicable diseases	837	625	1462
Injuries	49	21	70
Ill-defined diseases	**208**	**195**	**403**
All causes	**1,333**	**1,081**	**2,414**

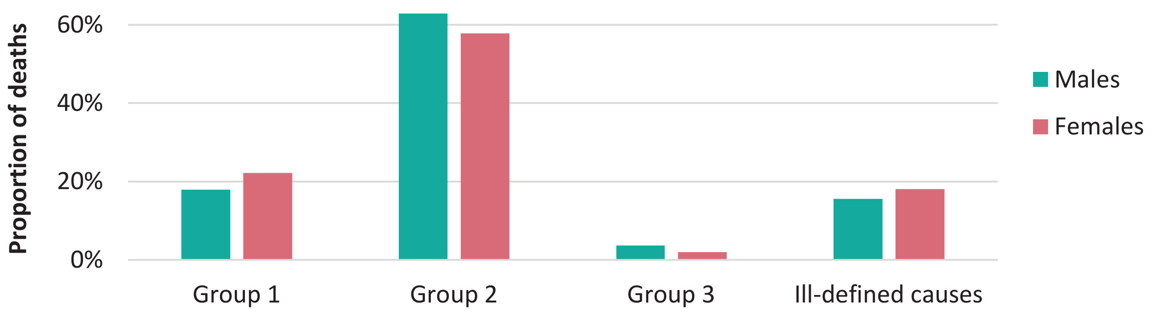
6.2 Number and proportion of deaths by major cause groups

An important step in checking the quality of cause of death data is to look at the distribution of deaths by three broad groups to assess if the observed pattern is consistent with what is known about the extent of the epidemiological transition in the country. The three groups are:

- Group 1: Communicable, maternal, neonatal, and nutritional diseases.
- Group 2: Non-communicable diseases, including mental health conditions.
- Group 3: External causes and injuries (e.g., accidents, homicide, suicide, war deaths and natural disasters.[8]

Based on the input data, approximately 20% of recorded deaths during 2020–2024 were due to infectious diseases, 60% due to NCDs, and 3% external causes, with the remaining 17% assigned an ill-defined cause (***Figure 14***).

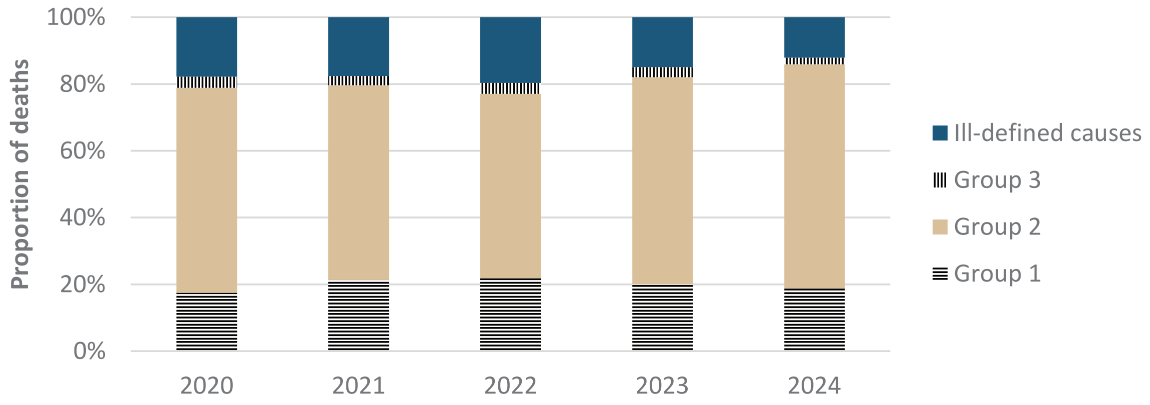
Figure 14. *Proportion of deaths by GBD Group, including ill-defined causes, 2020–2024 (all deaths combined)*



8 *Murray CJL, Lopez AD (eds.). The Global Burden of Disease. A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Boston, USA: Harvard School of Public Health on behalf of the World Health Organization and The World Bank; 1996.*

As shown in ***Figure 15***, there has been a noticeable decrease in the number of deaths assigned an ill-defined cause ("R codes"), from 18% in 2020 to 12% in 2024 – indicating improvements in medical certification and coding.

Figure 15. *Proportion of deaths by GBD Group, including ill-defined causes, 2020–2024*

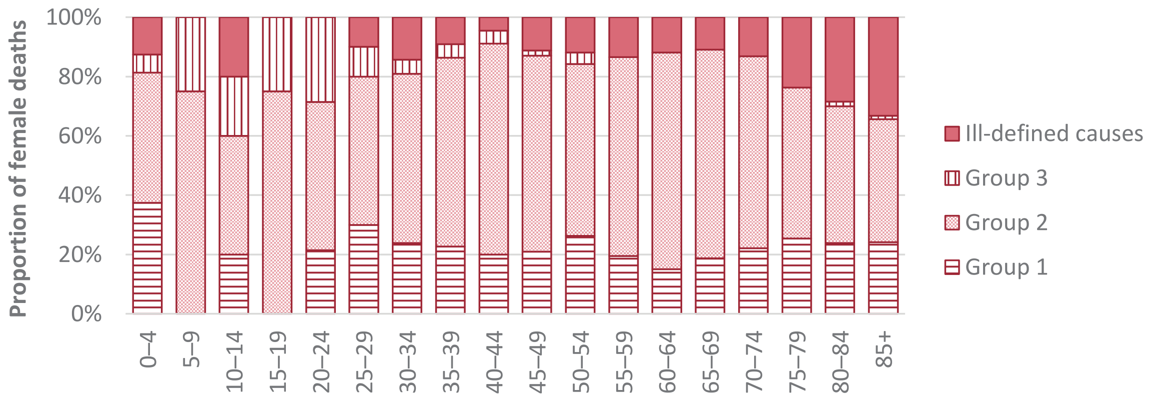


Data from this step can also be used to create a detailed visualisation of the distribution of deaths by sex and age group, highlighting any potential biases in the cause of death distribution through analysing the proportion of deaths with an ill-defined cause in each age group.

As shown in ***Figures 16 & 17***, mortality in Tonga generally follows the expected pattern, with a higher proportion of deaths due to communicable diseases among the youngest age groups and an increasing proportion of deaths due to NCDs among older adults. For both females and males, accident- and injury-related deaths are more common among younger adults; however, the relative share of these deaths is larger for males than for females. The proportion of deaths with an ill-defined cause, while an issue across all age groups, is particularly problematic among the oldest age groups, from approximately age 75 for females, and 80 for males.

Figure 16. *Proportion of male deaths by GBD Group, including ill-defined causes, by age group, 2020–2024 (all deaths combined)*

Figure 17. *Proportion of female deaths by GBD Group, including ill-defined causes, by age group, 2020–2024 (all deaths combined)*



7. Distribution of major causes of death

The causes of death in a population generally follow a predictable age pattern that has been identified from decades of global epidemiological observations. This is because the risk of dying from the different diseases and injuries covered in each group varies with age. The diseases in Group 1 (communicable diseases) are known to cause significant mortality in younger children (particularly those aged less than one year). After this, the proportion of deaths declines to a low level, and only gradually begins to increase again towards the older ages where people become more susceptible to infectious diseases such as pneumonia. Group 2 (NCDs) has very few deaths in younger age groups, with most deaths found in the adult and older age groups for both men and women. A clear gender difference is usually observed in Group 3, with more males dying from external causes during young adulthood, while the proportion of women dying remains relatively even at each age group.

Figures 18 & 19 show the age pattern of deaths found for each of the three groups together with comparator data for Tonga, as provided from the ANACoD tool. For males, comparator data (from WHO) highlight the potential under-reporting of deaths due to communicable diseases among children (0–14 years). From age 15 onwards, input data are comparatively higher than WHO estimates, particularly among older age groups. For Group 2 deaths (NCDs), input data are generally higher than WHO estimates until around age 55. For injuries and accidents (Group 3), comparator data from WHO indicate potential under-reporting from age 10 years and above.

The trend is similar for females, with potential under-reporting among children (0–9 years), and again among adolescents and young adults (15–24). From age 25–44, input data are similar to WHO comparators, while from age 45 onwards, input data are higher, indicting more deaths due to communicable diseases than would be expected for a country at this income level. Deaths due to NCDs are higher than comparator data from birth to age 44, with noticeable differences among the youngest age groups (0–19 years), indicating a substantial NCD-related burden. As with males, injury and accident deaths are lower than comparator data, suggesting under-reporting.

As previously discussed, these differences (between input data and the comparator) may be due to poor comparator selection, especially as Tonga's mortality profile does not align perfectly with mortality profiles for other upper-middle income countries.

Figure 18. *Proportion of deaths by GBD Group, males [with comparator], 2020–2024 (all deaths combined)*

Figure 19. *Proportion of deaths by GBD Group, females [with comparator], 2020–2024 (all deaths combined)*

8. Leading causes of death

All vital statistics systems should be able to produce a table showing the leading causes of death for the population. Such rankings are an important source of policy-relevant information. Overall, of the 2,414 death records analysed during 2020–2024, 15.4% of deaths were due to ischaemic heart disease, followed by lower respiratory infections (6.2%), and cerebrovascular diseases (5.4%). For children aged 0–4 years, the leading cause of death was prematurity and low birth weight (at 6.5%).

An analysis of the leading causes of death can also indicate the reliability and plausibility of the cause of death data. While this step in ANACoD provides a list of leading causes of death *after* redistribution, from a quality perspective, understanding the impact of ill-defined causes on leading causes of death is also important. As such, an additional software tool, ANACONDA,[9] was used to analyse causes of death *before* and after redistribution.

ANACONDA groups ill-defined causes according to their potential impact for guiding or misguiding public policy to prevent premature deaths. In this classification, four levels are defined, depending on how serious their impact is for misinforming public policy. Specific examples of these causes include "sepsis", "cardiac arrest", and "senility" – all of which provide no actionable information on the underlying cause of death.

In looking at the leading causes of death for females before redistribution, the impact of ill-defined causes is noticeable, with eight of the 10 leading causes ill-defined, including unspecified septicaemia, senility, and unspecified cardiac arrest. Following redistribution, only one ill-defined cause remains in the top 10, with the most probable leading causes of death for females being ischemic heart disease, stroke, and breast cancer (***Table 9***).

For males, seven of the top 10 leading causes of death were due to ill-defined causes, including unspecified septicaemia, unspecified cardiac arrest, and cardiogenic shock. Following redistribution, the most likely leading cause of death in the country is also ischemic heart disease – however, compared with deaths among females, this category takes up a much higher proportion of deaths – as much at 25% of all male deaths (***Table 10***).

Table 9. *Top 10 causes of female deaths, before and after redistribution of ill-defined causes, 2020–2024 (all deaths combined)*

Before redistribution (input data)				After redistribution	
Rank	**ICD code**	**Cause of death**	**Proportion of deaths (%)**	**Cause of death**	**Proportion of deaths (%)**
1	A41.9	Septicaemia, unspecified	15.1	Ischemic heart disease	15.8
2	I21.9	Acute myocardial infarction, unspecified	7.3	Stroke	8.9
3	R54.0	Senility	5.4	Breast cancer	4.8
4	I46.9	Cardiac arrest, unspecified	5.3	Sepsis, unspecified organism	4.6
5	R57.0	Cardiogenic shock	3.2	Chronic kidney disease	2.4
6	C50.9	Breast, unspecified	3.1	Diabetes mellitus	2.4
7	Z51.5	Palliative care	3.1	Cirrhosis and other chronic liver diseases	2.2
8	J18.9	Pneumonia, unspecified	3.0	Cervical cancer	2.0
9	I64.0	Stroke, not specified as haemorrhage or infarction	2.9	Tracheal, bronchus and lung cancer	1.8
10	N19.0	Unspecified renal failure	2.3	Uterine cancer	1.6
		Total (top 10 causes)	**50.7**	**Total (top 10 causes)**	**46.5**
		All other causes	**49.3**	**All other causes**	**53.5**

9 *Area Analysis of Causes of National Deaths for Action (ANACONDA). The University of Melbourne.*

Table 10. *Top 10 causes of male death, before and after redistribution of ill-defined causes, 2020–2024 (all deaths combined)*

Before redistribution (input data)				After redistribution	
Rank	ICD code	Cause of death	Proportion of deaths (%)	Cause of death	Proportion of deaths (%)
1	I21.9	Acute myocardial infarction, unspecified	14.3	Ischemic heart disease	25.2
2	A41.9	Septicaemia, unspecified	9.5	Stroke	9.4
3	I46.9	Cardiac arrest, unspecified	5.0	Tracheal, bronchus and lung cancer	3.5
4	J18.9	Pneumonia, unspecified	4.7	Chronic obstructive pulmonary disease	3.5
5	R57.0	Cardiogenic shock	3.7	Cirrhosis and other chronic liver diseases	2.3
6	R54.0	Senility	3.6	Chronic kidney disease	2.2
7	I50.9	Heart failure, unspecified	3.5	Diabetes mellitus	2.1
8	I64.0	Stroke, not specified as haemorrhage or infarction	2.9	Prostate cancer	2.0
9	Z51.5	Palliative care	2.7	Stomach cancer	1.3
10	N19.0	Unspecified renal failure	2.0	Colon and rectum cancer	1.1
		Total (top 10 causes)	51.9	**Total (top 10 causes)**	**52.6**
		All other causes	48.1	**All other causes**	**47.4**

Analytical software: ANACONDA

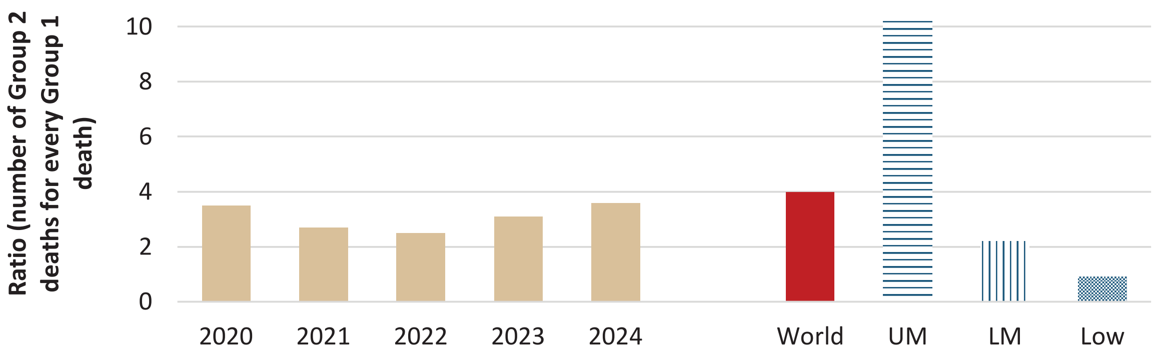
Key: [shaded] *Ill-defined causes of death (Severity level 1–3)*

9. Ratio of non-communicable to communicable causes of death

Step 9 produces the ratio of deaths from non-communicable diseases (Group 2) to communicable diseases (Group 1) and compares them side-by-side with the average ratios according to income grouping (both sexes combined).Globally, there are about four times as many deaths due to Group 2 causes as to Group 1. In high-income countries, NCDs account for 13 times as many deaths as communicable diseases. By contrast, in low-income countries, more deaths are due to communicable diseases (the ratio is less than 1). This reflects the fact that in high- 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 early childhood, due to infectious diseases conditions – pneumonia, diarrhoea, and vaccine preventable conditions – and perinatal causes.

As shown in ***Figure 20***, the ratio has remained relatively stable over the five-years analysed, at around 3.1 deaths due to NCDs for every death due to communicable diseases. This is much lower than the comparator (10.3), and closer to low-income countries, than upper-middle income. This indicates Tonga's mortality pattern may still reflect that of low-income countries.

Figure 20. *Ratio of deaths due to Group 2 to deaths due to Group 1 [with comparator], 2020–2024 (all deaths combined)*



Notes: *UM = upper-middle income; LM = low-income; Low = Low income*

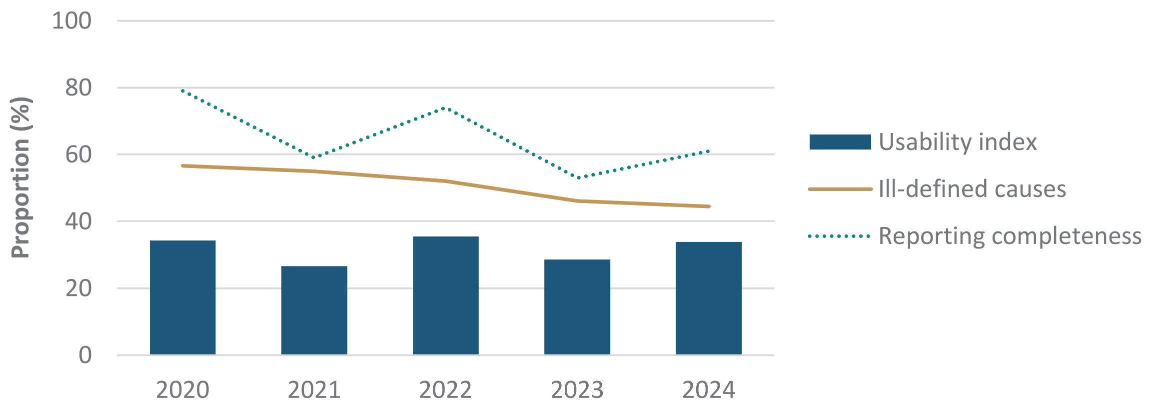
10. Quality of cause of death data

10.1 Cause of death data usability index

To assess the overall usability of cause of death data, WHO has developed a “usability index”, which is a function of reporting completeness (from ***Step 3***) and the proportion ill-defined causes of death (***Step 10.2***). The higher the usability index, the better quality the data and the more accurate the indicators.

While Tonga’s index shows little change over the past five years, this is primarily due to decreases in reporting completeness (from 79–61%), with the proportion of ill-defined causes showing a decrease during the same period (57–45%) (***Figure 21***). In essence, while the quality of cause of death data has improved (as reflected in a decrease in ill-defined causes of death), the overall usability index has remained unchanged due to decreases in reporting completeness over the period.

Figure 21. *Usability index, proportion of ill-defined causes of death, and reporting completeness, 2020–2024 (all deaths combined)*



10.2 Ill-defined causes of death

The ICD-10 classifies mortality codes into 22 broad chapters. In this step, ANACoD tabulates the number of ill-defined causes of death in each chapter and presents them by sex and age group. It should be noted that in this step, the definition of “ill-defined causes” has been expanded (again) to include unusable codes (as discussed in ***Step 1.7*** and ***Step 6.1***), along with all other insufficiently specified causes. As shown in ***Table 11***, this has substantially increased the number of ill-defined causes in the dataset from 403 (17%) to 1,238 (51.3%).

Table 11. *Number of ill-defined causes of death in each chapter by sex, 2020–2024 (all deaths combined)*

ICD-10 Chapter	Reported deaths		
	Males	Females	Total (%)
I Infectious and parasitic diseases	139	172	311
II Neoplasms	10	12	22
III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	0	1	1
IV Endocrine, nutritional metabolic diseases	1	1	2
IX Diseases of the circulatory system	166	120	286
X Diseases of the respiratory system	15	15	30
XI Diseases of the digestive system	9	8	17
XIV Diseases of the genitourinary system	50	49	99
XVI Certain conditions originating in the perinatal period	0	0	0
XVIII Sympoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	208	195	403
XX External causes of morbidity and mortality	46	21	67
Total ill-defined causes (% of ill-defined)	**644 (48.3)**	**594 (54.9)**	**1,238 (51.3)**

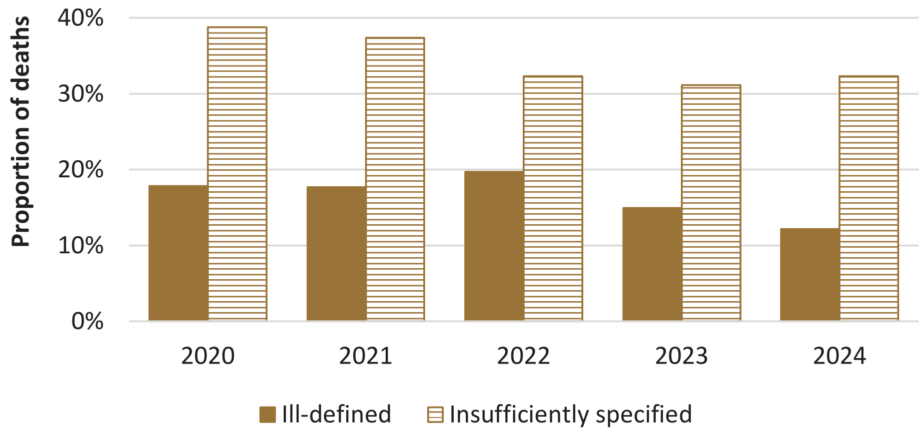
To better differentiate between these two sub-groups of ill-defined causes, the following categories have been applied (based on WHO guidance):

1. **Unusable causes**. These are causes that should not be used as the underlying cause of death, as found in Chapter XVIII, Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (also referred to as "R codes").
2. **Insufficiently specified causes**. These are vague or unspecified causes of death in other ICD-10 Chapters, and include ill-defined cancer sites, unspecified heart disease, and external causes of death where the intent cannot be determined.

While deaths classified to either of these two categories are of limited value for public health purposes, *insufficiently specified* causes can generally be used to identify the broad cause of death group. For example, a death with an ill-defined cancer site would still be classified as a cancer death, and able to be analysed as part of Group 2 (NCDs). While it would be more beneficial to know if the death was due to breast, cervical, prostrate, or lung cancer, knowing the death was due to cancer still has some policy value. Of more concern are the *unusable causes*, as they provide no information on the probable cause of death – thus having a major, negative impact on policy.

In looking at ***Figure 22*** below, along with a slight decrease in the proportion of deaths assigned an insufficiently specified cause (from 39-32%), there was also a decrease in the proportion of deaths assigned an unusable cause: from 18% in 2020, to 12% in 2024.

Figure 22. *Distribution of ill-defined causes of death by category, 2020–2024*



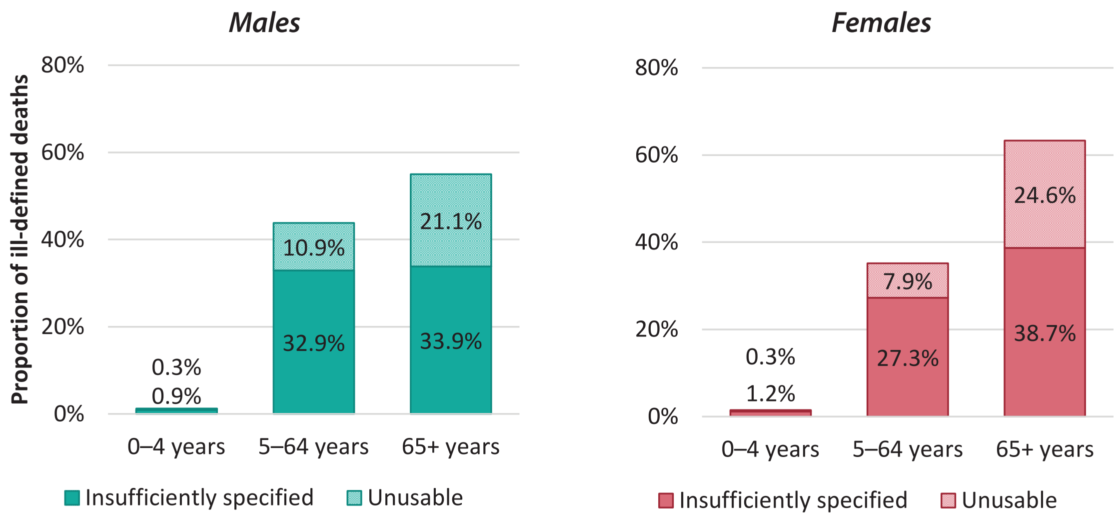
Examining the age and sex distribution of insufficiently specified and unusable codes helps to assess whether certain codes are more frequent in different population groups. As people age, they are more likely to suffer from diseases and conditions concurrently, making it difficult for physicians to identify a single underlying cause. 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 insufficiently specified and unusable codes should collectively not exceed 10% for deaths at ages 65 years and over and should be less than 5% for deaths in age groups below 65 years.[10]

As shown in ***Figure 23***, while the proportion of insufficiently specified and unusable codes is very low for young children (aged 0–4 years), it increases dramatically with age, accounting for 43.8% of deaths among males aged 5–64 years, and 55.0% for males aged 65+ years. For females, while the proportion is lower among women aged 5–64 years (at 35.2%), for older women the proportion is much higher: with 63.3% of all deaths either insufficiently specified or unusable.

10 *Analysing Mortality and Cause of Death (Version 3.1). Geneva: World Health Organization; 2021 (*https://www.who.int/standards/classifications/classification-of-diseases/services/analysing-mortality-levels-and-causes-of-death*).*

Within this broader category of "ill-defined", the proportion of deaths due with unusable causes (those with no value for policy or planning), at 21.1% for males and 24.6% for females aged 65+ years, indicates persistent data quality issues.

***Figure 23.** Proportion of deaths with unusable and insufficiently specified codes by sex and broad age group, 2020–2024 (all deaths combined)*



10.3 Ill-defined codes

This step tabulates specific ill-defined codes, ranks them by number and proportion of all deaths, and distributes them according to broad age groups. Information from this step can be used to improve medical certification and coding practices, by providing specific feedback on the most common ill-defined codes used. As shown in ***Table 12***, the most common ill-defined code is "septicaemia, unspecified", which is categorized as an unusable code, as it gives no indication of what caused the blood poisoning.[11] This is followed by "cardiac arrest, unspecified" and "senility".

***Table 12.** Number and proportion of ill-defined codes by age group, 2020–2024 (all deaths combined)*

ICD-10 code	Description	Number	Proportion of ill-defined (%)	Age group (years)			
				0–4	5–19	20–59	60+
A419	Septicaemia, unspecified	289	23.3	7	4	92	186
I469	Cardiac arrest, unspecified	124	10	0	2	36	86
R54	Senility	106	8.6	0	0	1	105
R570	Cardiogenic shock	84	6.8	0	2	26	56
Z515	Palliative care	69	5.6	1	1	15	52
I509	Heart failure, unspecified	65	5.3	0	0	25	40
N19	Unspecified renal failure	51	4.1	1	0	21	29
I10	Essential (primary) hypertension	40	3.2	1	0	0	39
N180	End-stage renal disease	32	2.6	0	0	13	19
R571	Hypovolaemic shock	32	2.6	1	0	9	22
—	**Total (top 10)**	**892**	**72.1**	**11**	**9**	**238**	**634**
—	All other codes	346	27.9	6	14	142	184
	Total (all ill-defined)	**1,238**	**100.0**	**17**	**23**	**380**	**818**

11 Septicaemia, or blood poisoning, can be caused by communicable diseases, NCDs, or injuries – meaning the death could be from any of the three broad GBD Groups, and as such, cannot be used for policy and planning.

Produced by the Pacific Community (SPC)
Pacific Community
B.P D5 – 98848 Noumea Cedex
New Caledonia

Telephone: +687 262000
Email: spc@spc.int
Website: www.spc.int