Institute
and Faculty
of Actuaries

IFoA Drivers of Mortality Dataset

Mapping drivers of mortality to England and Wales
Data for deaths occurring in 2001 to 2021

07 July 2024

Authors and Contributors

Name	Role	Contribution
Nana Asiamah	Longevity Actuary (Royal London)	Final Mapping of ICD codes to cause of death. Data construction and mapping of drivers. Documentation Content and structure
Naylor Guy	Pricing Actuary (Munich Re)	Initial Mapping of ICD codes to cause of death
Caroline Roberts	Senior Longevity Risk Manager (Phoenix)	Universe-of-Longevity-Catalysts-Call-For-Research.pdf (actuaries.org.uk)

Reviewers

Name	Role	Review Scope
Sacha Dhamani	Head of Longevity (Royal London) and Chair of the MRSC	Bottom-up review of mappings Review of documentation APS X2 Peer Review
Alison Yelland	Senior Longevity Actuary (Royal London)	Review of documentation
Sandesh Maharaj	Longevity Actuary (GenRe)	Independent replication of mappings Code review
Ben Rees	Longevity Actuary (LCP)	Independent review of age bands used in final data and spreading of deaths banned by age
Alan Kilpatrick	Analyst (M&G)	Review of mappings
Caroline Roberts	Senior Longevity Risk Manager (Phoenix)	Review of documentation

Contents

1   Introduction

1.1   Document Purpose

The paper documents the creation of an Institute and Faculty of Actuaries (IFoA) mapping of International Statistical Classification of Diseases version 10 (ICD-10) to 13 drivers of mortality on multi-mention England & Wales (EW) mortality data. The work was conducted in the Secure Research Service (SRS) network, part of the Office for National Statistics (ONS).

1.2   Audience

The intended audience of this paper are:

- SRS (ONS) who approved the publication of the "IFoA Drivers of Mortality Dataset".
- Institute and Faculty of Actuaries including:
  - Mortality & Morbidity Research Steering Committee (MRSC)
  - Population Mortality Improvements Working Party (initiated by the MRSC)
  - Continuous Mortality Investigation (CMI) Mortality Projections Committee (MPC)
- Wider public with interests in understanding the process followed in creating the drivers of mortality dataset.

1.3   Background

The CMI publishes the CMI Mortality Projections Model, which is a world-leading model that smooths recent trends in mortality and projects mortality improvements. However, there is no equivalent IFoA published robust and reliable method for deriving plausible population mortality improvement assumptions from more discrete data sources (such as a cause of death data), necessarily supported by subjective judgements.

The MRSC has set up a Population Mortality Improvements Working Party (the "Working Party") to produce a methodology for deriving plausible population mortality improvement assumptions for the UK general population. The methodology pursued by the working party is to:

- Create an IFoA mapping of cause of death data to the main Drivers of Mortality – i.e. groups of causes that are believed to respond similarly to changes in the underlying drivers of mortality such that the group of causes can be considered to be reasonably homogeneous.
- Consider how the patterns of mortality may change in the future as a result of longevity 'catalysts' e.g. ban on smoking. As an initial step to understanding these longevity catalysts and their potential impacts on different drivers of mortality, a paper 'Universe of Longevity Catalysts: Call For Research' has been written that sets out a universe of material catalysts and which drivers they are most likely to impact.

This paper covers the development of a 'drivers of mortality' dataset, using the ONS mortality data, with the aim of identifying the driver behind a death in the presence of multiple mentions of causes. There is an established method provided by WHO[1] to identify the primary cause of death, however this approach undergoes frequent changes which are not applied retrospectively, leading to discontinuities in the data. The proposed approach in

---

[1] ICD-10_2016_Vol2_PRINT.pdf (who.int)

this paper is applied consistently in all years, leading to a more consistent dataset for modelling purposes.

1.4 Document Scope

The intent of this document is to describe the:

- Methodology which covers reasons why this data is needed along with an overview of the process; see section 3.
- ONS mortality data which contains individual deaths data along with variables such as age, sex and list of recorded causes of death; see section 4.
- Mapping of ICD-10 codes to 'cause of death' including changes made to the data and classification of the ICD codes; see section 5.
- Mapping of the drivers of mortality to multi-mention data which covers the list of drivers and related causes, along with judgements made in the mapping exercise; see section 6.
- Specification of the dataset being published; see section 7.
- Analysis of the resultant dataset including an assessment of COVID and the new Co-morbidity driver; see section 8.
- Validation of the underlying data and process taken to derive the data being published; see section 9.
- List of limitations and potential further developments; see section 10.

Recommendations on catalysts of mortality or modelling of the data is out of scope of this paper. Drivers derived in this paper should not be used for modelling purposes without fully understanding the underlying diseases mapped and their relationship to the specific catalyst being modelled.

1.5 Actuarial Standards

The work carried out to prepare this document has been done in accordance with TAS 100: General Actuarial Standards Version 2.0.

1.6 ONS Disclaimer

This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

1.7 IFoA Disclaimer

This document has been prepared by and/or on behalf of the IFOA. The IFoA does not accept any responsibility and/or liability whatsoever for the content or use of this document. Whilst care has been taken during the development of the document, IFoA does not (i) warrant its accuracy; or (ii) guarantee any outcome or result from the application of this document (whether contained in or arising from the application of this document or otherwise). You assume sole responsibility for your use of this document, and for any and all conclusions drawn from its use. IFoA hereby excludes all warranties, representations, conditions and all other terms of any kind whatsoever implied by statute or common law in relation to this document, to the fullest extent permitted by applicable law. If you are in any doubt as to using anything produced by IFoA, please seek independent advice.

2 Executive Summary

2.1 Introduction

This section provides a summary of the process followed to create the drivers of mortality dataset and covers the following elements:

- Methodology
- Data
- Mapping of ICD-10 Codes to Cause of Death
- Mapping of the Drivers of Mortality
- IFoA Drivers Of Mortality Dataset
- Analysis of Dataset
- Validation
- Limitations and Potential Further Development

2.2 Methodology

This paper covers the development of a 'drivers of mortality' dataset, using the ONS mortality data, with the aim of identifying the driver behind a death in the presence of multiple mentions of causes. There is an established method provided by WHO[2], however this approach undergoes frequent changes which are not applied retrospectively leading to discontinuities in the data. The proposed approach in this paper is applied consistently in all years leading to a more consistent dataset for modelling purposes.

2.3 Data

The drivers of mortality dataset has been developed using:

- International Statistical Classification of Diseases (ICD), a core statistically based classificatory diagnostic system for health care related issues[3]. It provides a common language that allows health professionals to share standardised information across the world[4]. ICD version 10 is used in this project due to its availability in the ONS mortality data. A further development will be allow for changes when ICD11 is adopted by the ONS.
- The ONS mortality data is a series of annual datasets, that capture deaths that occurred in England and Wales along with variables including age, sex, calendar year (year of occurrence) and up to 15 mentions of ICD-10 code per death.

2.4 Mapping of ICD-10 Codes to Cause of Death

All ICD code mentions (up to 15 on each death) are used in the process. The ICD codes are mapped to a Cause of Death classification created by the Working Party. The mapping of each ICD-10 code mention to a cause of death expands upon the official ONS mapping[5]and the main differences:

- Over 95% of deaths are mapped consistently.

---

[2]ICD-10_2016_Vol2_PRINT.pdf (who.int)
[3]International Classification of Diseases (who.int)
[4]WHO's new International Classification of Diseases (ICD-11) comes into effect
[5]Leading causes of death in England and Wales (revised 2016) - Office for National Statistics (ons.gov.uk)

- Approximately 1.4% of deaths remain unmapped based on the ONS classification which does not cover all ICD Codes. This number is low enough to be ignored for the intended use of the dataset.
- Differences are observed for 3.2% of deaths. The main reasons for the differences are:
  - Over 54% of the differences are due to a decision to move the ICD code R540 which represents senility [6] away from "Other" classification into "Alzheimer's Disease and Dementia".
  - Approximately 33% of the differences are due to better placement of cancers.

See section 6 for more information.

2.5   Mapping of Cause of Death to Drivers of Mortality

The mapping approach is to take the cause of death for each available mention and determine an overall Driver for each death, the underlying reason for the death to have occurred.

2.5.1   List of Drivers

The drivers identified for this project are presented in Table 1.

Table 1 – The Drivers of Mortality

Driver	Description
Non-Disease	Covers deaths due to accidents, suicide and other external causes. It is differentiated due to its distinctive influences i.e. not driven by disease.
Cancer – Unmet need	Cancer Research UK (CRUK) identified 4 cancers of unmet need [7] (lung, pancreatic, oesophageal and brain) which share poor five-year survival rates. These 4 cancers have been grouped together to reflect late diagnosis, poor survival rates and limited catalysts likely to improve death rates in the short term.
Cancer – Smoking Related	Aside from medical advancements (which are likely to impact cancer across the board), the main risk factors for cancer are smoking [8] and obesity [9]. The CRUK website lists the cancers that are linked to smoking and obesity and these lists have been used to create the related cancer groups.
Cancer – Obesity Related	
Cancer – Other	The remaining cancers are grouped together as 'other' with Prostate cancer making up over 21% of these cancers; other notable cancers in his group include Lymphoma (11%), Mesothelioma (6%), Melanoma (6%) and skin cancer (2%).
Respiratory Disease – smoker related	Covers respiratory causes of death with a strong link to smoking i.e. Chronic obstructive pulmonary disease (COPD), Bronchial Disease and Emphysema.
Alzheimer's Disease and Dementia	… is differentiated based on its magnitude and significance.
Diabetes and Obesity	… is differentiated based on its recent and potential future significance.
Cardiovascular Disease	Due to similarities of the drivers leading to stroke and coronary heart disease, causes under these are clustered together.

---

[6] ICD-10 Version:2019 (who.int)
[7] Research opportunities in cancers of unmet need | Cancer Research UK
[8] How does smoking cause cancer? | Cancer Research UK
[9] How does obesity cause cancer? | Cancer Research UK

Driver	Description
Infection	Covers all infections such as influenza and HIV. It has been included to reduce the size of the "Other" group.
COVID	… is differentiated due to its significance and inflated levels in recent years.
Other	Captures all other causes including liver disease, other respiratory conditions, neurological conditions, kidney damage, asthma and newborn conditions.
Co-morbidity	A co-morbidity driver group has been created which is defined as where a death has three or more distinct mentions in selected categories; see section 6.7.

These drivers are discussed further in section 6.4.

2.5.2   Mapping Overview

The steps taken in mapping a driver of mortality to a death are:

- The order of the mentions provided by the ONS is as listed on the death certificate and requires re-ordering to prioritise the underlying cause of death as separately recorded on the death certificate.
- The re-ordered list of the mentions is used in determining the driver classification e.g. respiratory (smoker) disease is used for deaths where it is mentioned before, say, "Alzheimer's Disease and Dementia".
- In determining the proposed driver classification for a death, distinctions are made between events i.e. events that occur as a result of an underlying cause, e.g. cardiac arrest and drivers i.e. what triggered the event of death e.g. cancer.
- Certain causes are assumed to be more dominant e.g. a mention of cancer, excluding prostate or benign cancer, overshadows other mentions e.g. pneumonia.
- Multiple and distinct mentions of significant causes are mapped to Co-morbidity.

2.6   IFoA Drivers of Mortality dataset

The fields included in the dataset ("IFoA Drivers of Mortality Dataset"), planned to be released are:

- Mapping i.e. covers the proposed mapping and alternative versions of the mapping e.g. a version where COVID is treated similar to seasonal flu and one where it is its own driver.
- Age (Last birthday) is provided at an integer level up to where the death count is greater than 2 for each sex, year and driver combination. Below this, the ages are banded to derive clusters with more than 2 deaths.
- Sex is categorised as M (male) or F (female).
- Calendar Year i.e. the year death occurred. An alternative dataset has been created using year of death registration to demonstrate consistency with ONS publications, and the data used by the CMI in its projection model.
- Driver Of mortality.
- Number of deaths in each age, sex, year and driver cluster for each dataset.

2.6.1   Year of Occurrence

The year of death occurrence is believed to be a more suitable/reliable for the modelling of drivers however its use requires consideration of incurred, but not reported (IBNR) deaths . Further details can be found in section 9.6.

2.7 Analysis of Dataset

The following observations are noted in the drivers of mortality dataset:

- Cardiovascular Disease is the leading driver of mortality (20% of all deaths) for both sexes occurring in the period 2001-2021; there has been a decline in deaths associated with this driver in more recent years.
- Post 2014, Co-morbidity is the leading driver with the group "Alzheimer's Disease and Dementia" being the single leading cause of death.
- A quarter of all deaths are linked to Cancer with Cancer – Obesity Related deaths being most prevalent for females and Cancer – Unmet Need deaths being most prevalent for males.
- Most of the drivers peak in 2020 and fall in 2021 as more deaths are recorded under the COVID driver.
- The Non-disease and Other driver groups are dominant at younger ages. At older ages, the leading drivers shift from cardiovascular and cancers during earlier years to "Alzheimer's Disease and Dementia" and Co-Morbidity in more recent years.
- Over 50% of deaths occur between ages 80-99 as a significant proportion of the population survives to these ages where they become increasingly susceptible to disease.

2.7.1 Co-Morbidity

The most significant change from dataset produced by the ONS is the inclusion of a Co-Morbidity driver and a significant and increasing number of deaths are allocated to this driver.

The Proposed Mapping approach attributes close to two million deaths to the co-morbidity driver; 15% of all deaths. The Co-morbidity driver absorbs deaths from other drivers and in its absence, these deaths would have fallen under other drivers. Co-morbidity trends upwards and is the driver with the highest number of deaths from 2014 onwards; over 20% of deaths post-2013 are classed under this driver.

2.8 Validation

Validation:

- Checks were performed on the data to assess consistency of variables e.g. sex (M/F) and with ONS's "21st century mortality" dataset.
- Checks have been performed on the end-to-end process, to ensure deaths were not added or excluded for no valid reason.

The mapping approach, which is formulaic, has been subject to manual spot-checks on a significant subset of the data by Sacha Dhamani (SD) and meets his expectations with few discrepancies. These discrepancies are noted as a difference in opinion with low materiality.

- In comparison to a dataset with drivers formulated using the ONS Primary ICD Code:
  - Consistency in mapping ranges from 36.5% (Infection) to 100.0% (COVID). This is due to the proposed mapping focusing on the drivers of mortality opposed to the effects with COVID prioritised in both mapping.
  - Ignoring COVID the cancers show the highest consistency led by Cancer - Unmet Need (90.9%).

o Co-Morbidity, which is not a driver in the ONS approach, absorbs the most deaths (15.2%) with deaths classified as Diabetes and Obesity deaths by the ONS showing the most significant movement (53.2%) into Co-Morbidity.

2.9 Limitations and Potential Further development

2.9.1 Limitations:

- Information stored on death certificates evolve over time due to increasing demands for better data collection. For example, the change in trend relating to number of mentions recorded on death certificates could be due to recording changes and result in inaccurate classification of drivers e.g. Co-morbidity which relies on the number of mentions or the absence of a mention which outranks another e.g. a cardiomyopathy death with diabetes, which outranks cardiovascular disease, not recorded.
- The IFoA Cause of Death mapping list covers all ICD codes in the mortality data however it does not cover all ICD codes in the WHO library. The list needs to be expanded and reviewed periodically or when changes are made to the WHO library.

2.9.2 Potential Areas of Further Developments:

- Consider alternative mapping rules e.g. alternative definitions of co-morbidity – such as splitting by the most significant cause indicated.
- Deaths reported pre-2001 are excluded due to a lack of ICD-10 mapping. This could be resolved by, for example, creating a cause of death mapping for ICD-9 and other previous versions[10].

2.10 Future Releases

This dataset is intended to be periodically updated for additional deaths experience, but the operational process is yet to be determined. It is expected that feedback on the usefulness of the data will be captured over 2024 which will inform how frequently the data will be updated.

---

[10]International Classification of Diseases (ICD) (who.int)

3   Methodology

3.1   Introduction

This section outlines the methodology in developing the IFoA Cause of Death dataset and covers the following elements:

- Need for an IFoA Cause of Death Dataset
- ICD Code Mapping
- Driver Mapping

3.2   Need for an IFoA Cause of Death Dataset

The ONS has published datasets containing cause of death over many decades and different ICD regimes. This data has proved to be useful to actuaries in understanding the variation in mortality improvements by different causes.

However, there are some limitations with the data provided that the development of an IFoA Cause of Death Dataset is seeking to address:

- There is an established method provided by the WHO[11], to identify the cause of death in the presence of multiple mentions. This approach undergoes frequent changes which are not applied retrospectively by the ONS, leading to discontinuities in the data. The proposed approach in this paper is applied consistently in all years for the ICD10 regime leading to a more reliable dataset for modelling purposes.
- The WHO methodology used by the ONS, currently under reports deaths for some key drivers such as "Alzheimer's Disease and Dementia" and Diabetes - although these are key drivers in the UK and other developed nations, that is not necessarily the case in the global context. By recognising the presence of these drivers at lower mentions this underreporting can be addressed.
- The ICD code regime is intended for use for a number of purposes and as a result the mapping of ICD code to cause may not be appropriate for cause of death modelling purposes. The ONS ICD code mapping is amended for the purposes of the IFoA cause of death dataset.

3.3   ICD Code Mapping

The first step in creating the IFoA Cause of Death Dataset is the mapping of ICD codes (for each mention recorded for each death) to a subset of causes. This mapping is described further in section 5 including the changes from the ONS mapping. This creates a more manageable classification of deaths based upon which the underlying driver of death can be identified.

3.4   Driver Mapping

The second step in creating the IFoA Cause of Death Dataset is the mapping of causes at a mention level to a single Driver for each death. These drivers are intended to:

- Capture the material variations in the sub-trends of mortality improvement.
- Be sufficiently few in number such that the deaths split by these drivers are sufficiently credible to allow modelling of mortality improvement by driver.

---

[11] ICD-10_2016_Vol2_PRINT.pdf (who.int)

The Driver mapping is described further in section 6 and follows the following principles:

- All ICD code mentions (up to 15 on each death) are used in the process. The ICD codes are mapped to the IFoA Cause of Death classification.
- The order of the mentions provided by the ONS is as listed on the death certificate and requires re-ordering to prioritise the underlying cause of death as separately recorded on the death certificate.
- The order of the mentions, after reordering, is used in determining the driver classification e.g. respiratory (smoker) disease is used for deaths where it is mentioned before, say, "Alzheimer's Disease and Dementia".
- In determining the proposed driver classification for a death, distinctions are made between:
  - Events i.e. events that occur as a result of an underlying cause, e.g. pneumonia and influenza, including events that are typically mentioned but not the underlying driver of death e.g. prostate cancer.
  - Drivers i.e. what triggered the event of death e.g. cancer.
- Certain causes are assumed to be more dominant e.g. a mention of cancer, excluding prostate or benign cancer, overshadows other mentions e.g. pneumonia.

The proposed approach resultants in more consistent trends which are covered in section 9.7.

4 Data

4.1 Introduction

This section discusses the source data and covers the following elements:

- ICD-10 Classification
- ONS Mortality Data
- Number of Mentions

along with limitations and further development.

4.2 ICD-10 Classification

International Statistical Classification of Diseases (ICD) is a core statistically based classificatory diagnostic system for health care related issues[12]. It provides a common language that allows health professionals to share standardised information across the world[13].

ICD version 10 is used in this project due to its availability in the ONS mortality data. The approach the ONS follows to coding ICD-10 to the raw data undergoes periodic updates which alters the interpretation of death certificate for a small number of deaths. Deaths registered from 1st January 2020 use MUSE 5.5 and from 1st January 2022 MUSE 5.8. The latest generation of ICD codes (version 11) is not yet available in the ONS data. For more information on version 10 see ICD-10 Version:2019 (who.int).

4.3 ONS Mortality Data

The ONS mortality data is a series of annual datasets that capture deaths that occurred in England and Wales along with variables including:

- Age (Last birthday) is provided at an integer level. Nulls occurring in the 2001 dataset are removed; 5,739 deaths (1.00% of this particular file or 0.05% of the entire dataset) which also have variables like sex and death date missing. See section 9.2 for more details.
- Sex is banded M (male) or F (female). Unknown/Missing variables are defaulted to the Male sex as done by the ONS in its published datasets; 138 deaths (0.00%).
- Calendar Year i.e. the year death occurred, which is more suitable for the modelling of drivers. Note that the CMI use the year a death was registered/reported due to lags in death reporting and for consistency reasons with the ONS[14], who note that 99% of deaths are registered within a year. Due to reporting delays, this paper recommends the 2021 data is discarded or an adjustment is made to make it more useable; this is explored further in section 9.6.. Although the ONS mortality data covers deaths reported pre-2001, this project utilises deaths from 2001 onwards. This is because older deaths have not been assigned an ICD-10 classification; this will be explored as a further development.
- ICD-10 classification: All ICD-10 code mentions, up to 15 codes on each death, is used to create the data; see section 5 for more details. Note that:

---

[12] International Classification of Diseases (who.int)
[13] WHO's new International Classification of Diseases (ICD-11) comes into effect
[14] Impact of registration delays on mortality statistics in England and Wales - Office for National Statistics (ons.gov.uk)

- o Unknown/Missing ICD-10 codes linked to new-born deaths, are defaulted to either the ONS Primary ICD Code or "Other" if the primary is also blank; 41,337 deaths (0.38%).
- o The ONS primary is determined based on rules provided by the WHO[15].

The ONS primary is determined by linking each mention to an ICD-10 code and passing the data through MUSE 5.8 software noting that there are regular updates to this system which are not applied retrospectively. New versions of the ICD framework are introduced at the start of a year, so each full year of data is coded to that standard. Tweaks occur during the year, for example where inconsistencies or errors are identified. However, once an annual dataset has been published it will only be updated in exceptional circumstances, and never to update it to a new version of ICD. In addition, the determination of the primary cause of death undergoes changes[16] which again is not applied retrospectively. These issues with the ONS primary code are addressed by this project. For more information on the ICD framework contact: health.data@ons.gov.uk.

The data has been validated (by age band, sex and year of reporting) against the ONS' 21st Century Mortality Dataset and the two sets of data are consistent; see section 9.2.

A full list of variables in this dataset can be found in Appendix A – ONS Mortality Data.

4.4 Number of Mentions

Figure 1 to Figure 4 show the proportion of mentions by year of death for the age groups all ages, under 60, 60-79 and 80+ respectively. For all ages, it suggests a convergence in the number of mentions to three mentions, which is of concern, if due to reporting changes. There is a distortion in the trends in 2020-21 due to COVID. Split by age, there is evidence this convergence has a strong age component:

- Under 60s: The change in pattern is observed primarily for two mentions. Other mention clusters appear flat or have a gentler slope. No mentions occur for under 60s only with most related to newborn causes.
- 60-79s: The steepest slope is observed for two mention cases with is in decline. All other mention clusters are rising however more gently.
- 80+: The convergence observed at an aggregate level is driven by this age group which shows steep declines in one and two mentions with rises observed for higher mentions.

It is difficult to decompose the components that give rise to these patterns. It could be a result of ageing and improvement in longevity (single conditions are easier to manage); some of the patterns could be driven by reporting changes. This uncertainty is highlighted for the user of the data to consider how they interpret the results.

---

[15]ICD-10_2016_Vol2_PRINT.pdf (who.int)
[16]Cause of death coding in mortality statistics, software changes - Office for National Statistics (ons.gov.uk)

Figure 1 – The Proportion of Mentions by Year of Death
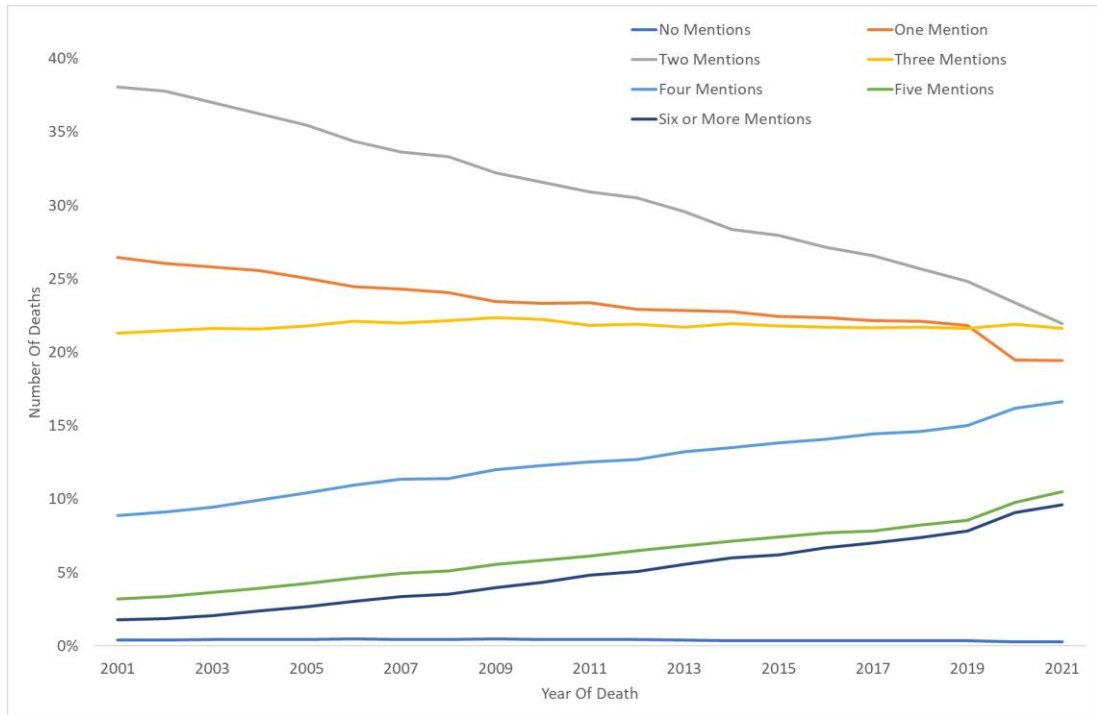
Figure 2 – The Proportion of Mentions by Year of Death – Under 60s
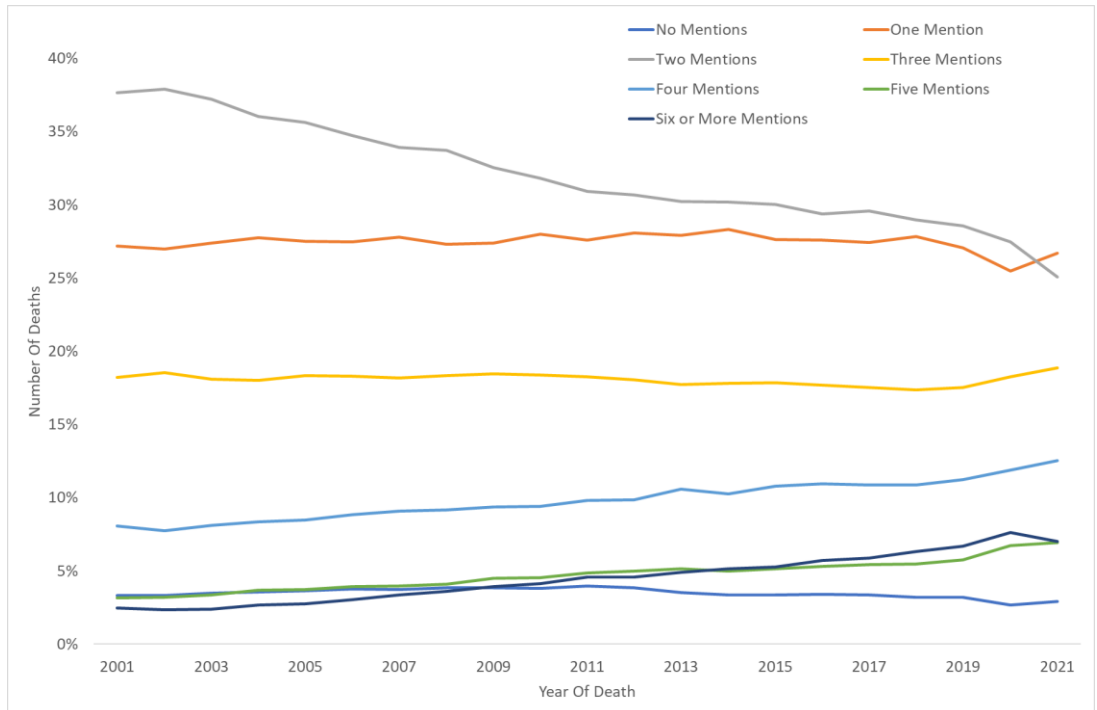
Figure 3 – The Proportion of Mentions by Year of Death – 60 to 79
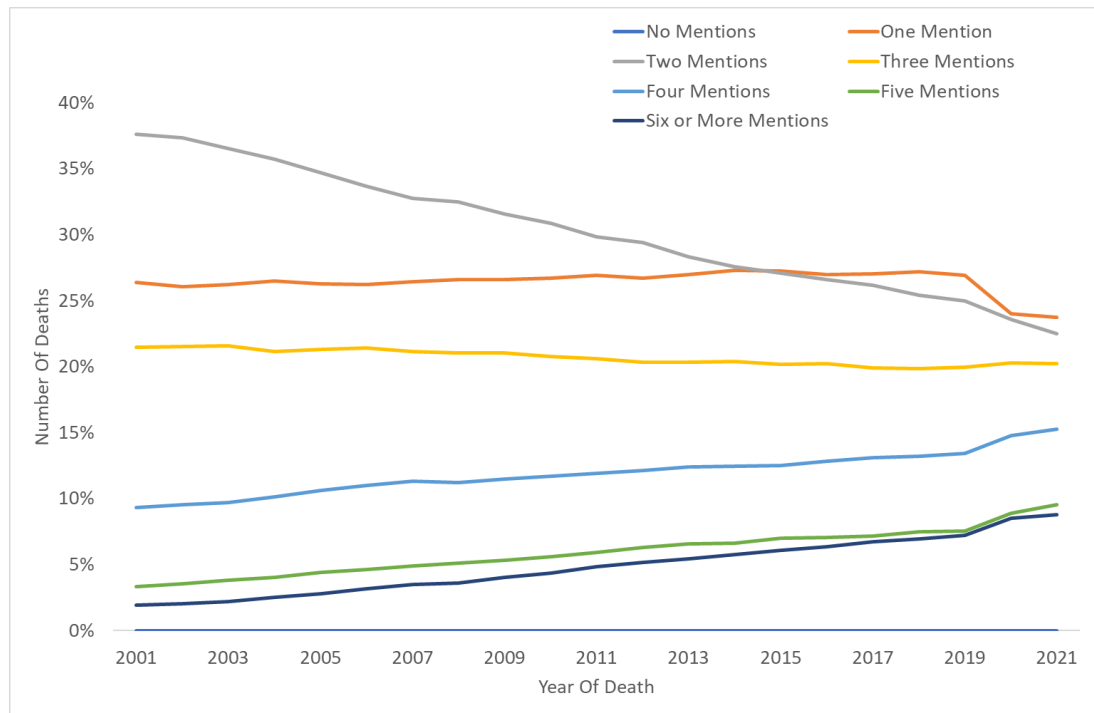
Figure 4 – The Proportion of Mentions by Year of Death – Over 80s
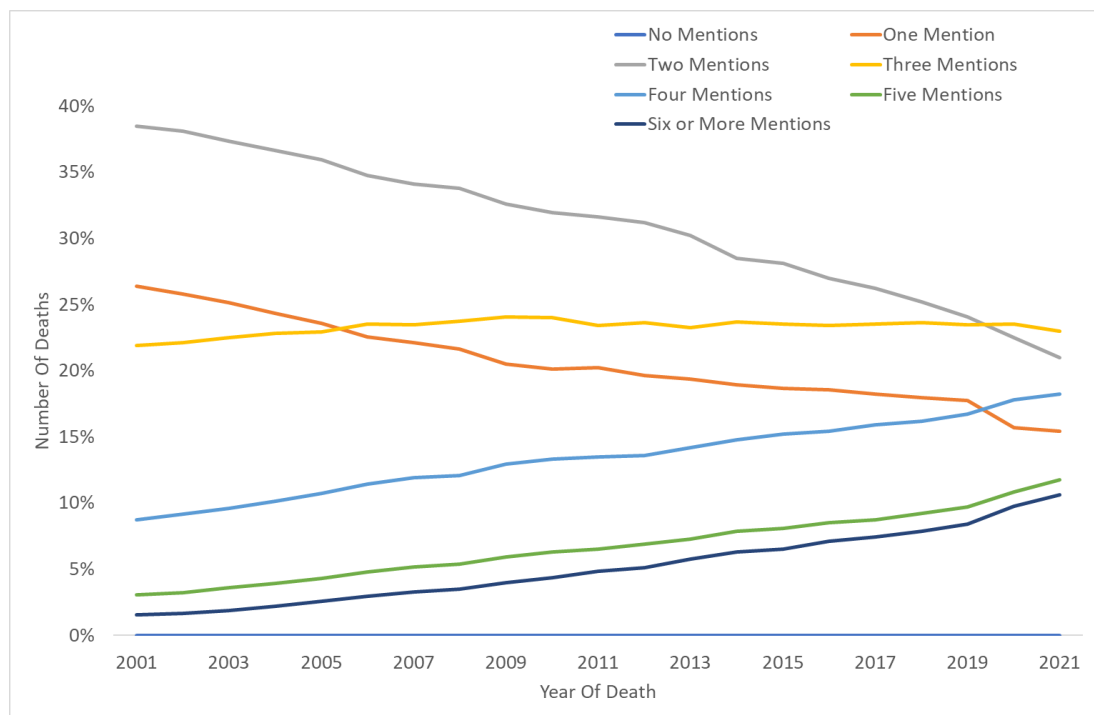
4.5   Limitations

Information recorded on death certificates evolve over time due to increasing demands for better data collection. For example, the change in trend relating to number of mentions recorded on deaths could be due to recording changes and result in inaccurate classification of drivers e.g. Co-morbidity which relies on the number of mentions or the absence of a

mention which outranks another e.g. a cardiomyopathy death with diabetes, which outranks cardiovascular disease, not recorded.

4.6   Further Development

Deaths reported pre-2001 are excluded due to a lack of ICD-10 mapping. This could be resolved by:

- Creating a cause of death mapping for ICD-9 and other previous versions[17]
- A manual exercise or application of a free text classification program to map death descriptions to the ICD-10 codes.

---

[17] International Classification of Diseases (ICD) (who.int)

5 Mapping of ICD-10 Codes to Cause of Death

5.1 Introduction

This section covers the mapping of ICD-10 codes (12,000+) to 87 identified of causes of death. The classification builds on a version[18] created by the ONS (see Appendix B – ICD Code to ONS Cause of Death) and in this section we cover:

- Background
- Renaming and aggregation of the causes
- Analysis of Change
- Limitations

The IFoA cause of death mapping can be found in Appendix C – ICD Code to IFoA Cause of Death.

5.2 Background

The IFoA have created a comprehensive list of ICD- 10 mapping to cause of death. This work was initiated by Guy Naylor (GN) and reviewed by Sacha Dhamani (SD). Unallocated deaths were discovered after mapping of this list to the ONS mortality data and the additional mapping rules were created by Nana Asiamah (NA), with SD acting as a reviewer. These gaps were primarily ICD-codes starting with T, V, W, X and Y used for injury, medical treatment complications, vehicular accidents, accidents, natural disasters, suicide, drug-alcohol, and external causes. An ONS file found within the SRS was used to provide descriptions leading to a satisfactory identification of the 'cause of death'. Further gaps were discovered by Alan Kilpatrick (AP) after a comparison of the list against the WHO library[19], although these gaps are ICD codes which do not appear in the England and Wales dataset and will be considered as a further development.

Compared to ONS cause of death classification, the IFoA's list aims to:

- Add more granularity by using the full ICD code to determine the cause of death in all instances.
- Aggregate diseases of less relevance to the England and Wales e.g. Tuberculosis is now covered within Infection.
- Disaggregate important causes e.g. Hodgkin's Disease, Lymphoma, Myeloma or Leukaemia classed by the ONS in a single cause is now split into separate causes.
- Rename the causes to make them more intuitive to readers without a medical background .
- Ensure coverage of all ICD codes found in the data e.g. the ONS classification does not cover X33 which maps to Natural Disasters.

5.3 Renaming and Aggregation of Causes

---

[18]Leading causes of death in England and Wales (revised 2016) - Office for National Statistics (ons.gov.uk)
[19]ICD-10 Version:2019 (who.int)

Table 2 shows the ONS cause of death and what it's been mapped to in the new classification created by the IFoA.

Table 2 – Renaming of the ONS 'Cause of Death' Classification

ONS Cause of Death	IFoA Cause of Death
Intestinal infectious diseases	Infection
Tuberculosis	Infection
Vector–borne diseases and rabies	Infection
Vaccine-preventable diseases	Infection
Meningitis and meningococcal infection	Infection
Septicaemia	Infection
Human immunodeficiency virus [HIV] disease	Infection
Malignant neoplasms	Cancer (with granularity added)
Malignant neoplasm of oesophagus	Oesophagus Cancer
Malignant neoplasm of stomach	Stomach Cancer
Malignant neoplasm of colon, sigmoid, rectum and anus	Colo-Rectal Cancer
Malignant neoplasm of liver and intrahepatic bile ducts	Liver Cancer
Malignant neoplasm of gallbladder and other parts of biliary tract	Gallbladder Cancer
Malignant neoplasm of pancreas	Pancreatic Cancer
Malignant neoplasm of larynx	Head and Neck Cancer
Malignant neoplasm of trachea, bronchus and lung	Cancer of trachea, bronchus and lung
Malignant neoplasms of bone and articular cartilage	Bone Cancer
Melanoma and other malignant neoplasms of skin	Melanoma or Skin Cancer
Malignant neoplasm of breast	Breast Cancer
Malignant neoplasm of uterus	Uterine Cancer
Malignant neoplasm of ovary	Ovarian Cancer
Malignant neoplasm of prostate	Prostate Cancer
Malignant neoplasm of kidney, except renal pelvis	Kidney Cancer
Malignant neoplasm of bladder	Bladder Cancer
Malignant neoplasm of brain	Brain Cancer
Malignant neoplasms, stated or presumed to be primary of lymphoid, haematopoietic and related tissue	Hodgkin's Disease, Lymphoma, Myeloma or Leukaemia
In situ and benign neoplasms, and neoplasms of uncertain or unknown behaviour	Cancer (with granularity added)
Diabetes	Diabetes
Malnutrition, nutritional anaemias and other nutritional deficiencies	Blood Disorders or Other
Disorders of fluid, electrolyte and acid–base balance (incl. dehydration)	Metabolic Disorders or Other
Dementia and Alzheimer disease	Alzheimer's Disease and Dementia
Mental and behavioural disorders due to psychoactive substance use	Mental Health Conditions or Drug-Alcohol
Systemic atrophies primarily affecting the central nervous system	Neurological Conditions or Other
Parkinson disease	Parkinson's Disease
Epilepsy and status epilepticus	Neurological Conditions
Cerebral palsy and other paralytic syndromes	Cerebral Palsy and other paralytic Syndromes
Chronic rheumatic heart diseases	Valve Disorder or Other CHD
Hypertensive diseases	Other CHD
Ischaemic heart diseases	Ischaemic Heart Diseases incl. AMI
Pulmonary heart disease and diseases of pulmonary circulation	Pulmonary Heart Disease

ONS Cause of Death	IFoA Cause of Death
Nonrheumatic valve disorders and endocarditis	Valve Disorder, Inflammation of the Heart or Other CHD
Cardiomyopathy	Cardiomyopathy
Cardiac arrest	Cardiac arrest
Cardiac arrhythmias	Other CHD
Heart failure and complications and ill–defined heart disease	Heart Failure or Other CHD
Cerebrovascular diseases	Cerebral Infarction, Intracranial Haemorrhage, Stroke or Other Cerebrovascular Disease
Atherosclerosis	Atherosclerosis
Aortic aneurysm and dissection	Aneurysm
Acute respiratory infections other than influenza and pneumonia	Respiratory Infection, Other Respiratory Conditions or Bronchial Disease
Influenza and pneumonia	Influenza and Pneumonia
Chronic lower respiratory diseases	Emphysema, COPD, Asthma or Other Respiratory Conditions
Pulmonary oedema and other interstitial pulmonary diseases	Other Respiratory Conditions
Respiratory failure	Other Respiratory Conditions
Appendicitis, hernia and intestinal obstruction	Other
Cirrhosis and other diseases of liver	Liver Disease
Diseases of the musculoskeletal system and connective tissue	Musculo Skeletal
Diseases of the urinary system	Kidney Disorder or Urological Conditions
Pregnancy, childbirth and the puerperium	Pregnancy Related
Certain conditions originating in the perinatal period	Newborn Conditions
Congenital malformations, deformations and chromosomal abnormalities	Congenital Malformation
Accidents	Accident
Land transport accidents	Vehicle Accident
Accidental falls	Accident
Non-intentional firearm discharge	Accident
Accidental drowning and submersion	Accident
Accidental threats to breathing	Accident
Accidental poisoning	Accident
Suicide and injury/poisoning of undetermined intent	Suicide
Homicide and probable homicide	Other
Symptoms, signs and ill-defined conditions	Other

5.4   Analysis of Change

The ONS classification is mostly based on the first 3 characters of an ICD code or broader classification using the first letter. The original ONS classification has been modified to make it more comparable to the IFoA list with assumptions:

- Where the ONS classification maps to multiple causes it is broken down as per the IFoA approach.

- Codes starting with V, W and X representing accidents are broken down as per the IFoA approach.

- Codes starting with A and B and not in original list would have been mapped to "Infection" by the ONS.

- Codes starting with J and not in original list would have been mapped to one of the respiratory classifications.
- Codes starting with K and not in original list would have been mapped to either Liver Disease or Digestive Disorders.
- Codes starting with L would have been mapped to skin diseases.
- COVID is mapped based on the ICD codes U07, U099 and U10.
- The ICD codes I739 and I330 are assigned to cardiovascular conditions.
- The ICD codes G350, G318, G931 and G231 is assigned to Neurological conditions.

Table 3 details the results of the comparison based on the 21st century mortality data. In summary:

- Over 95% of deaths are mapped consistently.
- Approximately 1.4% of deaths remain unmapped based on the ONS classification and assumption above. This number can be brought down by making more assumptions on how the ONS would have mapped the ICD-codes: however, the number is low enough to be ignored.

Table 3 – Comparison of the ONS Original List and Assumed Mapping to the IFoA Cause of Death

	Original List	Assumed Mapping	Total
Consistent	9,635,821	798,979	10,434,800
Different	330,991	16,677	347,668
Unmapped		155,271	155,271
Total	**9,966,812**	**970,927**	**10,937,739**

Approximately 3.2% of deaths are mapped differently. Table 4 shows these differences and in summary:

- Over 54% of the differences are due to the IFoA's decision to move the ICD code R540 which represents senility[20] away from "Other" classification. This is a material group which has been clustered with "Alzheimer's Disease and Dementia" in determining the driver of mortality. Senility mentions are weighted towards earlier periods; 30% of primary mentions in the study period 2001-2021 occurred in 2001-2005. Its inclusion with the "Alzheimer's Disease and Dementia" helps to lift the pattern upwards in earlier periods where this group is known to have been under reported; see section **Error! Reference source not found.**.
- Approximately 33% of the differences are due to better placement of cancers.
- Approximately 7% of the differences are due to the use of "Unknown" opposed to "Other" to separate out deaths pending investigation and/or where they have been ill-defined.
- The remaining 6% are due to introduction of more granularity by the IFoA e.g. splitting out codes that fits more naturally under "Other" from Digestive Disorders e.g. Hernia.

---

[20]ICD-10 Version:2019 (who.int)

Table 4 – Differences between the IFoA and ONS classification

IFoA Cause of Death	Cancer Other	Digestive Disorders	Infection	Ischaemic Heart Diseases	Other	Suicide	Valve Disorder
Cancer of trachea, bronchus and lung	1,141						
Colo-Rectal Cancer	7,990						
Digestive Disorders					783		
Disorders of Psychological Development					8		
Gynaecological Cancer	12,448						
Head and Neck Cancer	49,560						
Kidney Cancer	678						
Lymphoma	80						
Mental Health Conditions					58		
Mesothelioma	42,365						
Neurological Conditions			30		35		
Other		16,677					
Other CHD				2,069			380
Other Respiratory Conditions					311		
Respiratory Tract Conditions					154		
Unknown					24,937		
Urological Conditions					6		
Uterine Cancer	25						
Vehicular Accident						13	
Senility					187,920		

5.5   Limitations

The IFoA Cause of Death mapping list covers all ICD codes in the mortality data however it does not cover all ICD codes in the WHO library. The list needs to be expanded and reviewed periodically or when changes are made to the WHO library.

6 Mapping of the Drivers of Mortality

6.1 Introduction

This section covers the mapping of drivers to the reordered causes of death. This section covers:

- Overview of Approach
- Reordering of Mentions
- Drivers of Mortality
- Single Mention Mapping
- Multiple Mention Mapping
- Co-Morbidity
- Alternatives
- Judgements

6.2 Overview of Approach

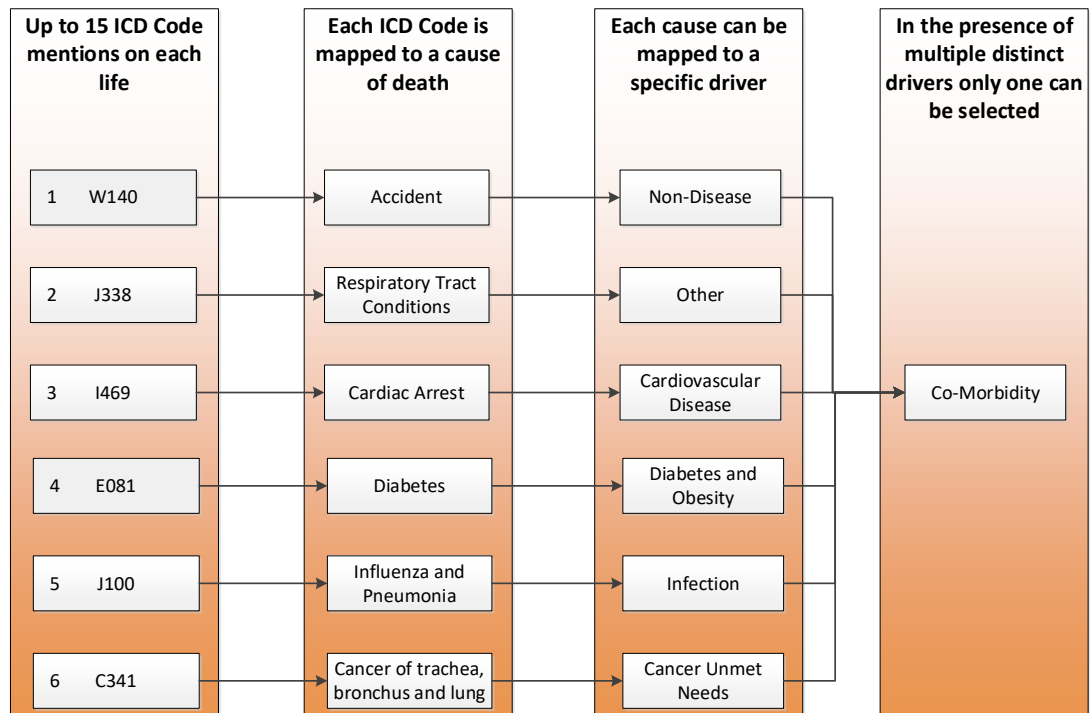The steps taken in mapping a driver of mortality to a death are:

- The order of the mentions provided by the ONS is as listed on the death certificate and requires re-ordering, see section 6.3, to prioritise the underlying cause of death as recorded in the death certificate.
- The order of the mentions is used in determining the driver classification e.g. respiratory (smoker) disease is used for deaths where it is mentioned before, say, "Alzheimer's Disease and Dementia".
- In determining the proposed driver classification for a death, distinctions are made between:
  - o Events i.e. events that occur as a result of an underlying cause, e.g. pneumonia and influenza, including events that are typically mentioned but not the underlying driver of death e.g. prostate cancer.
  - o Drivers i.e. what triggered the event of death e.g. cancer.
- Certain causes are assumed to be more dominant e.g., a mention of cancer, excluding prostate or benign cancer, overshadows other mentions e.g. pneumonia.
- Multiple and distinct mentions of significant causes are mapped to Co-morbidity.

A decision tree is used to identify the driver for a particular death. To avoid repetition, an example of process is shown in Figure 5; the full representation can be found in section 6.6.

In this example:

- Accident, cardiac arrest and Influenza & Pneumonia are treated as events due to the presence of more significant causes.
- Cancer of trachea, bronchus and lung is assumed to be more dominant than the mention of diabetes.
- The final driver is, however, Co-morbidity due to the mention of multiple significant causes i.e. Cancer of trachea, bronchus and lung, Diabetes and Respiratory Tract Conditions.

Figure 5 – An example of the mapping approach

Up to 15 ICD Code mentions on each life	Each ICD Code is mapped to a cause of death	Each cause can be mapped to a specific driver	In the presence of multiple distinct drivers only one can be selected
1 W140	Accident	Non-Disease	
2 J338	Respiratory Tract Conditions	Other	
3 I469	Cardiac Arrest	Cardiovascular Disease	Co-Morbidity
4 E081	Diabetes	Diabetes and Obesity	
5 J100	Influenza and Pneumonia	Infection	
6 C341	Cancer of trachea, bronchus and lung	Cancer Unmet Needs	

6.3   Reordering of the Mentions

The mentions in the ONS mortality data are ordered as recorded on a death certificate. Figure 6 shows an example of a death certificate and in this instance the mentions as recorded within the raw ONS data are:

- Mention 1: Cerebral haemorrhage
- Mention 2: Metastasis of the brain
- Mention 3: Breast cancer
- Mention 4: Arterial hypertension
- Mention 5: Diabetes mellitus

Figure 6: Example of a death certificate

In this example, the primary cause of death appears in the middle of the record. This project places significance on the order of the mentions in determining the driver classification and it thus important the mentions are reordered.

Note that, in the raw data, the mentions are not split into parts 1 and 2 as done on the death certificate; this presents a challenge which is addressed by reordering the mentions. This is achieved by identifying the location of the ONS primary ICD code within the mentions and modifying the order of the mentions e.g.

- **Raw data:** Mention 1, Mention 2, Mention 3 (Primary Location), Mention 4, Mention 5
- **Reordered data:** Mention 3, Mention 2, Mention 1, Mention 4, Mention 5

Note that the mentions after the location of the primary ICD-code which relates to "other significant" mentions are not reordered. This exercise is achieved by identifying the first position of the primary ICD-code noting that there are few cases where:

- The primary ICD-code is not mentioned in the mentions e.g. where it has been determined during a coroner investigation. In this scenario the mentions are not reordered.
- The primary appears in multiple positions. In this scenario the reordering is based on the first position and subsequent mentions are not reordered.

6.4 Drivers of Mortality

The drivers of mortality have been assessed based on the underlying data. The breakdown with justifications is as follows:

- Cancer is broken into four categories in such a way that each cancer in that group will be impacted by longevity catalysts in a similar way.

- o Cancer – Unmet Need: In 2014, Cancer Research UK (CRUK) identified 4 cancers of unmet need[21] (lung, pancreatic, oesophageal and brain) which share poor five-year survival rates. These 4 cancers have been grouped together to reflect late diagnosis, poor survival rates and limited catalysts likely to improve death rates in the short term. Note that lung, pancreatic and oesophageal cancers are also strongly related to smoking. Unmet needs represents cancers with poor diagnosis, survival rates and limited improvements in treatment expected over the short term. Grouping is irrespective of link to smoking, obesity and other lifestyle risk factors.
  - o Cancer – Smoking Related: Aside from medical advancements (which are likely to impact cancer across the board), the main risk factors for cancer are smoking[22] and obesity[23]. The CRUK website lists the cancers that are linked to smoking and obesity and these lists have been used to create the related cancer groups. Where there is an overlap the Population attributable fractions (PAF)[24] score is used to determine whether to place it into smoking or obesity.
  - o Cancer – Obesity Related: As above, causes linked to obesity are identified based on a list from the CRUK website and the PAF score where there is an overlap.
  - o Cancer – Other: The remaining cancers are grouped together as 'other' with Prostate cancer making up over 21% of these cancers; other notable cancers in his group include Lymphoma (11%), Mesothelioma (6%), Melanoma (6%) and skin cancer (2%). Most deaths observed in this driver group thus result from highly treatable cancers.

- Respiratory Disease – Smoker Related: Covers respiratory causes of death with a strong link to smoking i.e. Chronic obstructive pulmonary disease (COPD), Bronchial Disease and Emphysema. These causes are clustered together based on their related factor i.e. smoking, and differentiated from other respiratory diseases like Asthma which are not smoker-related.

- "Alzheimer's Disease and Dementia" is differentiated based on its magnitude and significance.

- Diabetes and Obesity is differentiated based on its recent and potential future significance.

- Cardiovascular Disease: Due to similarities of stroke and coronary heart disease, causes under these are clustered together.

- Infection: Covers all infections such as influenza and HIV. It has been included to reduce the size of the "Other" group.

- COVID: Covers COVID-19 which is differentiated due to its significance and inflated levels in recent years.

- Non-Disease: Covers deaths due to accidents, suicide and other external causes. It is differentiated due to its distinctive influences i.e., not driven by disease. In addition, its shape by age is distinctive compared to disease-based drivers.

- Other: Captures all other causes including liver disease, neurological conditions, kidney damage, asthma to newborn conditions. The original list covered only the causes Parkinson's Disease, Liver Disease and Pancreatitis.

---

[21]Research opportunities in cancers of unmet need | Cancer Research UK
[22]How does smoking cause cancer? | Cancer Research UK
[23]How does obesity cause cancer? | Cancer Research UK
[24]The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015 — UK Health Security Agency (ukhsa.gov.uk)

- A co-morbidity driver group has been created which is defined as where a death has three or more <u>distinct</u> mentions in selected categories; see section 6.7.

It is important to note that the divisions of deaths into distinct cause and drivers is a matter of judgement and should change over time – so the drivers listed above should not be viewed as definitive.

6.5   Single Mention Mapping

The categorisation of 'cause of death' to drivers of mortality groupings, for single mentions, is presented in Table 5. For multiple mentions of 'cause of death', a different approach is required, and this is detailed in section 6.6.

Table 5 – Classification of 'Cause of Death' into a Driver Group

Driver	Cause of Death
Non – Disease	- Accident - Drug-Alcohol - External Causes - Injury - Medical Treatment Complications Natural Disaster - Suicide - Vehicular Accident
Cancer – Unmet need	- Lung cancer - Pancreatic cancer - Oesophageal cancer - Brain Tumours (including meningioma)
Cancer – Smoking Related	- Cancers of the mouth (including Pharynx & Larynx) - Nasal and paranasal sinus cancer - Bladder cancer - Cervical cancer - Acute myeloid leukaemia (AML) - Stomach cancer - Gynaecological Cancer
Cancer – Obesity Related	- Breast cancer - Colorectal cancer - Uterine cancer - Thyroid cancer - Gallbladder cancer - Myeloma - Kidney cancer - Liver cancer - Ovarian cancer - Bone Cancer
Cancer – Other	Captures all other cause of deaths linked to cancer
Respiratory Disease – smoker related	- Chronic obstructive pulmonary disease (COPD) - Bronchial Disease - Emphysema
Alzheimer's Disease and Dementia	"Alzheimer's Disease and Dementia"

Driver	Cause of Death
	• The ICD-code R54, which represents senility, frailty/old age is categorised with this driver
Diabetes and Obesity	• Diabetes • Obesity
Cardiovascular Disease	Coronary Heart Disease (CHD): • Coronary heart disease • Peripheral arterial disease • Rheumatic heart disease • AMI • Aneurysm • Atherolsclerosis • Cardiac Arrest • Cardiomyopathy • Heart Failure • Inflammation of the Heart • Ischaemic Heart Diseases • Other CHD • Phlebitit and Thrombophlebitis • Pulmonary Heart Disease • Valve Disorder Stroke: • Stroke • Transient Ischaemic Attack • Cerebral Infarction • Intracranial Haemorrhage • Other Cerebrovascular Disease • Subarachnoid Haemorrhage • Stroke Other
Infection	• Antibiotic Resistance • Infection • Influenza and Pneumonia
COVID	COVID-19
Other	Captures all other cause of deaths including: • Respiratory conditions • Parkinson's Disease • Liver disease • Pancreatitis
Co-morbidity	Not applicable to single mentions; see next sub section for definition

6.6 Multiple Mention Mapping

Figure 7 visualises the process taken to map the drivers for multiple mentions. The principles applied are:

- All ICD code mentions (up to 15 on each death) are used in the process. Deaths with a high number of mentions tend to have a high number of repetition of causes and are skewed towards accidents.

- A mention of COVID triggers classification into the COVID driver group.
- If the conditions for co-morbidity are met and COVID is not mentioned the death is classed as co-morbidity.
- Non-Disease is prioritised where a significant disease is not mentioned, or a lower significant disease is listed in the 3rd or later mentions.
- The order of the mentions, after reordering, is used in some cases to determine its classification e.g. respiratory (smoker) disease is used for deaths where it is mentioned before, say, "Alzheimer's Disease and Dementia".
- In determining the proposed driver classification for a death, distinctions are made between:
  - Events i.e. events that occur as a result of an underlying cause, e.g. pneumonia and influenza, including events that are typically mentioned but not a primary cause of death e.g. prostate cancer.
  - Causes i.e. what triggered the event of death e.g. cancer.
  - Table 6 shows a list of events and causes with justification.
- Certain causes are assumed to be more dominant e.g. a mention of cancer, excluding prostate or benign cancer, overshadows other mentions e.g. pneumonia.

Table 6 – Events and Causes with Justification for their Treatment

	Cause of Death	Justification
Events	Infection	A person with the causes listed in second half of this table and an infection/stroke is more likely to have died from the causal mention e.g. cancer
	Stroke	
	Prostate cancer	Predominately a disease of older men [25] and mentioned on a high proportion of such deaths. In the formulation of a driver, it is treated more as a disease that men die with rather than a cause.
	Benign cancer	By definition is not the cause of death in the presence of other mentions
	Medical treatment complication (MTC)	A death via MTC is an effect of an underlying cause for which intervention by the medical profession failed or accelerated the death
	Heart Attack	In the presence of other mentions (excluding benign cancer and MTC), these conditions are observed to be effects from a weakened system e.g. an infection leading to a heart attack
	Cardiac Arrest	
	Unknown	In the presence of accidents including drug and alcohol mentions, these causes of death are treated more as effects as for example the drug and alcohol consumption lead to the Liver disease or the AMI/PHD was triggered by shock from a car crash. An unknown cause is also excluded in determining accident deaths.
	Liver Disease	
	Acute myocardial infarction (AMI)	
	Pulmonary Heart disease (PHD)	

[25] Prevalence | Background information | Prostate cancer | CKS | NICE

	Cause of Death	Justification
Causes	COVID	COVID is flagged as a major cause of death in this dataset primarily due to the media focus on it and uncertainty surrounding it.
	Accidents	Accidents in the absence of other major causes are generally treated as the driver. Examples where they are not include: • "Alzheimer's Disease and Dementia" patient being hit by a car • Cancer patient committing suicide • COPD mentioned first with accident appearing in lower mentions
	Cancers	Deaths from these causes are generally observed to be the driver with prioritisation given to say cancer over "Alzheimer's Disease and Dementia" as laid out in Figure 7.
	Alzheimer's Disease and Dementia	
	Respiratory Disease – smoker related	
	Diabetes and Obesity	
	Coronary Heart Disease	

Figure 7 – Mapping of Driver's Decision Tree

Key

| Decision | | Final Driver |

MTC= Medical Treatment Complication
BN= Benign Neoplasm
PC= Prostate Cancer
CHD= Coronary Heart Disease
CA= Cardiac Arrest
HF= Heart Failure
PHD= Pulmonary Heart Disease

Co-morbidity assigned where there are 3 distinct mentions in the categories:
-Cancer
-Diabetes and Obesity
-Dementia excluding Frailty
-CHD excl. HF/CA
-Respiratory Disease/Condition (all)
-Other excl. neurological, new born & pregnancy
-Neurological conditions
-Infection excl. COVID/Influenza

Excluding "Other" which is counted at a causal level, multiple mentions within a category are counted as one

6.7 Co-Morbidity

Figure 8 details the disease groups that contribute to co-morbidity. In this diagram:

- Multiple mentions of drivers/causes in a green container are counted as one.
- Multiple mentions of drivers/causes in a blue container are counted separately.
- Drivers/causes shaded grey do not count as a co-morbidity.

A few examples of co-morbidity are shown in Table 7. The classification of this driver group was discussed and approved by the Working Party.

Figure 8 – Drivers/Causes that Contribute to the Co-morbidity Driver
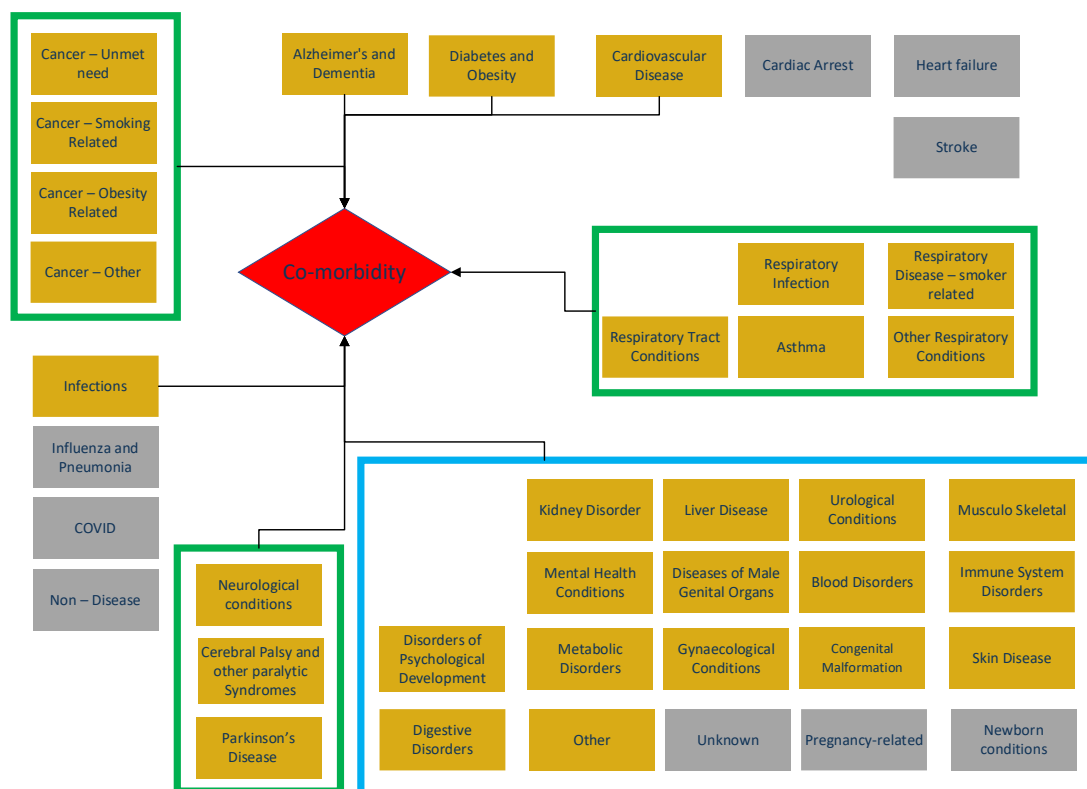
Table 7 – Examples of Co-morbidity Mentions

Co- morbidity	Not a Co-morbidity	Justification
Diabetes and Obesity, Bone cancer, "Alzheimer's Disease and Dementia"	Diabetes and Obesity, Bone cancer, Lung cancer	Multiple mentions of cancer are counted as one
Kidney Disorder, Liver Disease, Respiratory Infection	Kidney Disorder, Liver Disease, Influenza and Pneumonia	Influenza and Pneumonia is excluded from causes that count towards Co-morbidity
Tuberculosis, Ovarian Cancer, Parkinson's Disease	Tuberculosis, Neurological Conditions, Parkinson's Disease	Multiple mentions of neurological-related diseases are counted as one
"Alzheimer's Disease and Dementia", Colo-Rectal Cancer, Other CHD	"Alzheimer's Disease and Dementia", Colo-Rectal Cancer	Co-morbidity requires three or more mentions of certain diseases

Co- morbidity	Not a Co-morbidity	Justification
Metabolic Disorders, Bronchial Disease, Musculo Skeletal	Metabolic Disorders, Bronchial Disease, COPD	Multiple mentions of respiratory diseases are counted as one
Valve Disorder, Diabetes and Obesity, Bone cancer, Cardiomyopathy	Valve Disorder, Diabetes and Obesity, Cardiomyopathy	Cardiovascular diseases are counted as one

6.8 Alternatives

Alternative versions of the dataset have been created which:

- Include/ exclude the COVID and Co-morbidity drivers.
- Utilise the ONS Primary cause of death to identify the driver.

These versions of the dataset are included in the publication to help users understand the impact of judgements made and to take a different decision if desired. See section 7 for more details.

6.9 Judgements

Table 8 provides a list of judgements with rationale. This is added to aid users of the data understand judgements that have been made. No assessment has been conducted on the materiality of these judgments however some (e.g. introduction of co-morbidity), can be assessed by users using the alternative datasets.

Table 8 – Judgements

#	Judgement
1	Co-Morbidity is defined as where a death has three or more distinct mentions in the categories: • A cancer mention • "Alzheimer's Disease and Dementia" (excluding frailty) • Diabetes and Obesity • Cardiovascular Disease excluding stroke, heart failure and cardiac arrest • All respiratory diseases/ conditions • Other (excluding pregnancy-related, new-born cause and all neurological conditions) counted at a casual level • All neurological conditions • Infection excluding COVID/ influenza and Pneumonia • Non-Disease is the only driver group excluded entirely. Excluding "Other" which is counted at a causal level, multiple mentions within a category are counted as one.
2	Non – Disease (excluding medical treatment complication) is prioritised over other mentions excluding: • Cancer or "Alzheimer's Disease and Dementia" • A 1st or 2nd mention of either Other, Coronary heart disease, respiratory disease (smoker-related) excluding Liver disease, cardiac arrest, heart failure, pulmonary heart disease and AMI. The exclusions are applied because, for example, it is more likely that "Alzheimer's Disease and Dementia" (cause) lead to an accident (effect).
3	Due to recent attention on COVID, it is prioritised where mentioned for 2020 onwards. Note that COVID can be viewed as an effect as it tends to be mentioned with more

#	Judgement
	dominant disease e.g., cancer and a scenario is created where it is treated like seasonal flu.
4	Deaths with multiple significant mentions are mapped a Co-morbidity driver. Mapping of deaths with multiple significant mentions is highly subjective and a decision was taken to distinguish them from other deaths where the mapping is not straight forward.
5	Cancer excluding: • Prostate cancer which is mentioned in a high proportion of male deaths. • Benign cancer which by definition is not a cause of death. is prioritised over other mentions, and is ranked: • Unmet needs i.e., cancers with poor survival rates and limited improvements in treatment • Smoker-related defined based on causal associations • Obesity-related • Other
6	After cancer, respiratory disease – smoker related is mapped as the driver excluding where "Alzheimer's Disease and Dementia" is mentioned earlier.
7	The leading cause of death[26] is mapped as the "Alzheimer's Disease and Dementia" driver if mentioned.
8	A decision has been taken to recognise Diabetes and Obesity as a leading driver and it is mapped after "Alzheimer's Disease and Dementia" cases.
9	The order of mentions, after reordering, plays a vital role in the categorisation applied in determining if a death is related to Cardiovascular disease or Other. Where coronary heart disease CHD* is the first mentioned or the second mentioned with heart failure/cardiac arrest mentioned first, then Cardiovascular disease is prioritised. Where respiratory conditions is mentioned with an infection or "Other" is mentioned first the death is mapped as being driven by "Other". CHD* is then used if it's in the second mention, followed by Other in the second mention. For mentions 3 above CHD* is used followed by a mention of Other. E.g. • M1= Other M2= CHD then driver is Other • M1= CHD M2= Other then driver is Cardiovascular Disease • M1= Infection M2= Other M3=CHD then driver is Other • M1= Stroke M2= Infection M3= Other M4= CHD then driver is Cardiovascular Disease *Excluding heart failure and cardiac arrest*
10	Prostate cancer is mapped if it's the first mention. More weight is given to it based on the order of the mentions and lower significance of other causes that remain.
11	Infection excluding influenza and COVID is mapped if it's the first mention. More weight is given to it based on the order of the mentions and lower significance of other causes that remain.
12	Stroke, which is viewed as more of an effect than a cause is mapped at this stage if mentioned due to lower significance of other causes that remain.
13	A mention of prostate cancer is prioritised next regardless of position due to lower significance of other causes that remain.
14	A mention of Infection excluding influenza and COVID is prioritised next regardless of position due to lower significance of other causes that remain.

[26] Leading causes of death, UK - Office for National Statistics (ons.gov.uk)

#	Judgement
15	The ICD-code "R540", described as senility/frailty is mapped to "Alzheimer's Disease and Dementia". It appears on a large proportion of deaths at older ages and a judgement has been made to lump it with "Alzheimer's Disease and Dementia".
16	The following causes are then mapped in sequential order: • All infection mentions are mapped at this stage. • All remaining CHD i.e., heart failure and cardiac arrest are mapped next • Next, Benign cancer is mapped and then • Medical treatment complications These causes are lower in the process because they are considered to be more of an effect than a cause.
17	Where no mentions exist the process defaults to the ONS primary cause (if available) otherwise "Other". Note that the ONS primary cause, created from the mentions, is ignored otherwise.
18	Unknown/Missing ICD-10 codes linked to new-born deaths, are defaulted to either the ONS Primary ICD Code or "Other" if the primary is also blank.
19	As discussed in section 6.3, the primary ICD-code is used to reorder the mentions based on its first position in the mentions. The mentions in the ONS mortality data are ordered as recorded on a death certificate. The ordering suggests the primary cause of death appears in the middle of the record. This project places significance on the order of the mentions in determining the driver classification and it thus important the mentions are reordered.

7   IFoA Drivers of Mortality Dataset

7.1   Introduction

This section covers the:

- Specification of the dataset being published.
- Banding of Age

7.2   Specification of Data

The variables included in the data ("IFoA Drivers of Mortality Dataset") being released:

- Primary Mapping i.e. the data being proposed to be used to assess driver trends by year of death, sex and age. In addition to the primary mapping dataset, four other alternatives are being released:
  - Alternative Mapping (No COVID/No Co-Morbidity): An alternative version of the primary mapping with COVID treated similar to seasonal flu and the Co-Morbidity driver, which captures deaths with multiple significant causes, removed.
  - Alternative Mapping (No COVID): An alternative version of the primary mapping with COVID treated similar to seasonal flu.
  - Alternative Mapping (No Co-Morbidity): An alternative version of the primary mapping with the Co-Morbidity driver removed.
  - ONS Primary ICD Code: an alternative mapping based on only the ONS primary ICD code.
- Age (Last birthday) is provided at an integer level, where the death count is greater than 2 for each sex, year and driver combination. Below this, the ages are banded to derive clusters with more than 2 deaths. The banding of age is described in more detail in section 7.3.
- Sex is banded M (male) or F (female). Unknown/Missing variables are defaulted to the Male sex, consistent with the approach adopted by the ONS in its published datasets.
- Calendar Year i.e. the year death occurred. An alternative dataset has been created using year of death registration to demonstrate consistency with ONS publications, and the data used by the CMI in its projection model; it is discussed further in section 9.6
- Driver Of Mortality: 13 driver groupings discussed in section 6.4.

7.3   Banding of Age

To meet SRS rules (i.e. each cluster must have more than 2 deaths) and to extract as much information as possible so the data can be used for multiple purposes, age is banded as follows:

- For ages 70+:
  - Locate clusters (by dataset, age, year, sex and driver) where death counts are below 3.
  - Find the minimum age where deaths fall below 3 which is the point where banding starts.
  - Reduce the minimum age to the point where the sum of deaths is greater than 2. This step is required due to gaps between the minimum age and the next age.
  - Aggregate the data to remove duplicates.
- For ages under 70
  - Locate clusters where death counts are below 3.

- o Find the maximum age where deaths fall below 3 which is the point where banding starts.
- o Shift the maximum age up so it's in the format X4 or X9 e.g. 43 becomes 44, 35 becomes 39. This is done to simplify the process and reduce the number of bands generated.
- o Aggregate the data to remove duplicates.

- Application of age bands to data:
  - o Join the above queries with the data to assign bands which are driver specific; see section Appendix E – Age Banding.
  - o Check the death count by the variables to ensure deaths have not been added/removed.

- The process allocates deaths at younger ages, that are below 3 after application of the X4/X9 bandings, into a band called "LowerBand" and in addition utilises bands that could be misinterpreted e.g. band 40-49 does not cover ages 40-49 but rather covers from where deaths fall below 3 within this age band. The final step is thus to clean up the labelling to avoid any misinterpretation. A further check is performed to ensure the death count is unchanged by this step.

8 Analysis of Dataset

8.1 Introduction

This section provides an assessment of the patterns observed in the dataset and covers the following elements:

- Analysis by Key Factors
- Impact of treating COVID as a seasonal flu
- Explores the co-morbidity driver.

8.2 Analysis by Key Factors

Analysis has been carried out assessing the resultant data set by the following factors:

- Sex
- Calendar Year
- Age

8.2.1 Sex

Table 9 shows the deaths (count and proportions) by sex and drivers. The following observations are noted:

- Cardiovascular disease is the leading driver of mortality for both sexes for deaths reported in the study period (2001-2021); there has been a decline in deaths associated with this driver in more recent years. Post 2014, Co-morbidity is the leading driver.
- A quarter of all deaths are linked to Cancer with females dying more due to Cancer – Obesity Related and males Cancer – Unmet need.
- Drivers like "Alzheimer's Disease and Dementia" and Cancer – Obesity Related are heavily weighted towards females due to the age distribution at which the disease peaks with females living longer and diseases like breast cancer included in the classification, respectively.
- Drivers like non-disease, which covers accidents and suicides, and Cancer – Other (which includes prostate cancer) are weighted towards males.

Table 9 – Deaths by Sex and Drivers

Driver	Deaths			% of column total		
	Female	Male	ALL	Female	Male	ALL
Alzheimer's Disease and Dementia	808,474	348,691	1,157,165	14.3%	6.5%	10.5%
Cancer – Obesity Related	576,270	272,451	848,721	10.2%	5.1%	7.7%
Cancer – Other	250,354	420,728	671,082	4.4%	7.9%	6.1%
Cancer – Smoking Related	128,942	209,485	338,427	2.3%	3.9%	3.1%
Cancer – Unmet need	406,960	536,247	943,207	7.2%	10.1%	8.6%
Cardiovascular Disease	1,137,463	1,133,988	2,271,451	20.1%	21.3%	20.7%
Co – Morbidity	815,901	849,225	1,665,126	14.5%	15.9%	15.2%
Covid-19	71,988	87,559	159,547	1.3%	1.6%	1.5%
Diabetes and Obesity	129,739	146,219	275,958	2.3%	2.7%	2.5%
Infection	167,957	111,191	279,148	3.0%	2.1%	2.5%
Non – Disease	129,964	234,364	364,328	2.3%	4.4%	3.3%

Driver	Deaths			% of column total		
	Female	**Male**	**ALL**	**Female**	**Male**	**ALL**
Other	743,162	654,806	1,397,968	13.2%	12.3%	12.7%
Respiratory Disease – smoker related	272,800	320,038	592,838	4.8%	6.0%	5.4%
Pre-2001 deaths	6,344	9,823	16,167	0.1%	0.2%	0.1%
Total	5,646,318	5,334,815	10,981,133	100.0%	100.0%	100.0%

8.2.2 Calendar Year

Figure 9 shows deaths by calendar year and drivers. Note that the trends are heavily influenced by increasing age. Figure 10, Figure 11 and Figure 12 show the same chart for age bands <60, 60-79 and 80+ respectively. The following observations are noted:

- Co-morbidity which is defined based on a death having 3 distinct significant mentions trends upwards, for older age groups, and is the leading driver of mortality in more recent years. It is possible that part of the trend is due to changes in recording of causes on medical certificates over time.

- Deaths aged under 60 are driven primarily by non-disease and other, both of which are on a downward trend with non-disease heavily influenced by late reporting; see section 9.6.

- The decline in cardiovascular related deaths is more visible at older ages. For ages 80+, the proportion of deaths driven by "Alzheimer's Disease and Dementia" increases significantly.

Figure 9 – Deaths by Calendar Year and Drivers
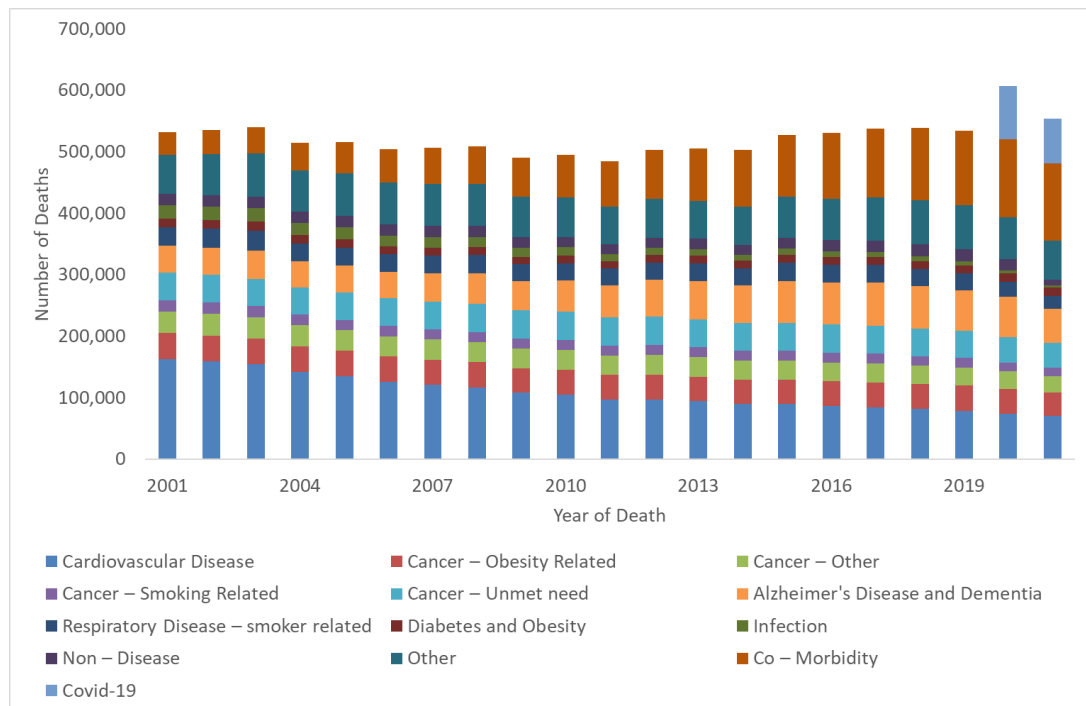
Figure 10 – Deaths by Calendar Year and Drivers (Under age 60)

Legend:
- Cardiovascular Disease
- Cancer – Obesity Related
- Cancer – Other
- Cancer – Smoking Related
- Cancer – Unmet need
- Alzheimer's Disease and Dementia
- Respiratory Disease – smoker related
- Diabetes and Obesity
- Infection
- Non – Disease
- Other
- Co – Morbidity
- Covid-19
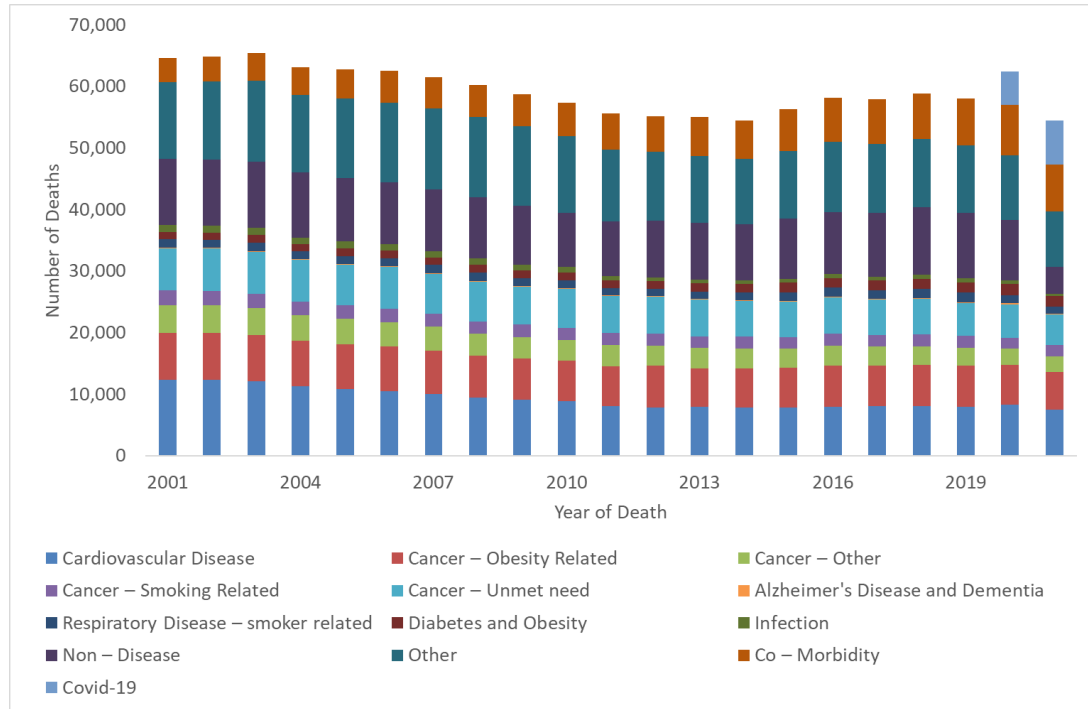
Figure 11 – Deaths by Calendar Year and Drivers (60-79)

Legend:
- Cardiovascular Disease
- Cancer – Obesity Related
- Cancer – Other
- Cancer – Smoking Related
- Cancer – Unmet need
- Alzheimer's Disease and Dementia
- Respiratory Disease – smoker related
- Diabetes and Obesity
- Infection
- Non – Disease
- Other
- Co – Morbidity
- Covid-19
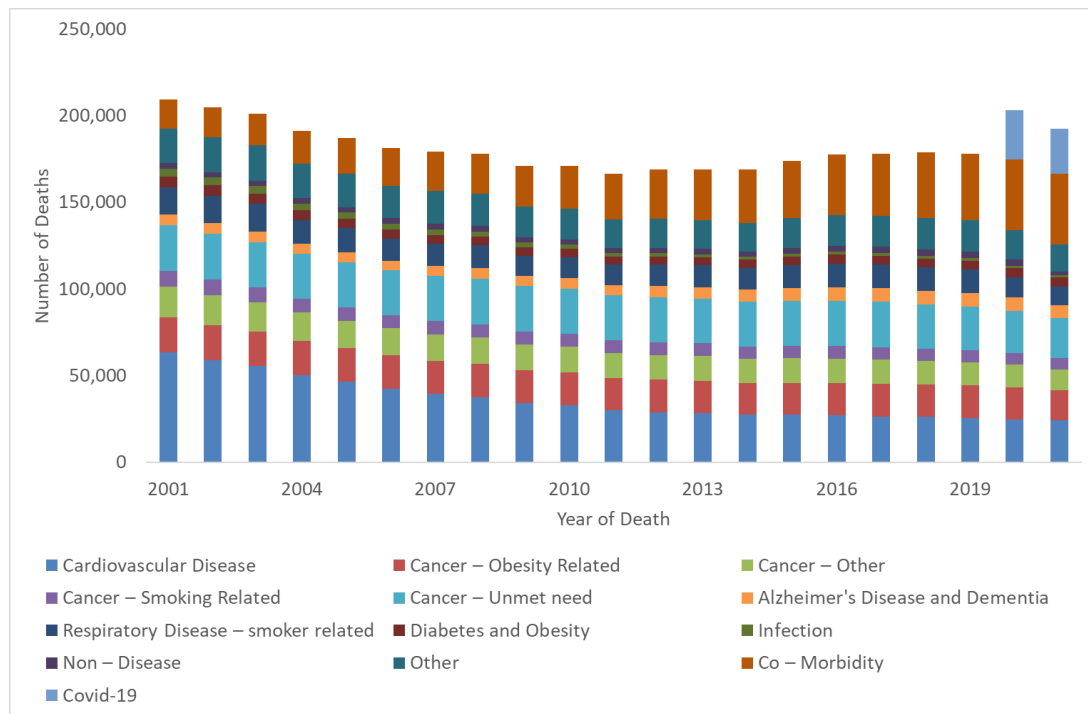
Figure 12 – Deaths by Calendar Year and Drivers (80+)
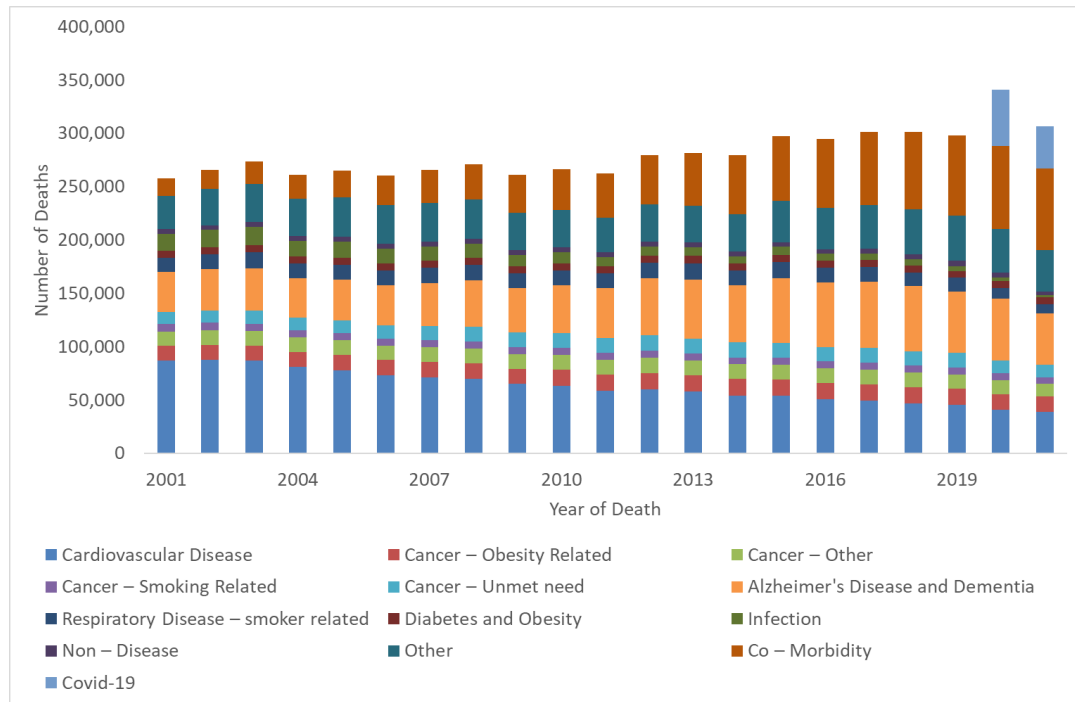
8.2.3   Age

Figure 13 shows deaths by calendar year and drivers. The following observations are noted:

- At younger ages (<20) deaths are driven by Other which includes new-born diseases and the non-disease driver.
- Above age 92, "Alzheimer's Disease and Dementia" surpasses Cardiovascular Disease as the leading cause of death.
- Over 50% of deaths occur between ages 80-99 due to ageing and increased susceptibility to disease.
- The peak of deaths varies by driver e.g., Cancer Unmet Needs peaks around age 77 whilst "Alzheimer's Disease and Dementia" peaks at age 90.

The patterns for older individual years (2001-06) reveal a bimodal peak by age however, combined, this pattern is not observed. The trough coincides with World War One era births and is observed in other datasets e.g., population estimates. It is a result of an uneven distribution of births due to World War One. This issue is not observed in the individual years from 2007 onwards. Note that this issue is present in the ONS Primary ICD code and is thus not a by-product of the proposed mapping approach.

Figure 13 – Deaths by Age and Drivers
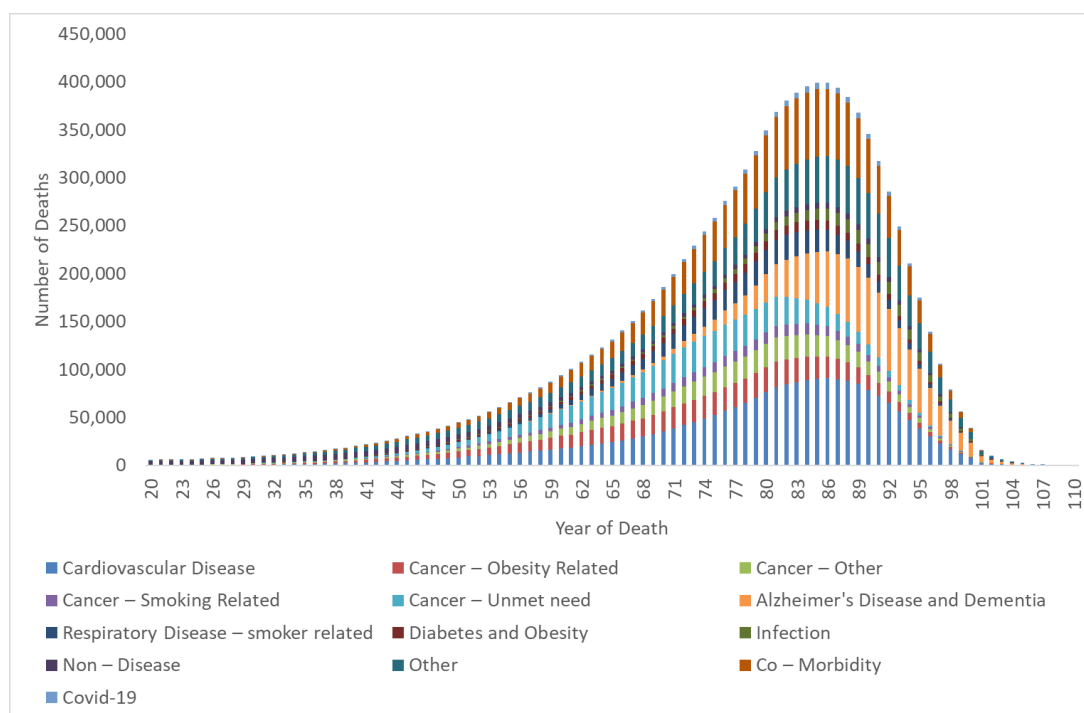
8.3   COVID as Seasonal Flu

Table 10 shows deaths in the period 2020-21 by driver for the proposed mapping and with COVID treated as a seasonal flu i.e. categorised with Infection as opposed to having its own driver group. As discussed in section 7, an alternative dataset has been produced without the COVID driver.

Infection-related deaths were most affected by COVID, followed by Diabetes, Respiratory and Dementia-related deaths.

Table 10 – 2020-21 Deaths Driven by Driver Showing Impact of COVID

Driver	Proposed Mapping	COVID as a seasonal flu	% Change
Alzheimer's Disease and Dementia	120,838	145,347	16.9%
Cancer – Obesity Related	77,028	80,386	4.2%
Cancer – Other	55,450	58,862	5.8%
Cancer – Smoking Related	28,805	30,836	6.6%
Cancer – Unmet need	82,175	85,284	3.6%
Cardiovascular Disease	144,779	162,961	11.2%
Co – Morbidity	252,185	291,365	13.4%
Covid-19	159,547	n/a	n/a
Diabetes and Obesity	27,417	39,826	31.2%
Infection	8,521	23,481	63.7%
Non – Disease	27,684	29,553	6.3%
Other	131,514	157,692	16.6%
Respiratory Disease – smoker related	44,131	54,481	19.0%

Driver	Proposed Mapping	COVID as a seasonal flu	% Change
Total	1,160,074	1,160,074	0%

8.4   Co-Morbidity

The Proposed Mapping approach attributes close to two million deaths to the co-morbidity driver; 15% of all deaths. As discussed in section 7, an alternative dataset has been produced without the co-morbidity driver. Figure 14 shows co-morbidity deaths by calendar year for the Proposed Mapping and with COVID treated as seasonal flu. The following observations are noted:

- Co-morbidity trends upwards and is the driver with the highest number of deaths from 2014 onwards; over 20% of deaths post 2013 are classed under this driver.
- Adding deaths which would have been classified as Co-morbidity, if COVID was treated like influenza, results in a spike of 13% in 2020-21.

Figure 14 – Co-morbidity Deaths by Calendar Year
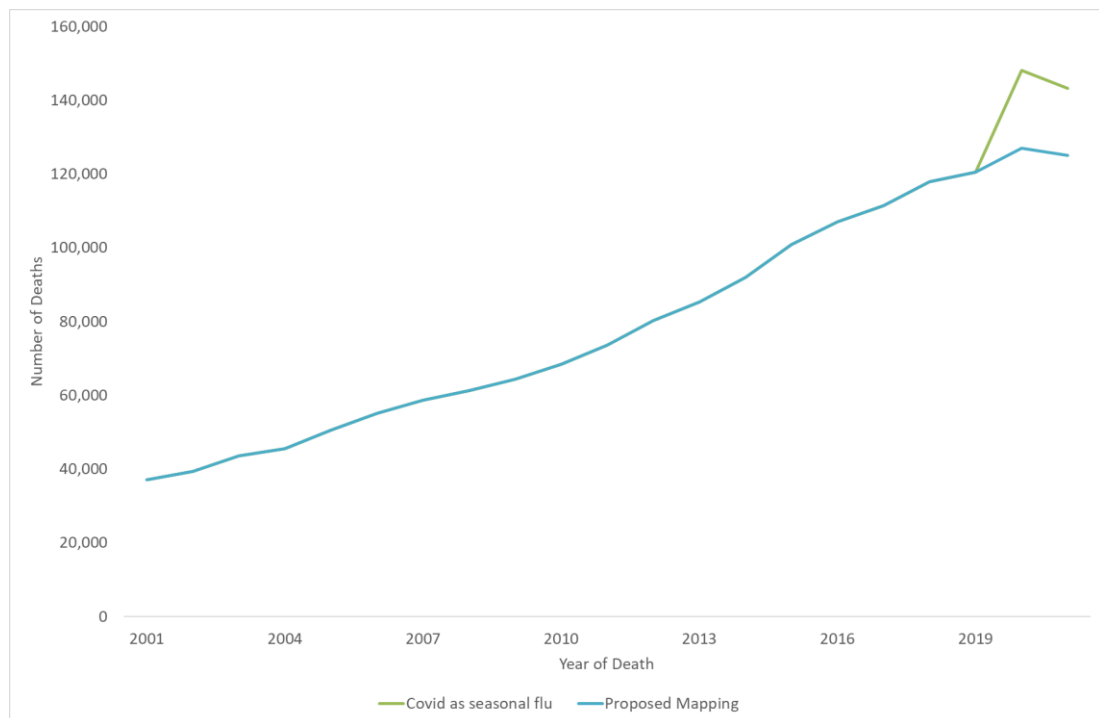
Table 11 places co-morbidity deaths in the driver they would have been assigned to if it was not created. The largest contributor to Co-morbidity is the Other driver with approximately:

- 40% being classed based on mentions with cardiovascular disease and other
- 60% being classed based on multiple mentions of Other e.g., a death with respiratory condition, liver disease and digestive disorder.

The trend in co-morbidity could be artificial in that it is due to an increase in the number of mentions being registered by medical practitioners over time or a result of improvement in diagnosing diseases.

Table 11 – Co-morbidity Deaths Split by Main Driver Group

Driver Group	Deaths	Percentage
Alzheimer's Disease and Dementia	294,886	17.7%
Cancer – Obesity Related	141,676	8.5%
Cancer – Other	116,156	7.0%
Cancer – Smoking Related	74,879	4.5%
Cancer – Unmet need	111,557	6.7%
Cardiovascular Disease	127,342	7.6%
Diabetes and Obesity	256,497	15.4%
Non – Disease	17,395	1.0%
Other	333,416	20.0%
Respiratory Disease – smoker related	191,322	11.5%
Total	1,665,126	100.0%

9 Validation

9.1 Introduction

This section details the validation carried out to ensure the accuracy of the data produced and covers the following elements:

- Data checks
- Process checks
- Independent review of data processing
- Independent review of the mapping
- Assessment of the Year of death occurrence vs. reported
- Compares the proposed mapping approach to a driver constructed using the ONS primary.

9.2 Data Checks

A number of checks were conducted by Nana Asiamah (NA) within the SRS network to ensure:

- Age is in the range 0-120
- Year of death is populated
- Sex is either M or F
- Primary ICD- Code is populated

on the raw datasets. These checks are detailed in this section and there are no issues of concern identified.

A summary of error count and as a proportion of deaths in the study period (2001-2021), is shown in Table 12. Note that the age errors have other variables missing and its count has not been repeated in other rows to avoid inflating the numbers. Figure 16 shows the errors spread across the study period. Excluding errors relating to age in 2001, none of these errors are filtered out of the data.

Table 12 – Errors in Raw Data

Field	Count	% Deaths
Age	5,379	0.05%
Year	0	0.00%
Sex	138	0.00%
ICD Unknown	39,486	0.38%

Figure 15 – Error Rate by Year for age, sex, and Primary ICD Code
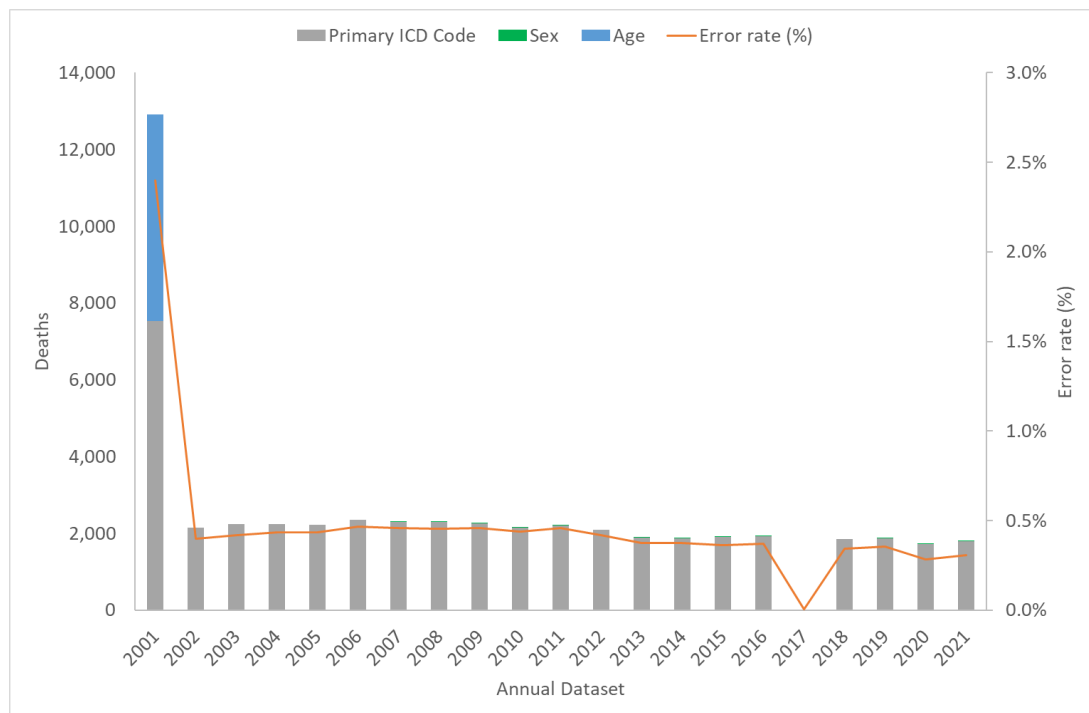
9.2.1 Age

Age (Last birthday) is provided at an integer level and is expected to be in the range 0-120. Nulls, suggested by the ONS to be newborn deaths, were observed for some records in the 2001 dataset; 5,739 deaths (1.00% of this particular file or 0.05% of the entire dataset). These records also have variables like sex and death date missing and have been removed from the data. Note that these records do not appear in the 21st Century Mortality Data which suggests their removal is consistent with how the ONS process the data. Figure 16 shows the deaths (under age 20) across the study period, the removed deaths (red dotted lines) do not fit the historical averages.

Figure 16 - E&W Deaths under age 20
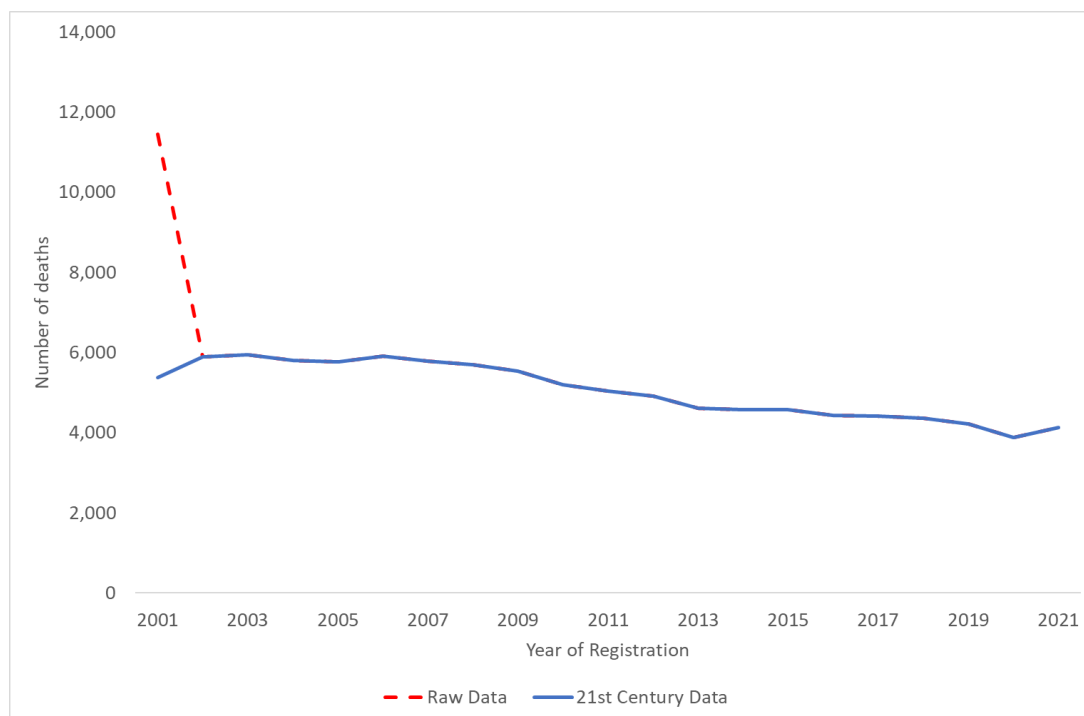
9.2.2 Year of Death

Year of death occurrence is deemed more suitable for the modelling of drivers. Excluding 2001, which has 5,379 deaths with no death date, no issues were identified with this variable. However, due to reporting delays, this paper recommends the 2021 data is discarded or an adjustment is made to make it more useable.

9.2.3 Sex

Sex is banded M (male) or F (female). Unknown/Missing variables are defaulted to the Male sex as done by the ONS in its published dataset. Excluding 2001, a total of 138 deaths (0.00%) are affected by this issue over the study period. Table 13 covers deaths affected by unknown/missing sex issues along with missing entries for ICD Code which is discussed in the next sub-section.

Table 13 – Deaths affected by Missing Data Items

Year	Sex	No Primary ICD Code	No Mentions (excl. Primary)
2001	5,379	7,527	7,529
2002	0	2,139	2,141
2003	0	2,248	2,279
2004	0	2,239	2,239
2005	0	2,220	2,275
2006	0	2,345	2,347
2007	7	2,294	2,295
2008	6	2,299	2,311
2009	4	2,252	2,256
2010	5	2,148	2,180

Year	Sex	No Primary ICD Code	No Mentions (excl. Primary)
2011	10	2,203	2,194
2012	0	2,098	2,111
2013	11	1,897	1,940
2014	8	1,870	1,818
2015	8	1,903	1,888
2016	21	1,929	1,963
2017	16	0	1,940
2018	0	1,851	1,880
2019	12	1,868	1,850
2020	14	1,719	1,665
2021	16	1,789	1,594

9.2.4   ICD- Code

All ICD-10 code mentions including the primary, up to 16 codes on each death, is used to create the data. There are deaths with no mentions including primary which has been suggested to be linked to newborns by the ONS. These deaths have an age of 0 assigned to them and have been left in the data.

For zero mention (excluding primary), the cause of death is determined using the ONS Primary ICD Code or "Other" if the primary is also blank. There are:

- 41,337 deaths (0.38%) with missing primary ICD-code excluding the 5,379 records filtered out.
- 48,772 deaths (0.44%) with no mentions ignoring the primary excluding the 5,379 records filtered out.

Table 13 shows deaths, including the 5,379 filtered deaths, with no primary ICD code and no mentions across the study period.

9.3   Process Checks

This section describes the end-to-end process with the goal of demonstrating the checking performed by Nana Asiamah (NA) within the SRS network. The end-to-end process is as follows:

- Each of the individual years mortality data (2001-2021) are cleansed, formatted and appended with the IFoA Cause Of Death classification in a separate database. The checks on age, sex, year and ICD-Code are performed at this stage.
- Combined dataset: All the annual datasets are combined into a single file and the sum of deaths is checked against the individual datasets. The combined dataset, described above, is validated against the 21st century mortality dataset[27] and produces consistent results by year of registration and age band excluding deaths in the 2001 registration file.
- Aggregated dataset: Due to access database memory limitations, the dataset is compressed into a smaller size with variables:
  - o Age

---

[27]Deaths registered in England and Wales – 21st century mortality - Office for National Statistics (ons.gov.uk)

- o Sex
- o Year of death/registration
- o Granular drivers (see section 14.3) associated with the primary ICD code
- o Granular drivers associated with the first mention
- o Granular drivers associated with second mention
- o Flags that identify the granular drivers on each of the clustered deaths

and the drivers mapped to this data based on the rules described in section 6.6. The rules used to map the drivers to the data has been independently reviewed; see section9.4. The total death count remains the same throughout the aggregation process.

The final data is then downloaded for export with some further aggregation to meet SRS rules (i.e. each cluster must have more than 2 deaths) and to extract as much information as possible so the data can be used for multiple purposes; see section 7.3.

9.4   Independent review of the data processing

Due to constraints around accessing the data within the ONS network an independent review, by Sandesh Maharaj (SM), has been performed outside the network by extracting all codes used in the process. A test dataset covering 3,000 deaths over the period 2016-18, each period covering a different data input format, has been used to verify consistency between the two processes. An exact match was produced by the two processes. In addition, the mapping approach detailed in section 6.6 has been independently replicated by SM.

An independent review of the cleanup process of the age bands including spreading of the deaths where banded is performed by Ben Rees (BR); see Appendix F – Redistribution of grouped data into single-year age bands.

9.5   Independent review of the mapping

The mapping approach, which is formulaic, has been manually reviewed from bottom-up by Sacha Dhamani (SD) and meets his expectations with few discrepancies. This review utilises a dataset grouped by mentions with clusters restricted to having deaths greater than 2 (88% of all deaths in the study period) as per SRS data release restrictions. Table 14 summaries the review results by mentions, percentage reviewed and with discrepancies. There are approximately 0.5% of the total number deaths where an alternative view could be taken.

Table 14 – Summary of Manual Review

Mention	Deaths	% Reviewed	% Differences
0	43,383	100%	0.0%
1	2,556,357	100%	0.0%
2	3,367,547	86%	0.0%
3	2,298,560	43%	1.6%
4	984,675	36%	1.2%
5	216,265	37%	0.3%
6	28,021	39%	0.7%
7	1,877	59%	1.7%
8+	1,606	100%	0.0%
Total	9,498,291	73.0%	0.5%

Examples of the discrepancies or disagreements are listed in Table 15.

The data and process used to perform this exercise is separate from the actual process and serves as an independent check. A test dataset with 65,586 deaths run through the two processes found 32 differences (0.05%). These differences are due to the treatment of deaths flagged as "Unknown" in the ONS primary and/or the first mention with subsequent mentions being accident or injuries. The actual process maps these cases to "Other" whilst the replicated approach maps these to "Non-Disease". This is noted as a difference with low materiality.

Table 15 – Examples of Differences Identified)

Final Driver	M1	M2	M3	M4	M5	SD Recommendation
Co – Morbidity	Alzheimer's Disease and Dementia	Urological Conditions	Infection			Alzheimer's Disease and Dementia
Co – Morbidity	Urological Conditions	Infection	Kidney Disorder			Other
Other	Urological Conditions	Infection	Senility, Frailty			Alzheimer's Disease and Dementia
Co – Morbidity	Cancer of trachea, bronchus and lung	COPD	Other CHD			Cancer - Unmet need
Co – Morbidity	Leukaemia	Infection	Blood Disorders	Other		Cancer- Other
Non – Disease	Accident	R540	Injury	Influenza and Pneumonia		Alzheimer's Disease and Dementia
Respiratory Disease – smoker related	Accident	COPD	Injury	Influenza and Pneumonia		Non – Disease
Cardiovascular Disease	Influenza and Pneumonia	Heart Failure	Kidney Disorder	Other CHD		Other
Non – Disease	Influenza and Pneumonia	Heart Failure	Other CHD	Injury	Medical Treatment Complications	Cardiovascular Disease

The differences are due to:

- The definition of co-morbidity which is formulaic and could be further refined.
- Different interpretations of what significance could be placed on a senility/frailty mention. This is currently mapped to "Alzheimer's Disease and Dementia" after other significant mentions have been used.
- Non-disease classification has some exclusions that are debateable e.g. was the accident that lead to a death driven by person having COPD or cancer?
- Where a heart condition is listed first or is the ONS primary cause of death, should other significant mentions like "Kidney Disorder" be ignored?

These are noted as a difference in opinion with low materiality.

9.6   Year of Death Occurrence vs Reported

This section assesses the difference between year of occurrence and year of reporting using similar dataset created using the year a death was reported. Table 16 shows the year of death occurrence and year of death reporting by calendar year and Table 17 shows the 2018 deaths by driver group for the two definitions of year. The following observations are noted:

- The proposed mapping uses the year of death occurrence which leads to a loss of 16,167 deaths which occurred pre-2001.
- Note that deaths reported in say 2021 cover deaths that occurred in that year including deaths reported in that year that occurred in previous years. For this reason, the deaths registered in the most recent year will be greater than the deaths that actually occurred. The year a death occurred in each registration period can be assessed in the ONS dataset Impact of registration delays on mortality statistics[28].
- The difference in yearly deaths is within 1% excluding 2021 where a difference of 5.6% is noted due to the impact of late of reporting of deaths; the ONS and CMI favour the year of death reporting in their annual refresh of the data for this reason.
- In addition, the CMI[29] note that there has been an unusual patterns of reporting delays in 2022 and 2023 to date which could lead to artificial trends and is something they are mindful of when considering the method and parameters for CMI_2023.
- The impact is more notable for the "Non-Disease" driver group. This due to the requirement for such deaths (e.g. suicide) to go through a coroner inquest. The differences for this driver group is significant in recent years 2016 (-2.4%), 2017 (-5.2%), 2018 (-2.6%), 2019 (0.1%), 2020 (3.1%) and 2021 (48.1%). The patterns suggest using the year of registration would lead to a misallocation of deaths by driver.

The year of death occurrence is thus more suitable/reliable for the modelling of drivers however its use requires the most recent year to be adjusted or discarded.

Table 16 – Year of Death Occurrence vs Reported

Year	Year of Death Occurrence	Year of Death Reporting	% Diff
<2001	16,167		
2001	531,745	532,498	0.1%
2002	535,319	535,356	0.0%
2003	540,282	539,151	-0.2%
2004	514,780	514,250	-0.1%
2005	515,115	512,993	-0.4%
2006	504,354	502,599	-0.3%
2007	505,956	504,052	-0.4%
2008	509,087	509,090	0.0%
2009	490,592	491,348	0.2%
2010	494,512	493,242	-0.3%
2011	484,299	484,367	0.0%

---

[28]Impact of registration delays on mortality statistics - Office for National Statistics (ons.gov.uk)
[29]Mortality monitor Week 2023-39 v01 2023-10-10.pdf (actuaries.org.uk)

Year	Year of Death Occurrence	Year of Death Reporting	% Diff
2012	503,099	499,331	-0.8%
2013	505,663	506,790	0.2%
2014	502,672	501,424	-0.2%
2015	527,630	529,655	0.4%
2016	530,400	525,048	-1.0%
2017	537,169	533,253	-0.7%
2018	538,600	541,589	0.6%
2019	533,618	530,841	-0.5%
2020	606,333	607,922	0.3%
2021	553,741	586,334	5.6%
Total	10,981,133	10,981,133	0.0%

Table 17 – Year of Death Occurrence vs Reported in 2018 for each Driver

Year	Year of Death Occurrence	Year of Death Reporting	% Diff
Alzheimer's Disease and Dementia	69,520	70,229	1.0%
Cancer – Obesity Related	40,561	40,580	0.0%
Cancer – Other	30,190	30,280	0.3%
Cancer – Smoking Related	15,349	15,411	0.4%
Cancer – Unmet need	44,288	44,450	0.4%
Cardiovascular Disease	81,169	81,768	0.7%
Co – Morbidity	117,948	118,900	0.8%
Covid-19	n/a	n/a	n/a
Diabetes and Obesity	13,080	13,151	0.5%
Infection	7,912	8,062	1.9%
Non – Disease	19,344	18,856	-2.6%
Other	71,147	71,380	0.3%
Respiratory Disease – smoker related	28,092	28,522	1.5%
Total	538,600	541,589	0.6%

9.7 Comparison with Alternative Formulations

Subsequent sub-sections explore the trends for each driver compared with:

- ONS Primary ICD Code (A): A driver gleaned using the ONS Primary ICD code ignoring all the mentions. Note that the ONS Primary ICD code records COVID as the primary cause in a high number of deaths but not all deaths with a COVID mention are classed as COVID.

- COVID as seasonal flu (B): The proposed mapping with COVID treated as a seasonal flu i.e. classified in the Infection driver group and as a consequence moved down the priority order.

- Co-morbidity redistributed (C): The proposed mapping with Co-morbidity redistributed to other drivers.

- Proposed Mapping (D): The proposed mapping defined in section 6.

Figure 17 presents a heat map showing the movement of deaths between drivers formulated using the ONS primary cause of death and the proposed mapping. In summary:

- Consistency in mapping ranges from 36.5% (Infection) to 100.0% (COVID). This is due to the proposed mapping focusing on the drivers of mortality opposed to the effects with COVID prioritised in both mapping.
- Ignoring COVID the cancers show the highest consistency led by cancer unmet needs (90.9%).
- Co–Morbidity, which is not a driver in the ONS approach, absorbs the most deaths (15.2%) with deaths classified as Diabetes and Obesity deaths by the ONS showing the most significant movement (53.2%) into Co-Morbidity.

Appendix G – Movement between formulations summaries the movement in deaths between alternative formulations. It shows for example where deaths would have gone if our proposal excluded drivers for Co-Morbidity or COVID.

Figure 17 – Heat map showing the movement of deaths between drivers formulated using the ONS primary cause of death and the proposed mapping

| ONS Primary	Proposed Mapping												
	Alzheimer's Disease and Dementia	Cancer – Obesity Related	Cancer – Other	Cancer – Smoking Related	Cancer – Unmet need	Cardiovascular Disease	Co – Morbidity	Covid-19	Diabetes and Obesity	Infection	Non – Disease	Other	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	82.4%	0.5%	0.3%	0.2%	0.1%	0.0%	15.0%	0.3%	0.0%	0.0%	0.0%	1.3%	0.0%
Cancer – Obesity Related	0.0%	88.8%	0.0%	0.4%	0.2%	0.0%	10.5%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
Cancer – Other	0.5%	1.1%	78.6%	1.2%	0.7%	1.5%	13.3%	0.2%	0.3%	0.1%	0.0%	2.2%	0.3%
Cancer – Smoking Related	0.0%	0.1%	0.0%	85.4%	0.6%	0.0%	13.7%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%
Cancer – Unmet need	0.0%	0.0%	0.0%	0.0%	90.9%	0.0%	9.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
Cardiovascular Disease	5.4%	0.9%	0.7%	0.4%	0.5%	64.5%	12.0%	0.2%	5.1%	0.7%	0.4%	5.4%	3.8%
Covid-19	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Diabetes and Obesity	3.2%	0.2%	0.1%	0.1%	0.1%	0.0%	53.2%	0.3%	42.3%	0.0%	0.0%	0.0%	0.5%
Infection	7.4%	0.3%	0.4%	0.2%	0.2%	15.4%	12.4%	0.0%	4.2%	36.5%	0.7%	21.8%	0.6%
Non – Disease	3.0%	0.4%	0.5%	0.3%	0.3%	1.9%	7.1%	0.2%	0.3%	0.1%	81.8%	3.5%	0.5%
Other	3.6%	0.6%	0.5%	0.3%	0.4%	0.1%	29.7%	0.2%	1.0%	0.0%	0.7%	61.0%	1.9%
Respiratory Disease – smoker related	0.0%	0.7%	0.7%	0.6%	1.2%	0.0%	17.5%	0.1%	0.0%	0.0%	0.1%	0.0%	79.1%

9.7.1 Alzheimer's Disease and Dementia

Figure 18 shows "Alzheimer's Disease and Dementia" deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 13% increase "Alzheimer's Disease and Dementia" deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 45% increase if COVID and Co-morbidity are distributed back to other drivers.
- "Alzheimer's Disease and Dementia" trends upwards in both scenarios however the Proposed Mapping, which benefits from consistent application in all years, addresses a large discontinuity in the ONS Primary ICD code pre-2011.
- Over half of deaths being recategorized as "Alzheimer's Disease and Dementia" were previously classed under cardiovascular disease.
- COVID and Co-morbidity absorb 22% of deaths that would have fallen under "Alzheimer's Disease and Dementia" in the proposed mapping (No COVID / No Co-morbidity).

Figure 18 – Alzheimer's Disease and Dementia Deaths by Calendar Year
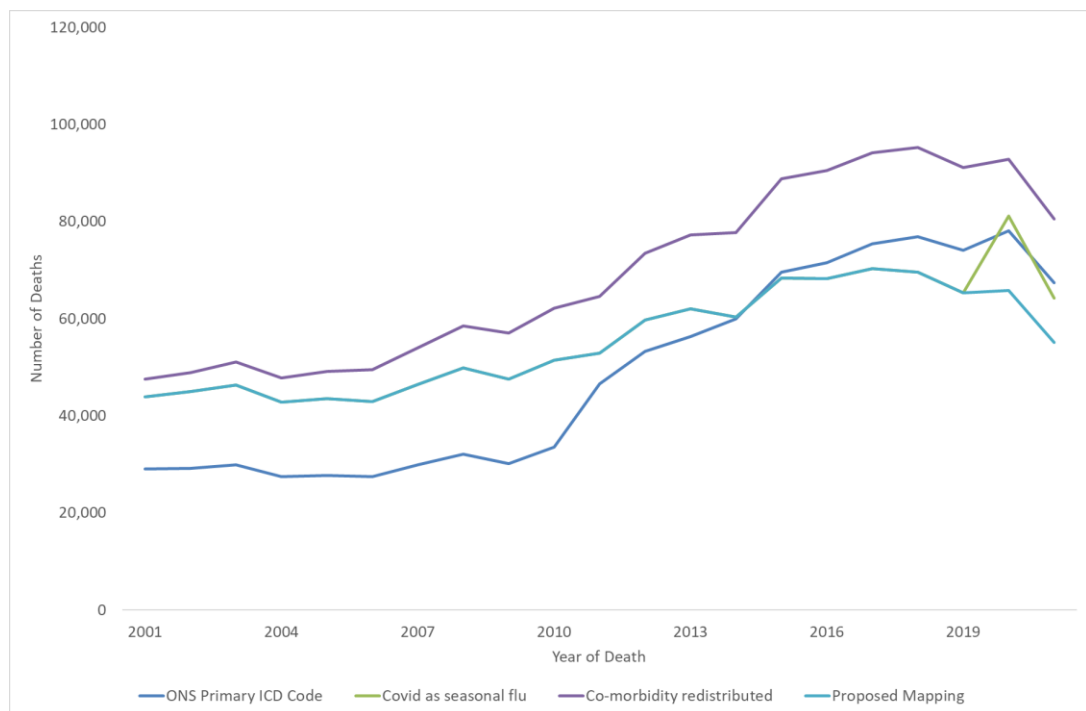
9.7.2 Diabetes and Obesity

Figure 19 shows Diabetes and Obesity deaths by calendar year for the scenarios described in the introduction.
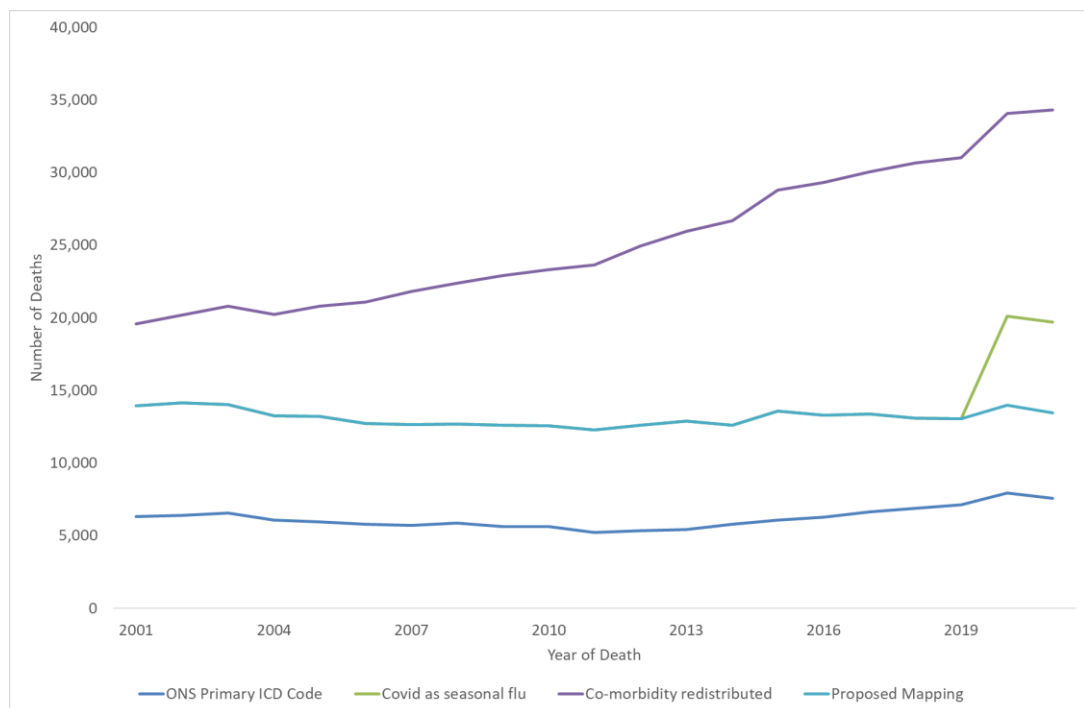
The following observations are noted:

- The Proposed Mapping approach leads to a 112% increase in Diabetes and Obesity deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 327% increase if COVID and Co-morbidity are distributed back to other drivers.

- Diabetes and Obesity trends appear stable although when the deaths shifted to co-morbidity are included, it trends upwards which is aligned with observations from the World health Organisation[30] which include:
  - The number of people with diabetes is on the rise.
  - Diabetes is a major cause of blindness, kidney failure, heart attacks, stroke and lower limb amputation.
  - Between 2000 and 2019, there was a 3% increase in diabetes mortality rates by age.
- Two-thirds of deaths being recategorized as Diabetes and Obesity were previously classed under cardiovascular disease.
- COVID and Co-morbidity absorb 50% of deaths that would have fallen under Diabetes and Obesity.

Figure 19 – Diabetes and Obesity Deaths by Calendar Year

9.7.3   Cardiovascular Disease

Figure 20 shows cardiovascular disease deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 32% decrease (over one million deaths) in cardiovascular disease deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 27% decrease if COVID and Co-morbidity are distributed back to other drivers.
- Cardiovascular disease trends downwards in both scenarios with the ONS Primary ICD Code derived driver suggesting a gentler slowdown in this trend from 2011.
- The Proposed Mapping shifts cardiovascular disease deaths to mainly Other, Diabetes and Obesity, "Alzheimer's Disease and Dementia" and Respiratory Disease – smoker related.

---

[30] Diabetes (who.int)

- COVID and Co-morbidity absorb 6% of deaths that would have fallen under cardiovascular disease; the percentage is small due to the shift of cardiovascular deaths into other significant drivers.

Figure 20 – Cardiovascular Disease deaths by calendar year

9.7.4 Non-disease

Figure 21 shows Non-Disease deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

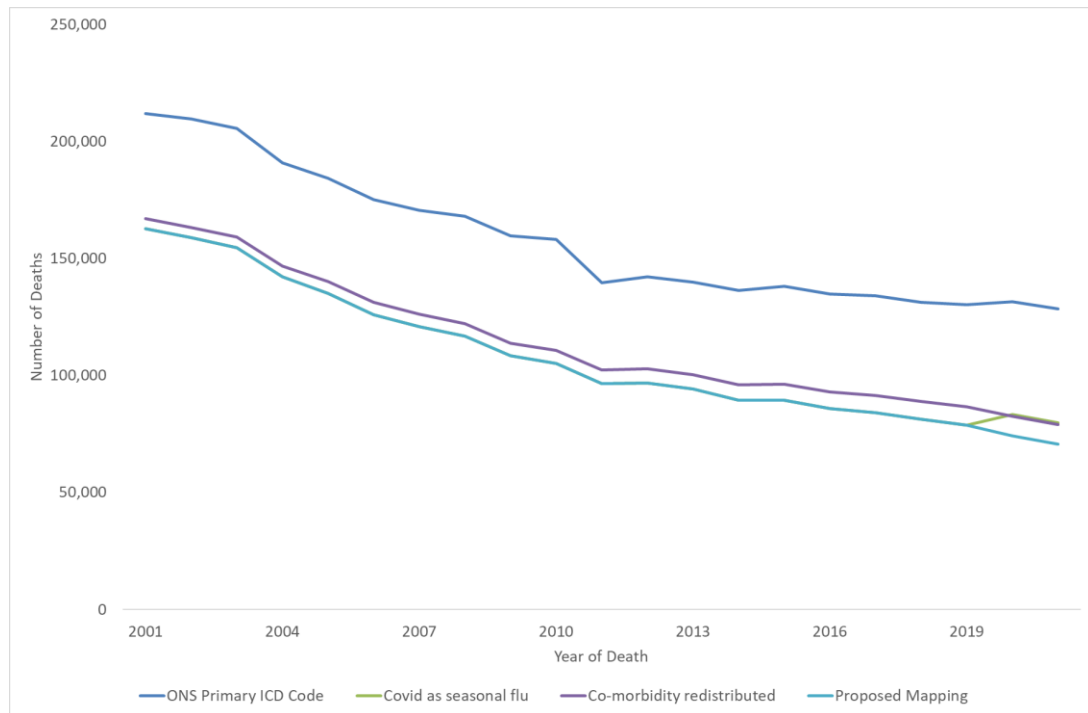- The Proposed Mapping approach leads to a 11% decrease in Non-Disease deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 6% decrease if COVID and Co-morbidity are distributed back to other drivers.

- The trends are similar in pattern and stable in earlier years: however, it diverges from 2011 onwards where the Proposed Mapping suggests a gentler increase in Non-Disease deaths compared to the ONS Primary ICD Code.

- There is a sharp fall in non-disease deaths during the COVID period (2020-21); the fall in the later year is driven primarily by late registration of deaths.

- The Proposed Mapping shifts Non-Disease deaths to mainly Other, "Alzheimer's Disease and Dementia" and Cardiovascular Disease. A judgement was made in the proposed approach that for example an accidental death with a "Alzheimer's Disease and Dementia" mention is driven by the disease.

- COVID and Co-morbidity absorb 5% of deaths that would have fallen under Non-Disease.

Figure 21 – Non-Disease deaths by calendar year

9.7.5   Cancer Unmet Needs

Figure 22 shows Cancer Unmet Needs deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 5% decrease in Cancer Unmet Needs deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 7% increase if COVID and Co-morbidity are distributed back to other drivers.

- The trends are similar in pattern in earlier years with the gap widening in recent years where the Proposed Mapping suggests a slowdown. However, note that when deaths being assigned to Co-morbidity are added back, an increasing trend is observed.

- COVID and Co-morbidity absorb 11% of deaths that would have fallen under Cancer Unmet Needs.

Figure 22 – Cancer (Unmet Needs) Deaths by Calendar Year

9.7.6   Cancer Smoker Related

Figure 23 shows Cancer Smoker Related deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 3% decrease in Cancer Smoker Related deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 19% increase if COVID and Co-morbidity are distributed back to other drivers.

- Compared to the ONS Primary ICD Code gleaned driver, the Proposed Mapping suggests a slowdown in recent years. However, note that when deaths being assigned to Co-morbidity are added back, an increasing trend is observed.

- COVID and Co-morbidity absorb 19% of deaths that would have fallen under Cancer Smoker Related.

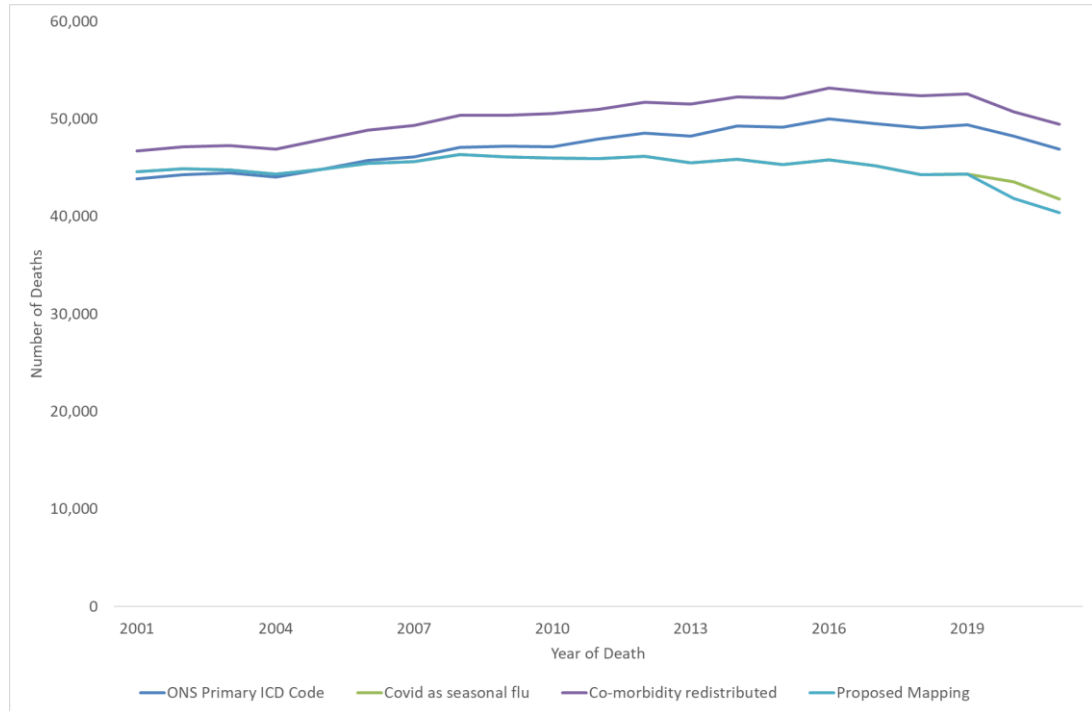**Figure 23 – Cancer Smoker Related Deaths by Calendar Year**

9.7.7 Cancer Obesity Related

Figure 24 shows Cancer Obesity Related deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 4% decrease in Cancer Obesity Related deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 13% increase if COVID and Co-morbidity are distributed back to other drivers.

- Compared to the ONS Primary ICD Code gleaned driver, the Proposed Mapping suggests a slowdown in recent years. However, note that when deaths being assigned to Co-morbidity are added back, an increasing trend is observed.

- COVID and Co-morbidity absorb 15% of deaths that would have fallen under Cancer Obesity Related.
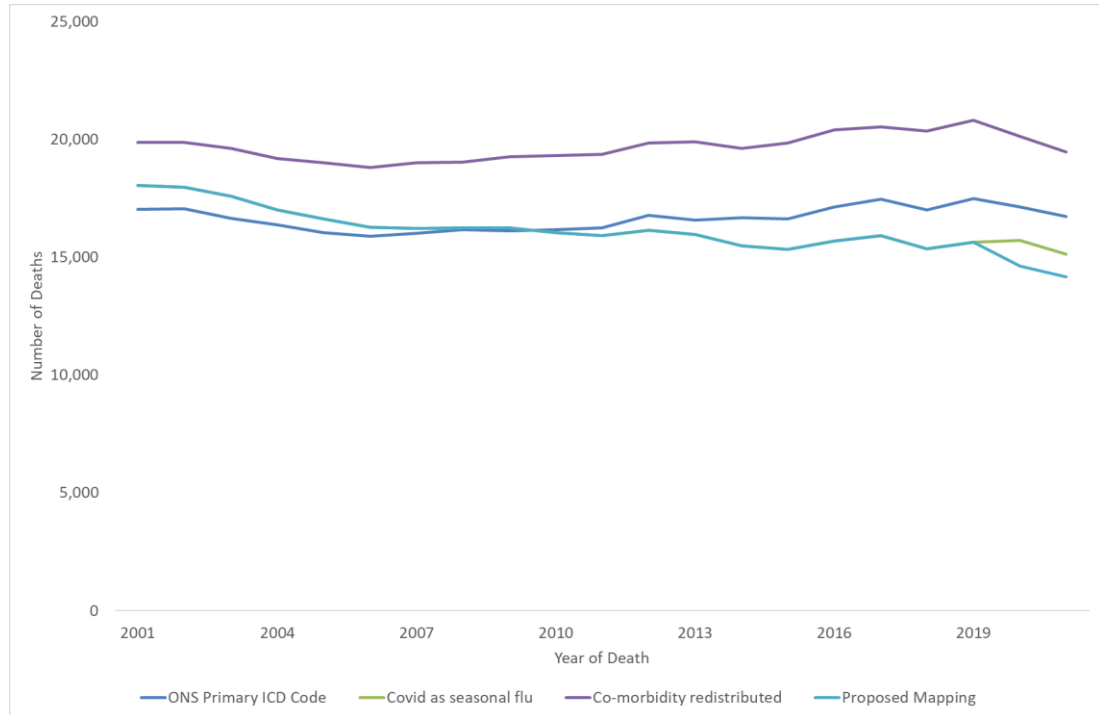
Figure 24 – Cancer Obesity Related Deaths by Calendar Year

9.7.8 Cancer Other

Figure 25 shows Cancer Other deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 16% decrease in Cancer Other deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 1% decrease if COVID and Co-morbidity are distributed back to other drivers.

- Compared to the ONS Primary ICD Code gleaned driver, the Proposed Mapping suggests a widening gap which is a result of movement of deaths into the co-morbidity driver.

- COVID and Co-morbidity absorb 15% of deaths that would have fallen under Cancer Other.
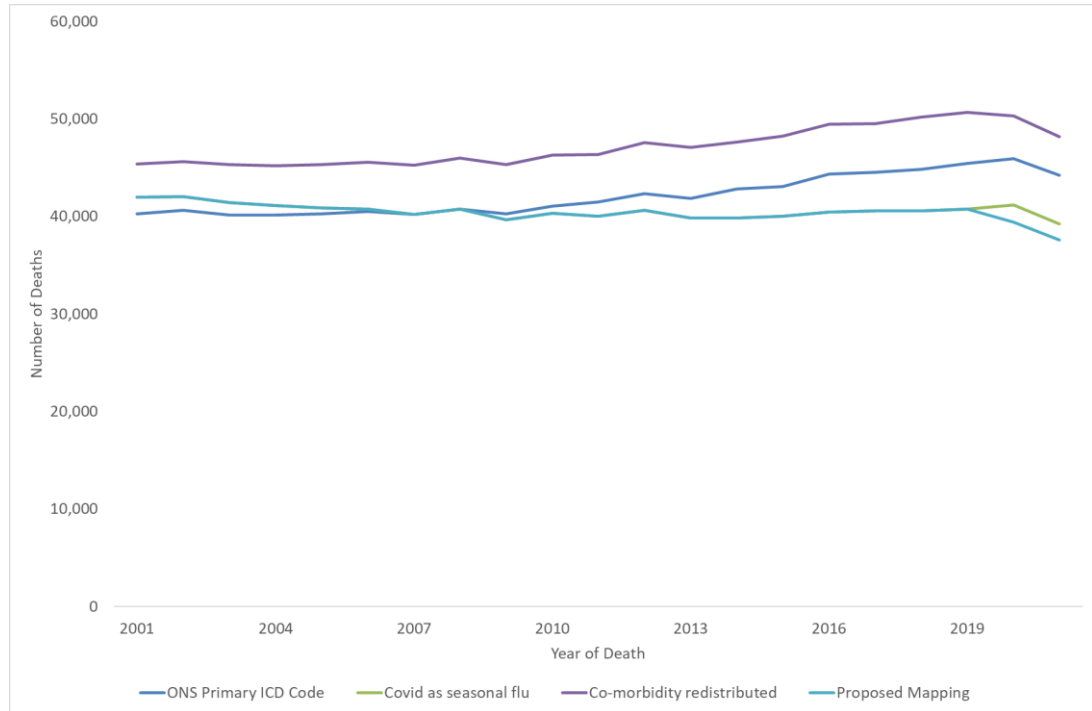
Figure 25 – Cancer Other Deaths by Calendar Year

9.7.9 Respiratory Disease – Smoker Related

Figure 26 shows Respiratory Disease- Smoker Related deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 10% increase in Respiratory Disease- Smoker Related deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 48% increase if COVID and Co-morbidity are distributed back to other drivers.

- Compared the ONS Primary ICD Code gleaned driver, the Proposed Mapping suggests an increasing trend in more recent years when co-morbidity deaths are factored.

- COVID and Co-morbidity absorb 26% of deaths that would have fallen under Respiratory Disease- Smoker Related.

Figure 26 – Respiratory Disease- Smoker Related Deaths by Calendar Year
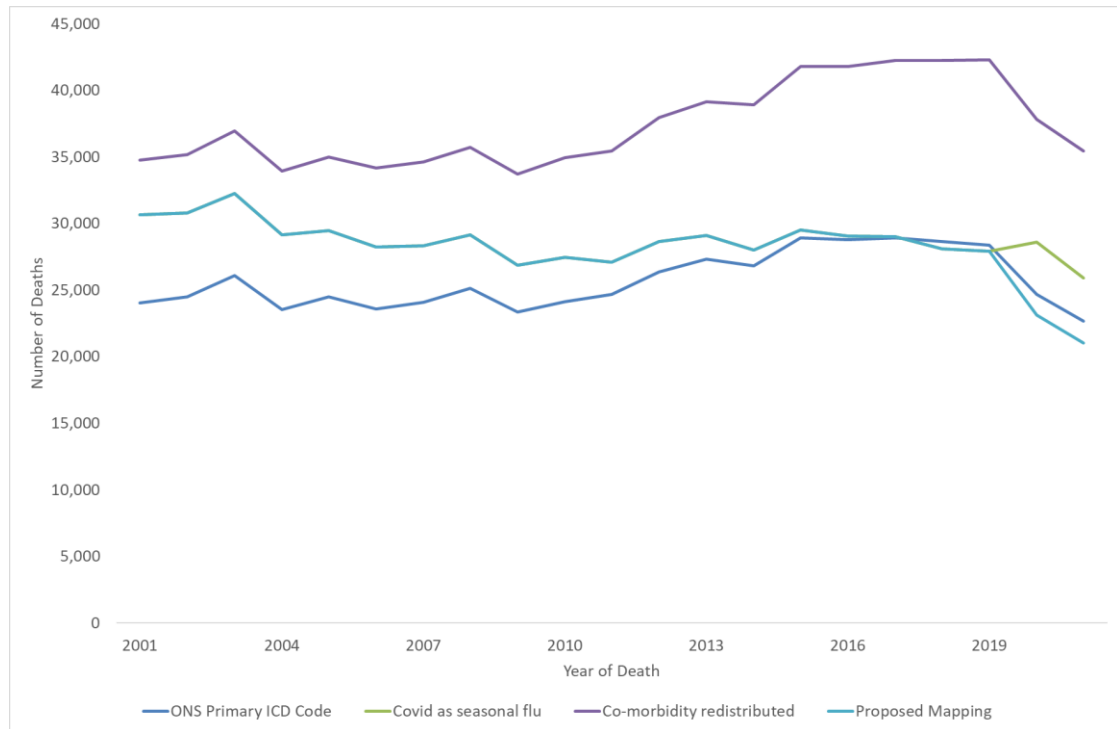
9.7.10 Infection

Figure 27 shows Infection deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 60% decrease in Infection deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 58% decrease if COVID and Co-morbidity are distributed back to other drivers.
- The Proposed Mapping driver trends downwards and no deaths are moved in the co-morbidity driver likely due to prioritisation process in the proposed mapping approach.
- COVID and Co-morbidity absorb 5% of deaths that would have fallen under Infection.
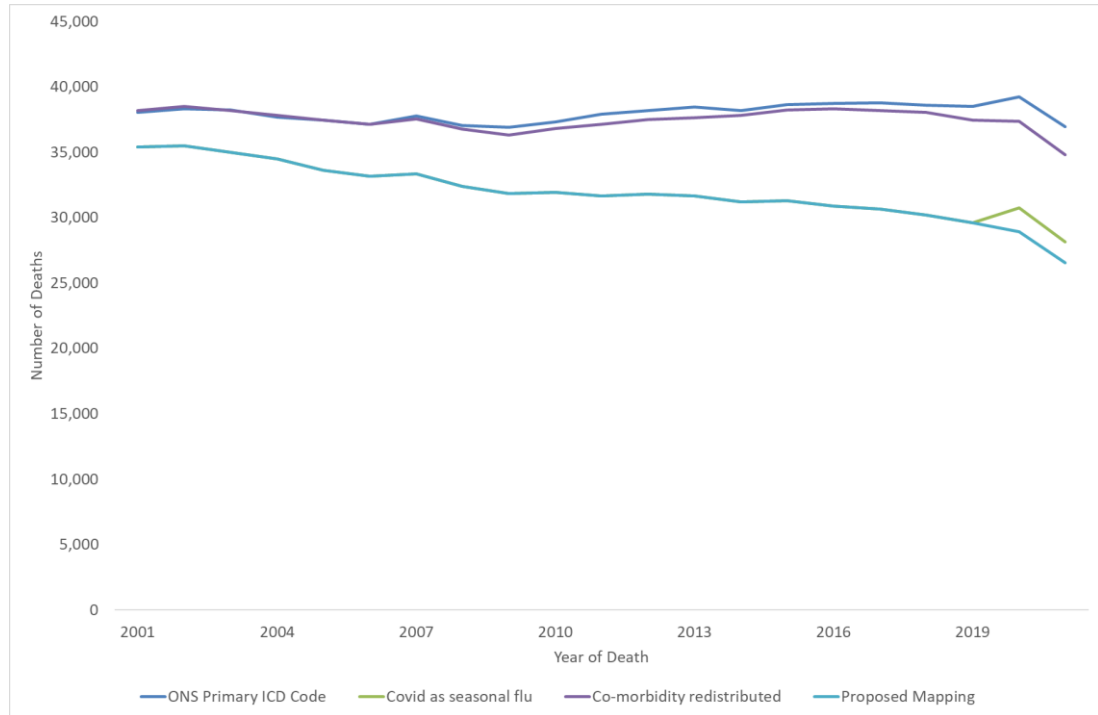
Figure 27 – Infection Deaths by Calendar Year
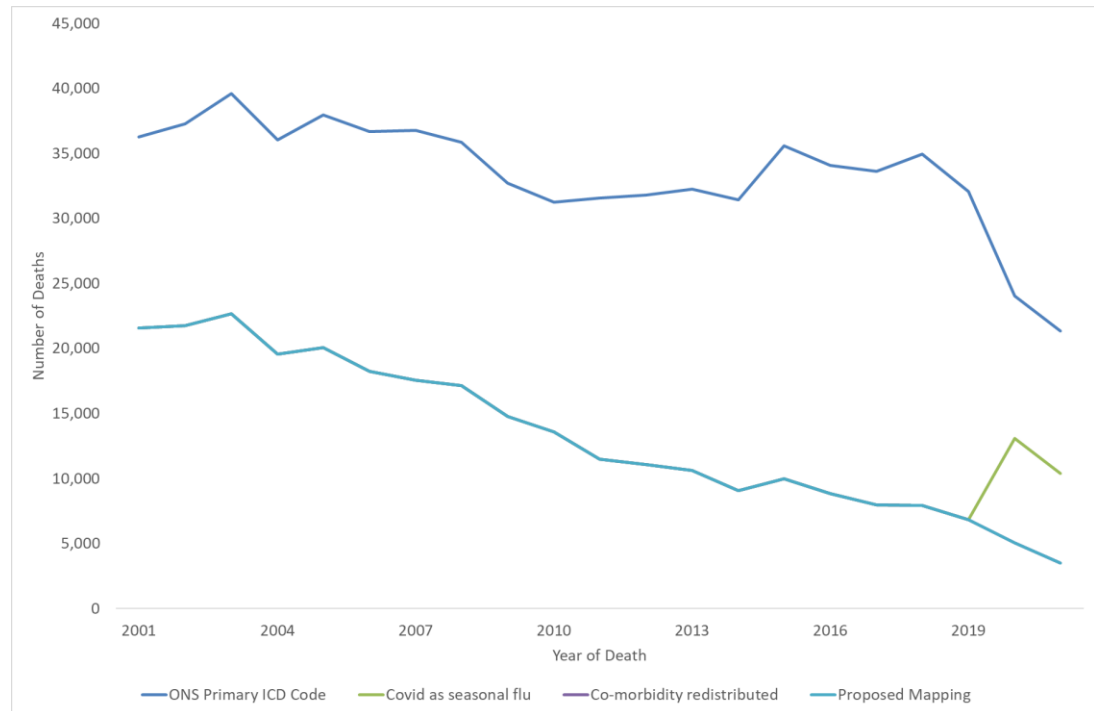
9.7.11 Other

Figure 28 shows Other deaths by calendar year for the scenarios described in the introduction.

The following observations are noted:

- The Proposed Mapping approach leads to a 16% decrease in Other deaths in comparison to a driver gleaned using only the ONS Primary ICD Code; 5% increase if COVID and Co-morbidity are distributed back to other drivers.

- The patterns are similar and trend upwards in more recent years; ONS primary shows a sharp increase in older periods which stabilise in 2008.

- Approximately 60% of deaths being recategorized as Other were previously classed under cardiovascular disease.

- COVID and Co-morbidity absorb 21% of deaths that would have fallen under Other.

Figure 28 – Other Deaths by Calendar Year
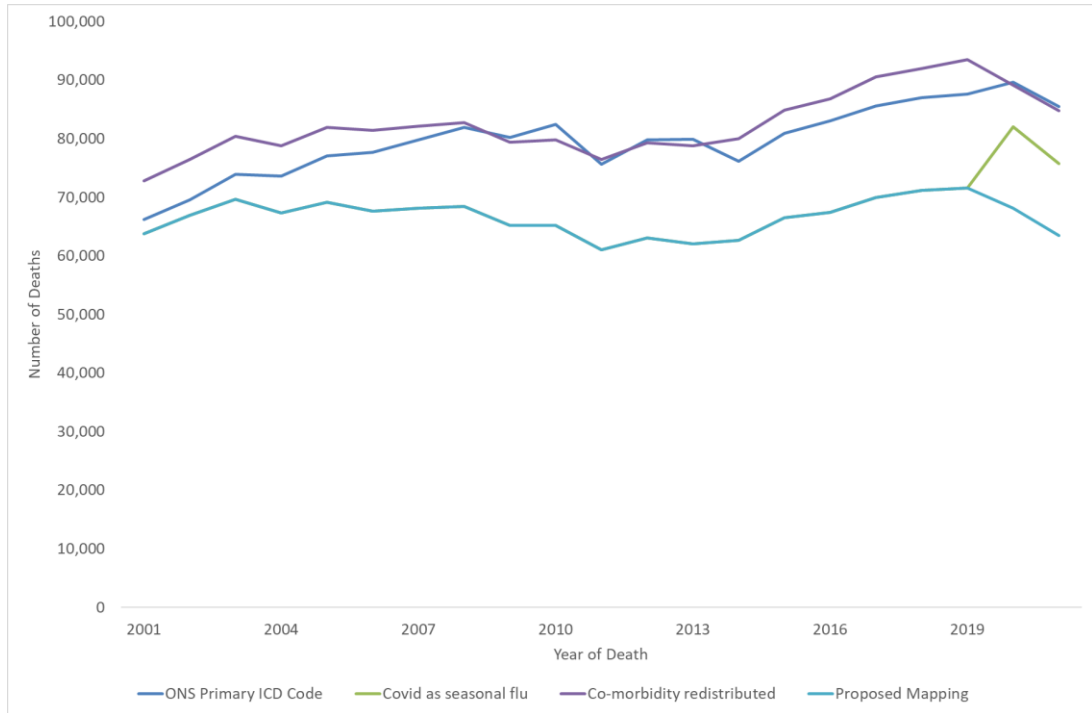
# 10 Limitations and Potential Further Development

10.1 Introduction

The section covers:

- A list of limitations accepted in the creation of the published dataset
- A list of areas that could be developed further.

10.2 List of Limitations

Limitations:

- Information stored on death certificates evolves over time due to increasing demands for better data collection. For example, the change in trend relating to number of mentions recorded on deaths could be due to recording changes and result in inaccurate classification of drivers e.g. Co-morbidity which relies on the number of mentions or the absence of a mention which outranks another e.g. a cardiomyopathy death with diabetes, which outranks cardiovascular disease, not recorded.
- The IFoA Cause of Death mapping list covers all ICD codes in the mortality data: however, it does not cover all ICD codes in the WHO library. The list needs to be expanded and reviewed periodically or when changes are made to the WHO library.

10.3 List of Potential Further Developments

Areas of further developments:

- Consider alternative mapping rules e.g.
  - Alternative definitions of co-morbidity
  - Breaking up the "Other" driver which covers 13% of deaths over the study period.
  - Sensitivity testing the prioritisation system deployed by the mapping rules e.g. what is the impact of placing "Alzheimer's Disease and Dementia" ahead the cancers.
  - Splitting up drivers e.g. cardiovascular disease can be split into coronary heart disease and stroke.
- Deaths reported pre-2001 are excluded due to a lack of ICD-10 mapping. This could be resolved by:
  - Creating a cause of death mapping for ICD-9 and other previous versions[31].
  - A manual exercise or application of a free text classification program to map death descriptions to the ICD-10 codes.

---

[31] International Classification of Diseases (ICD) (who.int)

11 Appendix A – ONS Mortality Data

11.1 Introduction

This section provides the specification for the ONS mortality data.

11.2 Raw Data Specification

The specification of the raw data is specified in the file "Deaths Data v2.docx" located on the SRS network. Table 18 summarises the descriptions of the fields listed in the above document.

Table 18 – ONS Raw Mortality Dataset Specification

Field	Description
AGEC	Calculated age of the deceased
AGEC	Calculated age of the deceased
AGECS	Calculated age of spouse
AGECUNIT	Calculated age unit
AGEINYRS	Calculated age in years
AGEU1DST	For deaths occurring under 24 hours, the registrar provides details of the age at death
CC10001	Cause of death column position mention 1-15
CC10002	Cause of death column position mention 1-15
CC10003	Cause of death column position mention 1-15
CC10004	Cause of death column position mention 1-15
CC10005	Cause of death column position mention 1-15
CC10006	Cause of death column position mention 1-15
CC10007	Cause of death column position mention 1-15
CC10008	Cause of death column position mention 1-15
CC10009	Cause of death column position mention 1-15
CC10010	Cause of death column position mention 1-15
CC10011	Cause of death column position mention 1-15
CC10012	Cause of death column position mention 1-15
CC10013	Cause of death column position mention 1-15
CC10014	Cause of death column position mention 1-15
CC10015	Cause of death column position mention 1-15
CC10F001	Final cause of death column position mention 1-15
CC10F002	Final cause of death column position mention 1-15
CC10F003	Final cause of death column position mention 1-15
CC10F004	Final cause of death column position mention 1-15
CC10F005	Final cause of death column position mention 1-15
CC10F006	Final cause of death column position mention 1-15
CC10F007	Final cause of death column position mention 1-15
CC10F008	Final cause of death column position mention 1-15
CC10F009	Final cause of death column position mention 1-15
CC10F010	Final cause of death column position mention 1-15
CC10F011	Final cause of death column position mention 1-15
CC10F012	Final cause of death column position mention 1-15
CC10F013	Final cause of death column position mention 1-15

Field	Description
CC10F014	Final cause of death column position mention 1-15
CC10F015	Final cause of death column position mention 1-15
CERTDETS	Type of medical certificate presented
CERTIFIC	When a cause of death has been certified an indicator is applied
CERTTYPE	Shows the type of death certification
CESTRSS	Communal establishment code
CESTRSSR	Communal establishment code of residence
CESTSTAY	Duration of stay in communal establishment
CR10001	Cause of death row position mention 1-15
CR10002	Cause of death row position mention 1-15
CR10003	Cause of death row position mention 1-15
CR10004	Cause of death row position mention 1-15
CR10005	Cause of death row position mention 1-15
CR10006	Cause of death row position mention 1-15
CR10007	Cause of death row position mention 1-15
CR10008	Cause of death row position mention 1-15
CR10009	Cause of death row position mention 1-15
CR10010	Cause of death row position mention 1-15
CR10011	Cause of death row position mention 1-15
CR10012	Cause of death row position mention 1-15
CR10013	Cause of death row position mention 1-15
CR10014	Cause of death row position mention 1-15
CR10015	Cause of death row position mention 1-15
CR10F001	Final cause of death row position mention 1-15
CR10F002	Final cause of death row position mention 1-15
CR10F003	Final cause of death row position mention 1-15
CR10F004	Final cause of death row position mention 1-15
CR10F005	Final cause of death row position mention 1-15
CR10F006	Final cause of death row position mention 1-15
CR10F007	Final cause of death row position mention 1-15
CR10F008	Final cause of death row position mention 1-15
CR10F009	Final cause of death row position mention 1-15
CR10F010	Final cause of death row position mention 1-15
CR10F011	Final cause of death row position mention 1-15
CR10F012	Final cause of death row position mention 1-15
CR10F013	Final cause of death row position mention 1-15
CR10F014	Final cause of death row position mention 1-15
CR10F015	Final cause of death row position mention 1-15
CTRYIR	Country of death (England, Wales, Scotland, Northern Ireland)
CTRYPOB	Country code of place of birth of deceased
CTRYR	Country code of usual residence of deceased
DEC_STAT_DOB	Date of birth of deceased
DESIGC	Senior Coroner, Acting Senior Coroner, etc…
DESIGR	Registrar, Deputy Registrar, Interim Registrar, Deputy Interim Registrar
DOBDY	Birthday
DOBMT	Birth month

Field	Description
DOBS	Date of birth of spouse
DOBYR	Birth year
DODDY	Death day
DODMT	Death month
DODYR	Death year
DOLSA	Date deceased last seen alive by the deceased's own doctor, does not apply to deaths certified by a coroner
DOR	Date the death was registered
DWIGS10	Office for National Statistics Hierarchical Classification of neonatal Deaths for ICD10
EMPSECDM	Employment status of deceased or mother of deceased juvenile for SOC
EMPSECHF	Employment status (ONS Code) of deceased's husband or father/parent of deceased juvenile for SOC
ESTTYPED	Establishment type where death occurred
FIC10IND	ONS primary ICD code
FIC10MEN1	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN10	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN11	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN12	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN13	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN14	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN15	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN2	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN3	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN4	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN5	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN6	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN7	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN8	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10MEN9	Final ICD10 code mention 1-15 (Used from 2017 onwards)
FIC10UND	This indicates if there is a final cause of death on the death record
FIMDTH10	ICD indicator field- not defined
I10P001	ICD provisional code mention 1-15
I10P002	ICD provisional code mention 1-16
I10P003	ICD provisional code mention 1-17
I10P004	ICD provisional code mention 1-18
I10P005	ICD provisional code mention 1-19
I10P006	ICD provisional code mention 1-20
I10P007	ICD provisional code mention 1-21
I10P008	ICD provisional code mention 1-22
I10P009	ICD provisional code mention 1-23
I10P010	ICD provisional code mention 1-24
I10P011	ICD provisional code mention 1-25
I10P012	ICD provisional code mention 1-26
I10P013	ICD provisional code mention 1-27
I10P014	ICD provisional code mention 1-28
I10P015	ICD provisional code mention 1-29

Field	Description
I10PF001	Final ICD provisional code mention 1-15
I10PF002	Final ICD provisional code mention 1-16
I10PF003	Final ICD provisional code mention 1-17
I10PF004	Final ICD provisional code mention 1-18
I10PF005	Final ICD provisional code mention 1-19
I10PF006	Final ICD provisional code mention 1-20
I10PF007	Final ICD provisional code mention 1-21
I10PF008	Final ICD provisional code mention 1-22
I10PF009	Final ICD provisional code mention 1-23
I10PF010	Final ICD provisional code mention 1-24
I10PF011	Final ICD provisional code mention 1-25
I10PF012	Final ICD provisional code mention 1-26
I10PF013	Final ICD provisional code mention 1-27
I10PF014	Final ICD provisional code mention 1-28
I10PF015	Final ICD provisional code mention 1-29
ICD10001	ICD10 code mention 1-15
ICD10002	ICD10 code mention 1-16
ICD10003	ICD10 code mention 1-17
ICD10004	ICD10 code mention 1-18
ICD10005	ICD10 code mention 1-19
ICD10006	ICD10 code mention 1-20
ICD10007	ICD10 code mention 1-21
ICD10008	ICD10 code mention 1-22
ICD10009	ICD10 code mention 1-23
ICD10010	ICD10 code mention 1-24
ICD10011	ICD10 code mention 1-25
ICD10012	ICD10 code mention 1-26
ICD10013	ICD10 code mention 1-27
ICD10014	ICD10 code mention 1-28
ICD10015	ICD10 code mention 1-29
ICD10F001	Final ICD10 code mention 1-15
ICD10F002	Final ICD10 code mention 1-16
ICD10F003	Final ICD10 code mention 1-17
ICD10F004	Final ICD10 code mention 1-18
ICD10F005	Final ICD10 code mention 1-19
ICD10F006	Final ICD10 code mention 1-20
ICD10F007	Final ICD10 code mention 1-21
ICD10F008	Final ICD10 code mention 1-22
ICD10F009	Final ICD10 code mention 1-23
ICD10F010	Final ICD10 code mention 1-24
ICD10F011	Final ICD10 code mention 1-25
ICD10F012	Final ICD10 code mention 1-26
ICD10F013	Final ICD10 code mention 1-27
ICD10F014	Final ICD10 code mention 1-28
ICD10F015	Final ICD10 code mention 1-29
ICD10SC	This field identifies the nature of injury when the underlying cause of death (ICD10U) is an external cause.

Field	Description
ICD10SCF	This field identifies the final nature of injury when the final underlying cause of death (ICD10UF) is an external cause
ICD10U	This field is a cause code that identifies the medical condition judged to be the underlying cause of death according to the rules of 10th Revision of the International Classification of Diseases
IMLINKNO	Not defined and blank
INDDMT	Education, gardening, sport, writing, leisure, - not defined
INDHFT	Not defined- sample Joinery, None, transport, Finance
INFQUAL	Codes- not defined
INQCERT	Codes- not defined
MARSTAT	Marital status (Single, Divorced, Widowed, Civil Partnership, Civil Partnership Widowed, Civil Partnership Dissolved)
MDTH10	Manner of death code indicates the verdict of intent or reason behind a traumatic death
MDTH10F	Final manner of death code indicates the verdict of intent or reason behind a traumatic death
NHSIND	NHS establishment indicator
OCCHFT	Sample: Cleaner, gardener, housewife, etc…
OCCTYPE	Occupation type
PCDPOD	Postcode of place of death
PCDR	Postcode of usual residence
PLOACC10	Place of accident
PMTYPE	Under whose authority the Post-mortem took place
POSTMORT	Indicator to show if a post-mortem has been held
REFCOR	Indicates if the death has been referred to the coroner by the registrar or doctor
REG_STAT_DOD	Start of registration? Not defined however similar to DOR but earlier
REGDETS	Details related to registration of deaths
RONUNQID	Some form of unique ID
SADIND	Indicates whether the body of the deceased was seen after death and whether this was by the certifying doctor or another doctor
SECCATDM	National Statistics Socio-economic Classification operational category for deceased or mother of deceased
SECCATHF	National Statistics Socio-economic Classification operational category for husband of deceased
SECCLRDM	Socio-economic classification for mother of child
SECCLRHF	Socio-economic classification for father/parent of child
SEX	Male= 1 Female=2 Unknown=3
SOC2KDM	Number- not defined
SOC2KHF	Standard Occupation Classification for husband or father/parent of deceased child
VMLID	Unique record identifier
WPLA10	Code indicating whether or not, in cases of violent death, the incident resulting in death took place at work.

12   Appendix B – ICD Code to ONS Cause of death

12.1   Introduction

This appendix details mapping from ICD-10 10 code to the ONS 'cause of death' classification; see Leading causes of death in England and Wales (revised 2016) - Office for National Statistics (ons.gov.uk)

12.2   Classification

Table 19 – ONS ICD Code Mapping

ICD10 Code	ONS Cause of Death
A00–A09	Intestinal infectious diseases
A15–A19, B90	Tuberculosis
A20, A44, A75–A79, A82–A84, A85.2, A90–A98, B50–B57	Vector–borne diseases and rabies
A33–A37, A49.2, A80, B01, B02, B05, B06, B15, B16, B17.0, B18.0, B18.1, B26, B91, G14	Vaccine-preventable diseases1
A39, A87, G00–G03	Meningitis and meningococcal infection
A40–A41	Septicaemia
B20–B24	Human immunodeficiency virus [HIV] disease
C00–C97	Malignant neoplasms
C15	Malignant neoplasm of oesophagus
C16	Malignant neoplasm of stomach
C18–C21	Malignant neoplasm of colon, sigmoid, rectum and anus
C22	Malignant neoplasm of liver and intrahepatic bile ducts
C23-C24	Malignant neoplasm of gallbladder and other parts of biliary tract
C25	Malignant neoplasm of pancreas
C32	Malignant neoplasm of larynx
C33-C34	Malignant neoplasm of trachea, bronchus and lung
C40–C41	Malignant neoplasms of bone and articular cartilage
C43-C44	Melanoma and other malignant neoplasms of skin
C50	Malignant neoplasm of breast
C53–C55	Malignant neoplasm of uterus
C56	Malignant neoplasm of ovary
C61	Malignant neoplasm of prostate
C64	Malignant neoplasm of kidney, except renal pelvis
C67	Malignant neoplasm of bladder
C71	Malignant neoplasm of brain
C81–C96	Malignant neoplasms, stated or presumed to be primary of lymphoid, haematopoietic and related tissue
D00–D48	In situ and benign neoplasms, and neoplasms of uncertain or unknown behaviour

ICD10 Code	ONS Cause of Death
E10–E14	Diabetes
D50–D53, E40–E64	Malnutrition, nutritional anaemias and other nutritional deficiencies
E86–E87	Disorders of fluid, electrolyte and acid–base balance (incl. dehydration)
F01, F03, G30	Dementia and Alzheimer disease
F10–F19	Mental and behavioural disorders due to psychoactive substance use
G10–G12	Systemic atrophies primarily affecting the central nervous system
G20	Parkinson disease
G40-G41	Epilepsy and status epilepticus
G80–G83	Cerebral palsy and other paralytic syndromes
I05–I09	Chronic rheumatic heart diseases
I10–I15	Hypertensive diseases
I20–I25	Ischaemic heart diseases
I26–I28	Pulmonary heart disease and diseases of pulmonary circulation
I34–I38	Nonrheumatic valve disorders and endocarditis
I42	Cardiomyopathy
I46	Cardiac arrest
I47–I49	Cardiac arrhythmias
I50–I51	Heart failure and complications and ill–defined heart disease
I60–I69	Cerebrovascular diseases
I70	Atherosclerosis
I71	Aortic aneurysm and dissection
J00–J06, J20–J22	Acute respiratory infections other than influenza and pneumonia
J09–J18	Influenza and pneumonia
J40–J47	Chronic lower respiratory diseases
J80–J84	Pulmonary oedema and other interstitial pulmonary diseases
J96	Respiratory failure
K35–K46, K56	Appendicitis, hernia and intestinal obstruction
K70–K76	Cirrhosis and other diseases of liver
M00–M99	Diseases of the musculoskeletal system and connective tissue
N00–N39	Diseases of the urinary system
O00–O99	Pregnancy, childbirth and the puerperium
P00–P96	Certain conditions originating in the perinatal period
Q00–Q99	Congenital malformations, deformations and chromosomal abnormalities
V01–X59	Accidents
V01–V89	Land transport accidents
W00–W19	Accidental falls
W32–W34	Non-intentional firearm discharge
W65–W74	Accidental drowning and submersion
W75–W84	Accidental threats to breathing

ICD10 Code	ONS Cause of Death
X40–X49	Accidental poisoning
X60–X84, Y10–Y34	Suicide and injury/poisoning of undetermined intent2
U50.9, X85–Y09, Y87.1	Homicide and probable homicide
R00-R99	Symptoms, signs and ill-defined conditions

13   Appendix C – ICD Code to IFoA Cause of Death

13.1   Introduction

This appendix details mapping from ICD-10 10 code to the IFoA 'cause of death' classification.

13.2   Classification

Table 20 shows the mapping of ICD-10 codes to the IFoA 'cause of death' to the drivers applicable to single mentions. The primary source for the description of the ICD-10 codes can be found in ICD-10 Version:2019 (who.int). For codes not found in above link:

- Information is gathered from search engines and other sites (e.g. BMJ),
- ONS datasets found within their SafePod network drivers or
- Based on the first 3 characters where the complete 4-character code could not be sourced

Table 20 – ICD Code Mapping

ICD10 Code	Cause of Death	Drivers[32]
F01-F03, G30	Alzheimer's Disease and Dementia	Alzheimer's Disease and Dementia
R540	Frailty (Senility, Frailty, Old age, …)	Alzheimer's Disease and Dementia
Series beginning with "A excluding A33-34", series beginning with B and G00-G07	Infection	Infection
U83	Antibiotic Resistance	Infection
Series beginning with V, Y32, Y85,	Vehicular Accident	Non – Disease
X33-X38	Natural Disaster	
Series beginning with W, X00-X32, X39-X52, X58-X59, Y10-Y31, Y33-Y34,	Accident	
Series beginning with S, T00-T14, T90-T98	Injury	
T15-T78, T80, X53-X57, X85-X99, Y00-Y09, Z57, Z58, Z77	External Causes	
F10-19, R78, Y90, Z714, Z715, Y90	Drug-Alcohol	
T81-T89, Y40-Y84, Y88, D78, E36, G973-G976, H591-H594, H952-H955, I97, J95, K91, K94, K95, L76, L80	Medical Treatment Complications	
X60-X84	Suicide	
U071, U072, U099	Covid-19	Covid-19
E08-E14	Diabetes	Diabetes and Obesity
E65, E66, Z683, Z684	Obesity	
G45-G46	Stroke Other	Cardiovascular Disease
I64	Stroke	

---

[32] Applicable to single mentions only. Multiple mentions follows decision tree; see section on mapping approach.

IFoA Drivers of Mortality Dataset

ICD10 Code	Cause of Death	Drivers[32]
I65-I69	Other Cerebrovascular Disease	
I01, I090-I092, I10-I20, I44-I45, I47-I49, I51-I52, I5A, I73-I78, I1791, I798, I81-I82, I87-I96, I99	Other CHD	
I05-I08, I098-I099, I34-I37	Valve Disorder	
I21	AMI	
I22-I25	Ischaemic Heart Diseases	
I26-I28	Pulmonary Heart Disease	
I30-I33, I38-I41	Inflammation of the Heart	
I42-I43	Cardiomyopathy	
I46	Cardiac Arrest	
I50	Heart Failure	
I60-I62	Intracranial Haemorrhage	
I63	Cerebral Infarction	
I70	Atherolsclerosis	
I71-I72, I790	Aneurysm	
I80	Phlebitis and Thrombophlebitis	
J09-J18	Influenza and Pneumonia	Infection
J20-J22, J40-J42	Bronchial Disease	Respiratory Disease – smoker related
J43	Emphysema	
J44	COPD	
C00-C14, C30-C32, C490, C760, D000, D020	Head and Neck Cancer	Cancer – Smoking Related
C15, D001	Oesophagus Cancer	Cancer – Unmet need
C33-C37, C39, D021-D024	Cancer of trachea, bronchus and lung	Cancer – Unmet need
C25	Pancreatic Cancer	Cancer – Unmet need
C71	Brain Cancer	Cancer – Unmet need
C67, D090	Bladder Cancer	Cancer – Smoking Related
C51-C52, C57 excl. C574, C58, D070-D073	Gynaecological Cancer	
C16	Stomach Cancer	
C91-C95	Leukaemia	
C64-C65	Kidney Cancer	Cancer – Obesity Related
C90	Myeloma	
C17-C21, D010-D014	Colo-Rectal Cancer	
C22	Liver Cancer	
C23-C24	Gallbladder Cancer	
C40-C41	Bone Cancer	
C53-C55, C574, D06	Uterine Cancer	
C56	Ovarian Cancer	
C50, D050-D059,	Breast Cancer	

ICD10 Code	Cause of Death	Drivers[32]
C43, D03,	Melanoma	Cancer – Other
C44, C4A, D04	Skin Cancer	
C45	Mesothelioma	
C77, C82-C86, C884	Lymphoma	
C61, D075	Prostate Cancer	
C26, C38, C46-C48, C491-C499, C49A, C60, C62-C63, C66, C68-C70, C72-C75, C761-C678, C78-C80, C880-C883, C887-C889, C96-C97, D002, D015-D019, D074, D076, D091-D099, D37-D39, D40-D49, C7A, C7B	Cancer Other	
C81	Hodgkin's Disease	
D10-D36, D3A	Benign Neoplasms	
D50-D77	Blood Disorders	Other
D80-D89	Immune System Disorders	
Series beginning with "P excl. P50, P56 and P57" and A33	Newborn Conditions	
Series beginning with O, P50, P56, P57, Z37, Z38, Z39, Z3A, N960 and A34	Pregnancy Related	
Series beginning with "Z excluding Z37, Z38, Z39, Z57, Z58, Z683, Z684, Z714, Z715, Z77", T79, Y35-Y38, Y86, Y87, Y89, Y92-Y99, Series beginning with H excl. "H591-H594, H952-H955", U01-U03, E00-E07, E15-E35, E40-E64, E67-E68, E86, F64-F66, R60-R61, R631-R636, G08-G09, G47, I00, I02, I83-I86, J30, K00-K23, K35-K46, K60-K62, K64, M30-M32, N60-N65, R00-R03, R10-R15, R20-R23, R262, R263, R293-R297, R370, R40-R43, R47-R61 excl. R540, R631-R636, R638-R779, R79-R95, R97	Other	
E70-E89 excl. E860, E861, E869,	Metabolic Disorders	
F04-F09, F20-F63, F68-F69, F70-F79, F99, R44-R46, R630, R636	Mental Health Conditions	
F80-F98, R62	Disorders of Psychological Development	
R96, R99, U509	Unknown	
G10-G18, G23-G26, G31-G44, G50-G73, G89, G970-G972, G978-G979, R25, R260, R261, R268-R292, R298, R299, G90-G96, G98-G99	Neurological Conditions	
G20-G21	Parkinson's Disease	
G80-G83	Cerebral Palsy and other paralytic Syndromes	
U049, U070, J19, J27, J47-J94, J96-J99, R05-R09	Other Respiratory Conditions	
J45-J46	Asthma	

ICD10 Code	Cause of Death	Drivers[32]
K25-K31, K50-K59, K63, K65- K68, K80-K90, K92, R16-R19	Digestive Disorders	
K70-K77	Liver Disease	
R300-R369, R39, N130-N135, N137-N139, N201, N21-N23, N30-N40	Urological Conditions	
M00-M27, M33-M99	Musculo Skeletal	
N00-N12, N136, N14-N200, N202, N209, N25-N29	Kidney Disorder	
N41-N53	Diseases of Male Genital Organs	
N70-N95, N97-N99	Gynaecological Conditions	
Series beginning with Q	Congenital Malformation	
Series beginning with L excluding L76 and L80	Skin Diseases	
J00-J06, J22, J31-J32	Respiratory Infection	
J33-J39, R04	Respiratory Tract Conditions	

14   Appendix D – Driver Grouping

14.1   Introduction

The document Universe-of-Longevity-Catalysts-Call-For-Research.pdf (actuaries.org.uk) describes the original categorisation of 'drivers' which have been modified in this paper, as shown in the next sub-section.

14.2   Driver Grouping Modification

Table 21 – Driver Mapping Modifications

Driver	Original	Modifications/Additions
Non – Disease	N/a- previously grouped with "Other"	• Accident • Drug-Alcohol • External Causes • Injury • Medical Treatment Complications Natural Disaster • Suicide • Vehicular Accident
Cancer – Unmet need	• Lung cancer • Pancreatic cancer • Oesophageal cancer • Brain Tumours (including meningioma)	n/a
Cancer – Smoking Related	• Cancers of the mouth (including Pharynx & Larynx) • Nasal and paranasal sinus cancer • Bladder cancer • Cervical cancer • Acute myeloid leukaemia (AML) • Stomach cancer	Gynaecological Cancer
Cancer – Obesity Related	• Breast cancer • Colorectal cancer • Uterine cancer • Thyroid cancer • Gallbladder cancer • Myeloma • Kidney cancer • Liver cancer • Ovarian cancer	Bone Cancer
Cancer – Other	• Hodgkin's Lymphoma • Malignant Melanoma • Prostrate • Testicular • Acute lymphoblastic leukaemia (ALL) • Non-Hodgkin's Lymphoma	• Benign Neoplasms • Cancer Other • Hodgkin's Disease • Melanoma • Mesothelioma • Skin Cancer

Driver	Original	Modifications/Additions
Respiratory Disease – smoker related	• Chronic obstructive pulmonary disease (COPD) • Bronchial Disease • Emphysema	n/a
Alzheimer's Disease and Dementia	Alzheimer's Disease and Dementia	The ICD-code R54, which represents senility, frailty/old age has been categorised with this driver
Diabetes and Obesity	• Diabetes • Obesity	n/a
Cardiovascular Disease (Original driver categorisation splits CHD and Stroke into two separate groups)	Coronary Heart Disease (CHD): • Coronary heart disease • Peripheral arterial disease • Rheumatic heart disease Stroke: • Stroke • Transient Ischaemic Attack • Subarachnoid Haemorrhage	Coronary Heart Disease (CHD): • AMI • Aneurysm • Atherolsclerosis • Cardiac Arrest • Cardiomyopathy • Heart Failure • Inflammation of the Heart • Ischaemic Heart Diseases • Other CHD • Phlebitit and Thrombophlebitis • Pulmonary Heart Disease • Valve Disorder Stroke: • Cerebral Infarction • Intracranial Haemorrhage • Other Cerebrovascular Disease • Stroke Other
Infection	N/a- previously lumped with "Other"	• Antibiotic Resistance • Infection • Influenza and Pneumonia
COVID	N/a- previously lumped with "Other"	COVID-19
Other	• Parkinson's Disease • Liver disease • Pancreatitis • …	Captures all other cause of deaths
Co-morbidity	Defined as old age/ frailty in original document	Redefined as where a death has three or more mentions that map to either: • A cancer mention (multiple cancer mentions counted as one) • Alzheimer's Disease and Dementia (excluding frailty) • Diabetes and Obesity • Coronary Heart Disease excluding heart failure and cardiac arrest. • All respiratory diseases/ conditions

Driver	Original	Modifications/Additions
		• Distinct mention of Other (excluding pregnancy-related and new-born cause) • Infection excluding COVID/ influenza and Pneumonia

14.3 Granular Driver used in mapping rules.

Driver	Granular Driver
Non – Disease	• Non – Disease excluding Medical Treatment Complication • Medical Treatment Complication
Cancer – Unmet need	Cancer – Unmet need
Cancer – Smoking Related	Cancer – Smoking Related
Cancer – Obesity Related	Cancer – Obesity Related
Cancer – Other	• Cancer – Other excluding Prostate Cancer and Benign Neoplasms • Prostate Cancer • Benign Neoplasms
Respiratory Disease – smoker related	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	• Alzheimer's Disease and Dementia • Senility/Frailty
Diabetes and Obesity	Diabetes and Obesity
Cardiovascular Disease	• Coronary heart disease • Stroke • Heart Failure • AMI • Cardiac Arrest • Other CHD • Pulmonary Heart Disease
Infection	• Infection excluding Influenza and Pneumonia • Influenza and Pneumonia
COVID	COVID
Other	• Blood Disorders • Digestive Disorders • Diseases of Male Genital Organs • Gynaecological Conditions • Immune System Disorders • Kidney Disorder • Liver Disease • Metabolic Disorders • Skin Diseases • Urological Conditions • Respiratory Conditions i.e. Asthma, Respiratory Tract Conditions, Respiratory Infection and Other Respiratory Conditions • Cerebral Palsy and other paralytic Syndromes, Neurological Conditions and Parkinson's Disease

Driver	Granular Driver
	• Congenital Malformation
	• Disorders of Psychological Development
	• Mental Health Conditions
	• Musculo Skeletal
	• Other
	• Newborn Conditions and Pregnancy Related
	• Unknown

15   Appendix E – Age Banding

15.1   Introduction

This section covers the driver specific age bands used in processing the data. The band "Lowerband" is used to catch deaths at younger ages where deaths fall below 3 after the X4 and X9 bandings are applied.

15.2   Driver Specific Age Bands

Table 22 – Driver Specific Age Bands

Driver	Age band
Alzheimer's Disease and Dementia	• LowerBand • 55-59
Cancer – Obesity Related, Cancer – Other, Cardiovascular Disease	• LowerBand • 20-29 • 30-34 • 35-39 • 40-44 • 45-49 • 50-54 55-59
Cancer – Smoking Related, Cancer – Unmet need, Infection	• LowerBand • 20-29 • 30-39 • 40-44 • 45-49 • 50-54 • 55-59
Co – Morbidity, Covid-19, Non – Disease and Other	• Lowerband • 20-24 • 25-29 • 30-34 • 35-39 • 40-44 • 45-49 • 50-54 • 55-59
Diabetes and Obesity	• LowerBand • 30-39 • 40-49 • 50-54 • 55-59
Respiratory Disease – smoker related	• LowerBand • 40-49 • 50-54 • 55-59

16   Appendix F – Redistribution of grouped data into single-year age bands

16.1   Introduction

This appendix details the transformation of the ONS output dataset into a version shown by SYOA and covers the following elements:

- Need for Transformations
- Approach

16.2   Need for Transformation

The cause of death data released by the Working Party has been extracted from data sourced from the ONS. To comply with restrictions placed by the ONS, the data released is grouped when individual combinations of number of deaths in a year, from a given cause for a given sex fell below three deaths.

In order to extract as much data as possible, the Working Party chose to apply different groupings for different causes, years and sexes. For example, if "Cancer – Unmet need" fell below three deaths for males in 2015 at age [105], then the final group is [over 105]. If in 2016, this didn't fall below three deaths until age [106], then the final group is [over 106].

Whilst this methodology has allowed the Working Party to extract as granular data as it was permitted to do so, it has created a dataset that requires some expertise to manipulate (given groupings are inconsistent between causes, years etc).

To provide a dataset that is easier to manipulate, the Working Party has provided a second dataset "IFoA Drivers of Mortality Dataset SYOA[33]". This dataset contains redistributed deaths from grouped bands to single year ages. The Working Party selected a method to do this such that it is simple and transparent, which is set out below. Note that:

- The split at younger ages do not reflect a best estimate view, but are a consequence of the method, which appears to work well at older ages.
- These ages are not material for financial services products and thus not worth delaying release to pursue a more complex approach to improve the results at those ages.
- An alternative is not publishing these ages, but it is helpful to include them so that totals can be reconciled.

The Working Party would encourage users who are using the dataset for studies at high or low ages to consider the reasonableness of our approach or whether they wish to redistribute the bands using an alternative method.

16.3   Approach

The Working Party's redistribution method can be summarised as follows:

- For data that has been banded with an upper and lower limit (e.g. ages 55-59) from minimum age, n_{min}, to maximum age, n_{max}, deaths have been evenly distributed over the period as per Equation 1.

---

[33] SYOA stands for Single Year of Age

Equation 1 – Spreading of deaths between an age band

Percentage allocated to age, n, where $n_{max} \geq n \geq n_{min}$ (0 otherwise):

$$\frac{1}{n_{max} - n_{min} + 1}$$

- For data that has been banded with a lower limit only (e.g.≥105), deaths have been distributed linearly over a 5-year period from the minimum age, n_{min}, as per Equation 2.

Equation 2 – Spreading of deaths banded with a lower limit

Percentage allocated to age, n, where $n \geq n_{min}$ (0 otherwise):

$$\max\left(0, \frac{5 - n + n_{min}}{15}\right)$$

- For data that has been banded with an upper limit only (e.g.<60), deaths have been distributed linearly over a 5-year period to the maximum age, n_{max}, as per Equation 3.

Equation 3 – Spreading of deaths banded with an upper limit

Percentage allocated to age, n, where $n < n_{max}$ (0 otherwise):

$$\max\left(0, \frac{6 - n_{max} + n}{15}\right)$$

17 Appendix G – Movement between formulations

17.1 Introduction

This appendix shows the movement in deaths between different formulations of the drivers.

17.2 ONS Primary vs. Proposed Mapping

ONS Primary	Proposed Mapping Alzheimer's Disease and Dementia	Cancer – Obesity Related	Cancer – Other	Cancer – Smoking Related	Cancer – Unmet need	Cardiovascular Disease	Co – Morbidity	Covid-19	Diabetes and Obesity	Infection	Non – Disease	Other	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	844,759	5,519	2,686	1,961	665	22	153,386	3,175	4		9	13,424	24
Cancer – Obesity Related		786,020	16	3,521	1,443	3	93,054	1,314	<3		3	18	4
Cancer – Other	4,251	9,093	628,372	9,737	5,412	11,925	106,148	1,333	2,261	738	61	17,614	2,120
Cancer – Smoking Related		289	24	298,436	2,180	<3	48,043	566		<3	<3	12	<3
Cancer – Unmet need		<3	40	19	900,861	<3	89,215	1,377				5	
Cardiovascular Disease	179,579	30,902	24,114	14,720	18,149	2,143,652	397,744	5,346	170,486	21,675	11,717	180,273	125,231
Covid-19								141,135					
Diabetes and Obesity	4,117	296	129	116	76	16	69,201	434	55,070		38	17	588
Infection	52,344	1,812	2,522	1,156	1,069	108,692	87,519	198	29,524	256,718	5,026	153,166	4,175
Non – Disease	12,599	1,759	2,160	1,293	1,386	8,097	29,557	828	1,270	432	340,939	14,425	1,941
Other	59,829	9,465	8,286	4,492	6,098	2,227	497,776	3,446	17,576	6	12,360	1,020,975	32,417
Respiratory Disease – smoker related	257	3,995	3,559	3,218	6,516	13	94,327	395			324	5	426,972

17.3 ONS Primary vs. Proposed Mapping (No COVID/ No Co-Morbidity)

ONS Primary	Proposed Mapping (No COVID/ Co-Morbidity) Alzheimer's Disease and Dementia	Cancer – Obesity Related	Cancer – Other	Cancer – Smoking Related	Cancer – Unmet need	Cardiovascular Disease	Diabetes and Obesity	Infection	Non – Disease	Other	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	988,908	11,070	6,133	4,326	1,550	38	52		12	13,486	59
Cancer – Obesity Related	<3	879,372	29	4,231	1,693	10	14		4	38	5
Cancer – Other	7,855	10,655	704,981	11,622	6,028	15,705	7,421	745	76	29,027	4,950
Cancer – Smoking Related		314	42	346,664	2,478	6	9	<3	<3	33	6
Cancer – Unmet need		3	49	20	991,432	<3			<3	10	4
Cardiovascular Disease	239,356	48,076	38,488	23,908	25,872	2,256,151	272,704	21,754	12,611	203,590	181,078
Covid-19	29,981	3,706	3,864	2,615	2,776	17,965	21,112	14,842	1,835	27,646	14,793
Diabetes and Obesity	14,146	2,284	1,408	972	652	35	105,143		135	92	5,231
Infection	60,107	4,189	4,900	2,351	1,963	120,210	52,319	256,749	5,932	185,765	9,436
Non – Disease	17,848	2,941	3,215	1,996	1,974	9,407	2,904	433	355,083	17,714	3,171
Other	127,269	26,259	23,512	12,351	13,040	3,222	93,984	6	14,228	1,286,871	74,211
Respiratory Disease – smoker related	1,209	7,236	6,646	5,840	10,281	20	23		493	36	507,797

17.4 Proposed Mapping (No COVID/ No Co-Morbidity) vs. Proposed Mapping

Proposed Mapping (No COVID/ Co-Morbidity)	Proposed Mapping Alzheimer's Disease and Dementia	Cancer – Obesity Related	Cancer – Other	Cancer – Smoking Related	Cancer – Unmet need	Cardiovascular Disease	Co – Morbidity	Covid-19	Diabetes and Obesity	Infection	Non – Disease	Other	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	1,157,735						294,931	34,014					
Cancer – Obesity Related		849,152					141,746	5,207					
Cancer – Other			671,908				116,221	5,138					
Cancer – Smoking Related				338,669			74,913	3,314					
Cancer – Unmet need					943,855		111,611	4,273					
Cardiovascular Disease						2,274,650	127,442	20,678					
Diabetes and Obesity							256,589	22,904	276,192				
Infection							14,960			279,570			
Non – Disease							17,475	2,459			370,478		
Other							333,633	30,741				1,399,934	
Respiratory Disease – smoker related							191,409	15,859					593,473

17.5 Proposed Mapping (No COVID/ No Co-Morbidity) vs. Proposed Mapping (No COVID)

Proposed Mapping (No COVID/ Co-Morbidity)	**Proposed Mapping (No COVID)**											
	Alzheimer's Disease and Dementia	Cancer – Obesity Related	Cancer – Other	Cancer – Smoking Related	Cancer – Unmet need	Cardiovascular Disease	Co – Morbidity	Diabetes and Obesity	Infection	Non – Disease	Other	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	1,182,244						304,436					
Cancer – Obesity Related		852,510					143,595					
Cancer – Other			675,320				117,947					
Cancer – Smoking Related				340,700			76,196					
Cancer – Unmet need					946,964		112,775					
Cardiovascular Disease						2,292,832	129,938					
Diabetes and Obesity							267,084	288,601				
Infection									294,530			
Non – Disease							18,065			372,347		
Other							338,196				1,426,112	
Respiratory Disease – smoker related							196,918					603,823

17.6 Proposed Mapping (No COVID/ No Co-Morbidity) vs. Proposed Mapping (No Co-morbidity)

Proposed Mapping (No COVID/ Co-Morbidity)	**Proposed Mapping (No Co-Morbidity)**											
	Alzheimer's Disease and Dementia	Cancer – Obesity Related	Cancer – Other	Cancer – Smoking Related	Cancer – Unmet need	Cardiovascular Disease	Covid-19	Diabetes and Obesity	Infection	Non – Disease	Other	Respiratory Disease – smoker related
Alzheimer's Disease and Dementia	1,452,666						34,014					
Cancer – Obesity Related		990,898					5,207					
Cancer – Other			788,129				5,138					
Cancer – Smoking Related				413,582			3,314					
Cancer – Unmet need					1,055,466		4,273					
Cardiovascular Disease						2,402,092	20,678					
Diabetes and Obesity							22,904	532,781				
Infection							14,960		279,570			
Non – Disease							2,459			387,953		
Other							30,741				1,733,567	
Respiratory Disease – smoker related							15,859					784,882