An Phríomh-Oifig Staidrimh
Central Statistics Office

# Implementation of IRIS software for the automated coding of deaths in Ireland

# Background

The production of cause of mortality statistics requires the selection of an underlying cause of death from death certificates. This underlying cause of death is defined as (a) the disease or injury which initiated the train of morbid events leading directly to death, or (b) the circumstances of the accident or violence which produced the fatal injury. As part of this process the CSO used to use the US Centre of Disease Control-developed Medical Mortality Data System (MMDS) suite of software but switched to using IRIS from 2018 onwards.

While not altering the classifications per se, a change in coding software can affect the assignment of the underlying cause of death codes. Therefore, as part of the implementation of IRIS, the CSO dual-coded the 2015 mortality data to see the effects of the change in software.

# What is IRIS?

IRIS is a European Union-led software package initially using components of the MMDS software. IRIS is the software tool preferred by Eurostat and it is used by many of the EU Member States. It was developed by a core group of European countries. *See link to IRIS website:* https://www.dimdi.de/dynamic/en/classifications/iris-institute/index.html

When processing a death certificate, IRIS will first try to code all the diagnostic expressions on a death certificate (of which there may be several) and then select the underlying cause of death according to the rules and guidelines published by the WHO in the ICD-10 classification. *See link to ICD-10 classification:* http://apps.who.int/classifications/icd10/browse/2016/en

IRIS will try to automatically code as many death certificates as possible. Following this automated process, coders can edit the text of the certificate to facilitate further processing or select the underlying cause of death manually.

IRIS requires manual intervention by coders for post-procedural disorders (E89, G97, H59, H95, I97, K91, M96, N99), pregnancy, childbirth and the puerperium (O00-O99), accidental poisoning (X40-X49), complications of medical and surgical care (Y40-Y84) and sequelae of medical and surgical care (Y880-Y883).

# Why move to IRIS?

The move to IRIS will lead to improvements in cause of death coding:

- IRIS is the most up-to-date software available for coding as MMDS is no longer being supported by the US Centre for Disease Control who have moved to join the IRIS development core group.
- IRIS software contains language-dependent tables that can be developed to suit individual jurisdictions e.g. we can include common phrases used in Irish death certificates which may not be included in the standard English dictionary released with the IRIS package.
- The IRIS user interface is more fluid, intuitive and user-friendly than the older MMDS system.
- The IRIS MUSE component provides detailed explanations on how the system arrived at the multiple and underlying cause codes, showing what WHO rules and instructions, it applied on a separate interface. This was not the case with MMDS.
- Increased automation in coding means less manual intervention required by coders thus mitigating human misinterpretation and error. The international rules and instructions for the selection of the underlying cause of death leaves space for individual interpretation, and can result in a certain degree of variability of the tabulated underlying code among coders (Harteloh et al., 2010).
  *See link:* https://medwinpublishers.com/EIJ/EIJ16000102.pdf

# Bridging study

The CSO used IRIS to select cause of death codes for the 2015 mortality records which were then compared to the original code assigned when using MMDS (the coding done under both MMDS and IRIS were obviously independent of each other). It is important to note that stillbirths and deaths identified as having external causes of death (suicides, homicides and accidents) continue to be coded manually by mortality coders.

There were 29,952 deaths registered in 2015 and each was attributed an underlying cause of death code in line with the WHO ICD-10 statistical classification of diseases and related health problems. Having removed unnatural deaths, a total of 27,502 mortality records were then processed in IRIS software, having previously been processed using MMDS. The CSO used the 2017 IRIS decision tables in the automatic coding process (55% of all records) and manually coded the remaining records (45%) using 2015 decision tables. Under MMDS the 2015 mortality data was coded using the 2011 decision tables and the uncoded mortality records were coded using the 2014 decision tables (which were the most up to date version available at that time).

## Comparing the MMDS and IRIS coding:

- 22,920 or 83.3% of records were attributed the exact same underlying cause of death code (UCOD) at four-digit ICD-10 level
- 24,509 or 89.1% had the same UCOD at three-digit level
- 25,262 or 91.8% had the same UCOD at two-digit level
- 26,158 or 95.1% had the same UCOD at one-digit level
- 1,344 or 4.9% had a different UCOD code in IRIS

The Comparability Ratio (CR) is defined as the number of deaths coded under MMDS for a particular ICD code divided by the number coded under IRIS. Obviously the closer this ratio is to one then the less change there is in the total number of deaths for that ICD code under IRIS compared to MMDS.

**Table 1 Comparability Rate (MMDS V's IRIS) 2015**

| ICD_10 | MMDS | IRIS | Comparability rate |
|---|---|---|---|
| **A** | 281 | 233 | 1.21 |
| **B** | 47 | 37 | 1.27 |
| **C** | 8,597 | 8,579 | 1.00 |
| **D** | 318 | 354 | 0.90 |
| **E** | 670 | 669 | 1.00 |
| **F** | 1,501 | 1,535 | 0.98 |
| **G** | 1,503 | 1,603 | 0.94 |
| **I** | 8,859 | 8,757 | 1.01 |
| **J** | 3,649 | 3,713 | 0.98 |
| **K** | 890 | 911 | 0.98 |
| **L** | 66 | 59 | 1.12 |
| **M** | 202 | 207 | 0.98 |
| **N** | 589 | 512 | 1.15 |
| **O** | 1 | 1 | 1.00 |
| **P** | 86 | 81 | 1.06 |
| **Q** | 155 | 171 | 0.91 |
| **R** | 88 | 80 | 1.10 |
| **Total** | 27,502 | 27,502 | 1.00 |

Per Table 1 above, the CRs show there is little or no difference in the main causes of death groups i.e., 'I' (diseases of the circulatory system), 'C' (neoplasms) and 'J' (diseases of the respiratory system) which have a CR of 1.01, 1.00 and 0.98 respectively.

The groups where the CR is further from one e.g., 'A' (certain infectious and parasitic diseases), which has a comparability rate of 1.21, are calculated on a much smaller number of records (e.g. Group A accounts for 1% of death records).

Looking in more detail at the differences between the two coding systems, Table 2 shows the changes between MMDS and IRIS at 1-digit level. For example there were 96.3% of records given a 'I' code (i.e. diseases of the circulatory system) both in MMDS and IRIS software while a further 1.7% of records that had been given an 'I' code in MMDS were attributed a 'J' code (i.e. diseases of the respiratory system) in IRIS.

## Table 2. Underlying cause of death by ICD-10 chapter MMDS(vertical) v Iris (horizontal)

| | IRIS | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **MMDS** | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **I** | **J** | **K** | **L** | **M** | **N** | **O** | **P** | **Q** | **R** | **Total (%)** | **No.** |
| **A** | **61.6** | 0.0 | 1.4 | 0.0 | 0.0 | 1.1 | 1.8 | 4.3 | 23.5 | 3.9 | 0.0 | 1.1 | 1.1 | 0.0 | 0.4 | 0.0 | 0.0 | 100 | 281 |
| **B** | 8.5 | **59.6** | 2.1 | 4.3 | 6.4 | 2.1 | 0.0 | 0.0 | 14.9 | 2.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 47 |
| **C** | 0.1 | 0.0 | **98.6** | 0.5 | 0.0 | 0.0 | 0.1 | 0.2 | 0.3 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 8,597 |
| **D** | 0.6 | 0.0 | 5.7 | **89.3** | 0.3 | 0.0 | 0.3 | 0.6 | 1.9 | 0.0 | 0.0 | 0.0 | 0.9 | 0.0 | 0.0 | 0.3 | 0.0 | 100 | 318 |
| **E** | 0.1 | 0.0 | 0.1 | 0.0 | **92.1** | 0.6 | 0.3 | 4.0 | 1.2 | 0.6 | 0.0 | 0.1 | 0.7 | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 670 |
| **F** | 0.1 | 0.0 | 0.2 | 0.1 | 0.1 | **94.9** | 1.9 | 1.5 | 0.5 | 0.2 | 0.0 | 0.1 | 0.1 | 0.0 | 0.0 | 0.3 | 0.0 | 100 | 1,501 |
| **G** | 0.1 | 0.1 | 0.0 | 0.1 | 0.2 | 0.9 | **96.1** | 0.5 | 0.7 | 0.1 | 0.1 | 0.2 | 0.1 | 0.0 | 0.0 | 0.9 | 0.1 | 100 | 1,503 |
| **I** | 0.1 | 0.0 | 0.2 | 0.1 | 0.2 | 0.3 | 0.3 | **96.3** | 1.7 | 0.2 | 0.0 | 0.1 | 0.2 | 0.0 | 0.0 | 0.1 | 0.0 | 100 | 8,859 |
| **J** | 0.5 | 0.1 | 0.7 | 0.1 | 0.2 | 1.5 | 2.0 | 0.9 | **92.6** | 0.5 | 0.0 | 0.5 | 0.2 | 0.0 | 0.0 | 0.1 | 0.0 | 100 | 3,649 |
| **K** | 0.2 | 0.3 | 1.1 | 0.4 | 0.6 | 0.1 | 0.3 | 1.8 | 1.0 | **93.1** | 0.1 | 0.1 | 0.7 | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 890 |
| **L** | 6.1 | 0.0 | 0.0 | 1.5 | 0.0 | 0.0 | 0.0 | 1.5 | 0.0 | 1.5 | **86.4** | 3.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 66 |
| **M** | 1.0 | 0.0 | 1.0 | 1.0 | 2.5 | 0.0 | 2.5 | 4.0 | 5.4 | 0.5 | 0.0 | **80.7** | 1.5 | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 202 |
| **N** | 1.0 | 0.0 | 1.4 | 0.2 | 1.2 | 0.7 | 0.7 | 12.1 | 3.4 | 0.7 | 0.0 | 0.8 | **77.9** | 0.0 | 0.0 | 0.0 | 0.0 | 100 | 589 |
| **O** | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | **100.0** | 0.0 | 0.0 | 0.0 | 100 | 1 |
| **P** | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.3 | 1.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | **89.5** | 7.0 | 0.0 | 100 | 86 |
| **Q** | 0.0 | 0.0 | 1.3 | 0.0 | 0.0 | 0.0 | 1.9 | 3.9 | 1.9 | 1.3 | 0.0 | 0.0 | 0.0 | 0.0 | 1.3 | **88.4** | 0.0 | 100 | 155 |
| **R** | 1.1 | 0.0 | 2.3 | 0.0 | 0.0 | 0.0 | 2.3 | 4.5 | 3.4 | 1.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | **85.2** | 100 | 88 |
| **Total** | | | | | | | | | | | | | | | | | | | 27,502 |

The primary analysis conducted by the CSO was concerned with the confirmation that the underlying cause of death code attributed by the IRIS software was correct rather than ascertaining the reason why it differed from that derived using MMDS. This was facilitated by the fact that IRIS, as part of its process, provides the mortality coder with an explanation (in text format) of how the WHO mortality coding rules are applied to each record.

As can be seen in the table above diseases of the circulatory system (8,859), neoplasms (8,597) and diseases of the respiratory system (3,649) accounted for the main causes of death in both systems and there was a high correlation between them with 96.3%, 98.6% and 92.6% cases respectively having the same selected underlying cause of death code at 1digit level.

Looking at some of the cases where there was a difference between both coding software packages, there were 23.5% (66 records) that were coded to 'A' (certain infestious and parasitic diseases) in MMDS and were assigned a 'J' code (diseases of the respiratory systems) in IRIS. Many of these were coded to 'Other specified respiratory disorders' (J988) due to the non-application of a WHO update of the term infection. This update was effective from the 1st January 2010 but had not been applied by coders to rejected MMDS records. This has now been corrected in the IRIS coding process and data quality checks are in place.

There were 7% (of 86 records) that were allocated a 'P' code in MMDS from the Chapter 'Certain conditions originating in the perinatal period' (P00-P96) which includes conditions that have their origin in the perinatal period even though death occurs later. These 7% were allocated a 'Q' code in IRIS. The exclusion note at the beginning of this chapter gives priority of assignment to Congenital malformations, deformations and chromosomal abnormalities (Q00-Q99) over conditions in the perinatal period. The selection of an underlying cause of death with perinatal mortality can be difficult to do due to certifiers not conforming to the hierarchy of the death certificate and general underreporting of conditions, particularly in relation to stillbirths.

There were also general reasons why there might be differences between MMDS and IRIS coding. Firstly, mortality coding is very complex and not all WHO ICD-10 updates were applied by the mortality coders in respect of manual coding in MMDS whereas they were applied when using IRIS. The IRIS software includes current WHO ICD-10 updates while the final update for international users of MMDS software was in 2011 and although updated decision tables could be referred to by coders after this, non-implementation of these decision tables in the production software meant a greater reliance on coder knowledge. All automated coding from 2011 used decision tables from that year without WHO updates included in the software. Edit checks were put in place to capture major updates by the WHO but not all updates were captured.

Secondly, the different levels of coder knowledge and training when using MMDS versus IRIS also had an impact on the results. In the last few years the CSO has initiated internal training workshops and we have also engaged international experts\trainers that have come on-site to upskill our coders and our senior coders have participated in training programmes abroad. In addition, senior coders engage with experts in the International Mortality Reference group to discuss and resolve complex cases. All this has led to increased coder knowledge over time.

Thirdly, there is a higher level of automated coding for IRIS as compared to MMDS and this may have led to situations where the IRIS system selected the mortality code in a more clinical and harmonised way than occurred using MMDS when coders selected the code themselves. In general, the IRIS coding system is an improvement due to increased automated coding and less reliance on manual intervention thus mitigating coding error.

Use of IRIS will allow us to focus on data quality and improving our processes. IRIS is constantly evolving, driven by worldwide user testing and is in line with WHO classification updates.