CRVS technical guide

Action guide on implementing automated coding (Iris)

This CRVS action guide describes seven key steps that countries should follow when implementing automated mortality coding.

1. Information gathering and process mapping
2. Mapping of Iris processes
3. Data dictionary and standardisation tables
4. Training
5. Rejected records
6. Sustainability
7. Evaluation

Background

Reliable knowledge on the mortality and causes of death in a population are critical for policy making. Ideally, analyses are based on the **underlying cause of death**,[1] which is recorded on the medical certificate of cause of death. However, it is estimated that 140 countries with 80% of the world's population do not have reliable cause of death data.

Mortality coding is a complex process by which all diseases and conditions recorded on a death certificate are transformed from text to alpha-numeric codes, following strict procedures as set out by the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10).[2] To be able to do this correctly **mortality coders have to be well trained in ICD-10 rules and regulations**.

Box 1: What are ICD-10 codes?

The International Statistical Classification of Diseases and Related Health Problems, 10th Revision, or ICD-10, is a statistical classification that groups similar diseases into mutually exclusive categories by translating text (cause of death) into alphanumeric codes. Coding is done for the purposes of storage, retrieval and analysis of COD data.

Using coding software to automate the coding process can improve the quality and consistency of data by applying standard coding rules. Additionally, because the selection of the underlying cause of death in automated coding is instantaneous, the process is more efficient and data dissemination can be timelier.

To improve the quality of coding, many countries now use the Iris software for coding multiple causes of death and for selecting the underlying cause of death.[3]

Iris is based on the International Form of Medical Certificate of Cause of Death provided by the World Health Organization (WHO) in Volume 2 of ICD-10. Iris was developed by Sweden and France and is now managed by the Iris institute[4] in Cologne which promotes its use and has a user forum that aims to further improve the utility of this language independent software.

1 The underlying cause of death is 'the disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury' (WHO 1994).

2 World Health Organization. International statistical classification of diseases and related health problems. 10th revision, 10th edition. Geneva, Switzerland: WHO; 2016.

3 Prieto JC, García MRG. Iris: International automatic coding system of causes of death. Its use in the Spanish mortality statistics. *Boletin de Estadistica e Investigacion Operativa* 2016; 32:130-147.

4 Iris Automated Coding System for Causes of Death: User's Reference Manual (Iris version V5.4.0S1). Cologne: IRIS Institute, German Institute of Medical Documentation and Information; 2017.

Box 2: What is Iris?

Iris is an automated coding software that allows death certificates to be coded according to ICD-10 rules and standards. Iris has been implemented in a number of European Union countries, as well as by the Office for National Statistics in the United Kingdom and by Statistics Canada. In the Asia-Pacific region it is used by the Australian Bureau of Statistics and the Fijian Statistical Office. The Philippines is the first country in Asia to implement Iris as part of the BD4H initiative.

Box 3: What is process mapping?

Process mapping is one of the tools used in enterprise architecture to describe and analyse the business architecture of a system. It is a systematic and standardised approach that civil registration and vital statistics (CRVS) stakeholders can use to understand, analyse and optimise processes within complex systems, to achieve intended system goals. A process map is a visual snapshot of the end-to-end activities, stakeholders and requirements of a CRVS system.[7]

Action guide – key tasks and challenges

1 2 3 4 5 6 7

Step 1: Information gathering and process mapping

Prior to implementing Iris, it is recommended to do a process map of the mortality data collection system (see **Box 3**) and to gather all the relevant background information on current certification and coding practices. Types of information to gather include:

- The government departments involved in the data collection, for coding the data and responsible for producing vital statistics;
- Where can the data be accessed and in what format they are stored;
- Degree of computerisation of the data;
- Who codes the death certificates and where (i.e. national, regional, local hospitals), and;
- Revision of ICD and language used
- Training that is available to coders.

Iris is a free, easy to install software[5] however, before implementation a set of key requirements must be in place[6] (**Box 4**).

1 2 3 4 5 6 7

Step 2: Mapping of Iris processes

Based on the process map, further develop the processes specific to implement Iris to determine:

- How the death certificates will be input into Iris (code entry or text entry mode);
- How and when the Iris batch processing software will be run;
- How to notify Iris which data dictionary to use (adults, fetal or neonatal);
- If the software should perform other calculations, such as the date of death of the decedent, and what other information is needed to do so;
- How rejected records will be handled or resolved, and;
- How the coded files from Iris will be imported back into the national civil registration and vital statistics (CRVS) system. Generally, an interface software program is designed in advance to perform these tasks, but it must be tested until the system flow is smooth and complete.

5 https://www.dimdi.de/dynamic/en/classifications/iris-institute/index.html

6 Mikkelsen L, Gamage S, Sarmiento C, Reolalas A, Orcilla E. Introduction of Iris automated mortality coding system in the Philippines. CRVS country perspectives series. Melbourne, Australia: Bloomberg Philanthropies Data for Health Initiative, Civil Registration and Vital Statistics Improvement, University of Melbourne; 2019

7 de Savigny D, Cobos Muñoz D. *Understanding CRVS systems: The importance of process mapping.* CRVS development series. Melbourne, Australia: Bloomberg Philanthropies Data for Health Initiative, Civil Registration and Vital Statistics Improvement, the University of Melbourne; 2017.

Box 4: Basic requirements for Iris implementation

- Use of the WHO International Form of Medical Certificate of Cause of Death;
- Use of the ICD-10 for mortality coding.
- Enough trained mortality coders (with access to ongoing training);
- Use of existing Iris dictionary or ability to develop one in local language, and;
- Local information technology (IT) support is available for implementation, maintenance, and development of interfacing computer programs.

1 2 3 4 5 6 7

Step 3: Data dictionary and standardisation tables

In Iris, a dictionary table gives the correspondence between a textual expression of a diagnosis and an ICD- 10 code.

A country-specific data dictionary needs to be created that assigns the corresponding ICD-10 code to the text-based cause of death - written in the country's official language.

English, French, German and Swedish versions of the dictionary have been developed and may be used freely. However, if one of these dictionaries is adopted, it is still necessary to complement it with additional and local terms used (text expressions) to describe causes of death, which will likely vary between the two countries.

It is strongly recommended that countries use and adapt Iris dictionaries that have already been developed by other countries using the same language. Iris has an iteration process whereby each rejection of an individual death record informs the development of the local dictionary so that once rectified, the dictionary can be expanded and tailored to the local context.

It is important that the dictionary remains manageable in size to make it easy to update and keep consistent. Mapping of local terms to standard expressions can be done manually or automatically. Iris proposes a way through embedded standardisation tables to do it automatically. The update of the standardisation tables to include common local diagnostic expressions must be done very carefully, in consultation with clinicians and professional bodies of the clinicians.

1 2 3 4 5 6 7

Step 4: Training

Coding staff must be trained in the use of Iris. Training is also useful for the IT staff who will be involved in installation, development of utility and interface programs, maintenance of standardisation and dictionary tables etc.

The suggested topics to be covered in the training are:

a. Introduction;
b. Installation guide;
c. How to use Iris;
d. Menus and tools;
e. Mortality medical data set (MMDS) in Iris;
f. Data bases (table and certificate);
g. Standardisation in Iris;
h. Table maintenance, and;
i. Practice exercises with several batches of death certificates including neonatal, child and adult deaths.

All coders should have basic ICD knowledge, and if there are no experienced coders, some should be given training in more advanced ICD coding.

1 2 3 4 5 6 7

Step 5: Rejected records

As part of the implementation, testing and improvement of Iris, it is common for many individual death certificates to be rejected in the beginning. What is important is that countries address the underlying problem causing rejection and have a system in place to manage such records. Common reasons for rejected death records include:

- Non-standard expressions for causes of death: this includes abbreviations and data entry errors when the record was created.
- Interval reasons: when several diseases are reported on the death certificate, but the duration of each disease is not stated, or the duration for one or more diseases is missing, or it is unclear to which diseases the duration listed applies to.

- Main injury missing in countries that require it: for example, if the death was due to a skull fracture from a road traffic accident, the underlying cause of death would be the road traffic accident. If the skull fracture is not also listed on the medical certificate as the main injury, the record will be rejected by Iris.
- Missing cause of death: if the dictionary is adapted from another country, but is not updated to local epidemiological circumstances, the causes of death listed in the data dictionary might not suffice.

Although some of the reasons Iris rejects individual death records will be resolved by updating the country data dictionary, there will always be other rejected records that will require experienced manual coders to revise them.

Step 6: Sustainability

To maintain the quality of automated coding it is important to retain some experienced coders. They are needed not only to resolve rejected records and perform periodic quality controls, but also to train new staff in use of Iris and basic ICD mortality rules.

A small committee comprising the senior coding and statistical staff should take the responsibility for managing the dictionary and standardization tables. The committee should also be responsible for creating local business rules for coding operations.

Step 7: Evaluation

Before switching over to fully automated coding with Iris, it is useful to double-code a certain amount of records (10-20%) to compare results and ensure that all is working correctly. After switching to automated coding, an in-depth evaluation of the complete annual dataset should be conducted to understand any significant differences between the Iris coded data and the manually coded data from the previous three years. This is essential to understand how the automated coding impacted the data. A technical note explaining any major change and inconsistency in the new data should be written to accompany the dissemination of the data.

Annual evaluations of the coded and compiled data should continue to be performed, focussing on;

- Identifying the typology of rejects and detection of common issues and areas for improvement;
- Comparing time series of statistical outputs for standard tabulation tables;
- Manual coding of a sample of auto-coded records to test the coding rules;
- Examining the trends in leading causes of deaths to check for inconsistencies.

Summary

Mortality coding can be a complex, time-consuming, and costly process. To address this, many countries are now using automated software to generate more reliable and timely cause of death data. One such software is Iris, which allows users to either enter ICD-10 codes, or to enter the causes in free text to then be assigned codes and an underlying cause of death.

There are a number of steps countries should take before implementing Iris, however. A complete system mapping and design process should be completed, along with the development of a country-specific data dictionary that maps causes of death in the language the local death certificates are completed in, to their corresponding ICD-10 codes. Additionally, countries will need to formulate a plan for dealing with rejected individual death records, and to correct some of the system processes that are causing these. The reward of using Iris is a coding system fully standardised according to ICD rules and regulations, that, as a result of efficiency gains, will significantly reduce coding time and dissemination of mortality data to users.

Bloomberg Philanthropies DATA FOR HEALTH INITIATIVE

Australian Government
Department of Foreign Affairs and Trade

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

Civil Registration and Vital Statistics partners:

Vital Strategies

CDC Foundation
Together our impact is greater

THE UNIVERSITY OF MELBOURNE

For more information contact:

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

CRICOS Provider Code: 00116K

Version: 0418-02