

RESEARCH

Open Access



Applying an ICD-10 to ICD-11 mapping tool to identify causes of death codes in an Alberta dataset

Chelsea Doktorchik¹, Danielle A. Southern^{2*}, James A. King^{2,3} and Hude Quan^{2,4}

Abstract

Background The most recent and 11th revision of the International Classification of Disease (ICD-11) is in use as of January 2022, and countries around the globe are now preparing for the implementation of ICD-11 and transition from the 10th revision (ICD-10). Translation of current coding is required for historical comparisons.

Methods We applied the World Health Organization (WHO) mapping tables to current Centers for Disease Control and Prevention (CDC) Lists of ICD-10 coding of underlying causes of death to assess what ICD-11 codes look like in an Alberta sample of causes of death (COD). We prepared frequency tables for a single year of COD in Alberta based on the CDC grouping of COD.

Results The mapping success rate at the ICD-10 code level for the adult population (> 18 years) was 96.6% and 100% for children (1–17 years) and infants (< 1 year). The mapping success rate by patient was 99.5% for the adult population patient deaths and 100% for children and infants. We mapped ICD-11 codes to identify the ten most frequently reported underlying COD in Alberta for 24,645 deaths in adults, children, and infants in 2017.

Conclusions Apart from two codes, all ICD-10 codes could be mapped to ICD-11 for underlying COD. These findings suggest that the ability to translate from the two iterations of coding will be feasible for future applications of health services data.

Keywords International Classification of Diseases, ICD11, Mortality

Background

A death certificate is completed at the time of death. Information collected includes immediate causes, underlying causes, and significant conditions. These causes of death (COD) are collected, processed, and presented in a standardized manner according to the International Classification of Disease (ICD) coding system [1]. The 10th revision of ICD (ICD-10), introduced in 1992 is a classification system used by over 150 countries around the globe [2]. Limitations of ICD-10 include code fidelity, coding accuracy, and location and activity information capture. Reporting of statistics using ICD coding has been found to underestimate disease and injury causes, largely due to missing/incomplete codes or a lack of specificity in the codes used [3–10]. Given the lack of

*Correspondence:

Danielle A. Southern
dasouthe@ucalgary.ca

¹ Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada

² Centre for Health Informatics, Cumming School of Medicine, University of Calgary, CWPB 5E43, 3280 Hospital Dr NW, Calgary, AB T2N 4Z6, Canada

³ Provincial Research Data Services, Alberta Health Services, Calgary, AB, Canada

⁴ Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

specificity and flexibility of the coding structure in ICD-10, as well as new and revised medical conditions that have arisen in our current medical system, the World Health Organization (WHO) updated and revised the ICD-10 system to improve specificity of its coding capacity.

Approximately 8,000 ICD-10 codes are available for underlying cause coding [11]. The 11th version of the ICD (ICD-11) was released in 2018, adopted at the World Health Assembly in May 2019, and has been in use globally since January 1st, 2022 [2]. As of May 2024, 132 member states are in various stages of the ICD-11 implementation process [2]. It allows for more underlying COD codes from a list that includes 55,000 more clinical concepts and external causes compared with ICD-10 [12]. Eynstone-Hinkins and Moran [13] reviewed Australian mortality data features and pointed out that ICD-11 allows for updated clinical terms and language, multiple parenting, clustering of codes with both mandatory and optional postcoordination, integrated terminology foundation and, most importantly, compatibility and integration with electronic records. To facilitate adoption of ICD-11, Popescu and Celik et al. [14] tested WHO automated selection system of underlying COD for ICD-11. Of 1248 death certificates, 62.8% were accurately assigned underlying COD. The remaining certificates, however, require expert intervention. Stausberg and Vogel [15] coded 200 death certificates with 1292 ICD-11 codes and 1068 ICD-10 codes. Specifying ICD-11 codes requires postcoordination. However, death certificates commonly lack details such as disease stages for chronic kidney disease stage 3a (GB61.2) and stage 3b (GB61.3) so that more ICD-11 codes were selected than ICD-10 codes for the same certificates. For a seamless transition, the WHO has developed mapping tools for international use to ensure no data is lost in the transition process [16]. Due to the international implementation of ICD-11, the mapping tools must be assessed using data from a “real-world” setting. Thus, the WHO will revise or refine mapping tools further.

This study translated ICD-10 COD codes to ICD-11 codes using the WHO mapping tool and identified unmapped ICD-11 codes using a dataset from Alberta, Canada. We then grouped ICD-10 and ICD-11 coded COD using the US Centre for Disease Control (CDC) mortality tabulation lists and compared the top underlying COD between both datasets.

Methods

Translating ICD-10 to ICD-11 COD codes

WHO mapping lists were used to find one-to-one translation of ICD-10 to ICD-11 for each code at the 4-digit level. The mapping identified a single ICD-11 code for

each ICD-10 code. Where there was no ICD-11 code to match the ICD-10 code, it was labeled as “no match”. See Additional file 1 for the complete translation. Categories of more than one code included all the ICD-11 codes found at the individual ICD-10 level. Some clusters of stem codes and extension codes are listed for systems where extensions codes may be coded as these clusters presented specific death information related to the grouping.

The CDC has published two Mortality Tabulation Lists (general, infant and child mortality) for categories within most ICD-10 chapters for conditions and external causes [17, 18]. We first translated the ICD-10 List of 358 Selected COD and 130 Selected Causes of Infant Death to ICD-11 [18]. To do this, we used computational and manual approaches to translation. We retrieved the initially selected underlying COD from the CDC ICD10 COD Lists for Tabulating Mortality Statistics [19]. The WHO crosswalk tables for ICD-10 to ICD-11 (the most recent mapping retrieved WHO crosswalk tables on February 5th, 2021) [4] were used to map codes from ICD-10 to ICD-11. We identified the corresponding ICD-11 code for each ICD-10 code and then populated an Excel spreadsheet row with the CDC grouping number, name, ICD-10 code, and the matching ICD-11 code. If the line from the grouping table only held one ICD-10 code, the mapped ICD-11 code was identified. If a range of codes was present, e.g. A00-A09, we included all mapped codes for the full range. We manually input missing codes from the spreadsheet by searching the code description from ICD-10 in the ICD-11 Coding Tool [20]. All analyses were done in Python [21].

Grouping ICD-11 COD codes

ICD-11 codes were then aggregated into groups mapped to the CDC COD groupings [5, 6] based on the original ICD-10 codes (Additional file 1). Finally, ICD-11 tabulation lists were created from the current ICD-10 tabulation lists.

Data sources

We used Vital Statistics data from Alberta, Canada, to identify the total number of deaths (number of individuals) and the frequencies of underlying COD at the fourth ICD digit level. The population consisted of Alberta residents who died in the fiscal year of 2017. Data from the 2017 fiscal year was used, as we were confident in the availability of comprehensive and complete data from vital statistics given that data extraction took place in 2020. Deaths and underlying COD were recorded in the Vital Statistics Database and held within the Analytics data warehouse of Alberta Health Services. In Canada, the tabulation of the COD is based on the underlying

COD, which the World Health Organization defines as the disease or injury that initiated the chain of events leading to death or as the circumstances of the accident/violence that caused a fatal injury. The underlying COD is selected from the causes and conditions listed on the medical certificate of COD completed by a medical professional, medical examiner or coroner. [22] Any individuals in the data that did not have an underlying COD code were excluded from the analysis. This study was approved by the Conjoint Health Research Ethics Board at the University of Calgary (REB20-1557). To access personally identifiable health information, a waiver of consent was provided on the conditions of Sect. 50 of the Health Information Act of Alberta, Canada.

Statistical considerations

Aggregated counts for each grouping based on ICD-10 coding were retrieved, and stratified by adults (age ≥ 18), children (age 1–17), and infants (age < 1 year). Frequencies were calculated for adults, children, and infants. For each age category, we identified unmapped codes, and the top ten COD in ICD-10 and 11 in Alberta were reported.

Results

Deaths in Alberta in 2017

In Alberta, there were 24,742 deaths in 2017. After removing all records with no documented COD, 24,645 records remained. We excluded one record as it did not have a documented age. Of the 24,644 deaths, 24,252 were adults, 147 were children, and 245 were infants.

Mapping results

The number of ICD-10 codes (rather than individual deaths) and total deaths (based on individuals counted) in mortality statistics for each age group are presented in Table 1. The mapping provided ICD-11 codes for 96.6% of the ICD-10 underlying COD codes for the adult population and 99.5% of the deaths were captured through mapping. The mapping provided 100% of the ICD-10 underlying COD codes and deaths. Only two ICD-codes were unmapped to ICD-11 for the adult population. These codes were I13.1 (hypertensive heart and renal disease with renal failure) and L08.9 (local infection of skin and subcutaneous tissue, unspecified).

COD in Alberta

We mapped ICD-11 codes and reported the ten most frequently reported underlying COD in Alberta for adults, children, and infants in 2017 (see Table 2). As the larger category of malignant neoplasms was reported as the top COD for adults in 2017, we undertook a separate analysis to determine the top ten categories within malignant

Table 1 2017 Alberta underlying COD by ICD-10 codes

Age category	Frequency type*	Count	Mapped ICD-11 code
Adults (≥ 18 years)	Unique ICD-10 Codes	1,039	1,033 (96.6%)
	Total number of deaths	24,252	24,134 (99.5%)
Children (1–17 years)	Unique ICD-10 Codes	87	87 (100.0%)
	Total number of deaths	147	147 (100.0%)
Infants (< 1 year)	Unique ICD-10 Codes	97	97 (100.0%)
	Total number of deaths	245	245 (100.0%)

* Unique ICD-10 codes represent the total number of CODs; total number of deaths represent the number of deaths that occurred in the Albertan population in the year 2017

neoplasms that contributed to this COD (see Table 3). Deaths caused by malignant neoplasms of the trachea, bronchus and lung occurred three times more than the second leading malignant neoplasm (breast). Since the ICD-11 results are simply mapped from the ICD-10 codes, ICD-10 results were not reported as they were the same.

Discussion

The release of ICD-11 brings new mortality codes for enhanced specificity and breadth of medical statistics and surveillance. The imminent transition to ICD-11 also requires mapping to previous versions of ICD-10 to ensure minimal data loss during the adoption period. The implication of the current study primarily included the identification of unmapped ICD codes when applied to a population-level, “real-world” dataset from Alberta, Canada. We also provided ICD-11 Mortality Tabulation Lists (see Additional file 1).

Overall, there was a good translation of ICD-10 to ICD-11 with over 95% of ICD-10 COD codes being mapped to an ICD-11 equivalent. We found two ICD-10 codes that were not mapped in ICD-11 and did not find identifiable codes through manual search. For these two specific codes, it may be related to the introduction of post-coordinated codes in ICD-11. Post coordinated codes is a new feature of ICD-11, which requires information on the sequence of disease from the COD certificate. This allows the provision of multiple values, for the axes of “associated with,” “causing condition,” and “medication,” as examples. For instance, ICD-10 has one code (I13.1) for hypertensive heart and renal disease with renal failure. ICD-11 has specific code BA01 for hypertensive heart disease and BA02 for hypertensive renal disease. Kidney

Table 2 Top 10 ICD-11 Underlying COD in Alberta in 2017

Causes of Death Adults (≥ 18 years)	N	Cause of Death Children (1–17 years)	N	Cause of Death for Infants (< 1 year)	N
Malignant neoplasms	6539	Intentional self-harm (suicide)	24	Newborn affected by maternal factors and by complications of pregnancy, labor and delivery	63
Ischemic heart diseases	3596	Malignant neoplasms	16	Disorders related to length of gestation and fetal malnutrition	33
Other diseases of the respiratory system	2179	Benign neoplasms	< 10	Accidents (unintentional injuries)	10
Other heart diseases	1354	Accidents (unintentional injuries)	< 10	Influenza and pneumonia	< 10
Cerebrovascular diseases	1117	Accidental poisoning and exposure to noxious substances	< 10	Malignant neoplasms	< 10
Accidental poisoning and exposure to noxious substances	739	Ill-defined and unknown causes of mortality	< 10		
Intentional self-harm (suicide)	601	Other diseases of the respiratory system	< 10		
Hypertensive diseases	545	Assault (homicide)	< 10		
Endocrine diseases	536	Event of undetermined intent	< 10		
Diseases of liver	473	Zoonotic and other bacterial diseases	< 10		

Table 3 Examples of Underlying COD within the Malignant Neoplasm Category for adults in Alberta in 2017

Malignant Neoplasm top 10 causes of death	N
Malignant neoplasms of trachea, bronchus, and lung	1565
Malignant neoplasm of breast	461
Malignant neoplasms of Colon	449
Malignant neoplasm of prostate	432
Malignant neoplasms of pancreas	419
Other malignant neoplasms of other and unspecified sites	368
Malignant neoplasm of brain	231
Non-Hodgkin's lymphoma	216
Malignant neoplasms of Rectosigmoid junction and rectum	185
Malignant neoplasms of esophagus	170

failure was coded as acute (GB60), chronic (GB61) and unspecified (GB6Z). GB60/BA02 is for coding acute kidney failure that is caused by hypertensive renal disease. When acute kidney failure is caused by both hypertensive heart disease and kidney disease, the ICD-11 code is GB60/BA02/BA01.

US CDC developed mortality tabulation lists for three age groups for ICD-10 data. These lists are mainly used for reporting and surveillance causes of death. The US CDC has published two Mortality Tabulation Lists (general mortality, and infant and child mortality) that have categories within most ICD-10 chapters for conditions and external causes. We mapped ICD-11 to the ICD-10 List of 358 Selected COD and the List of 130 Selected Causes of Infant Death, using computational mapping and manual approaches. Our mapping identified two ICD-10 codes that are subdivided into multiple codes in ICD-11. Thus, ICD-10 and ICD-11 tabulation lists are matched well. We reported the top 10 COD in ICD-11

and subdivided malignant neoplasms. The utility of ICD-11 mortality tabulation lists in Alberta data demonstrated coverage of concepts.

Our study has limitations. We do not have a dually coded mortality database. We translated ICD-10 into ICD-11 using WHO mapping and coding tools. Thus, the utility of the ICD-11 mortality tabulation lists is available, but we cannot provide validity of these lists as ICD-11 has not yet been implemented for coding death certificates in Canada and dually coded dataset is not available for studying at this time. We only studied underlying COD in this study. This was because most mortality statistics show a single COD for an individual, regardless of how many conditions are reported on the certificate. The underlying COD is the condition selected for such single-cause tabulation. Mortality indicators that use multiple causes of death, however, highlight health conditions that contribute considerably to death but are less likely to feature as underlying cause [23]. Multiple causes of death capture is a valuable tool for public health. This study however provides a more general use of ICD-11, where only single-cause tabulation is done [24]. While we have assessed the translation of ICD-10 to ICD-11 in one Canadian provincial dataset, we highly recommend these lists be tested in other datasets.

Conclusions

In conclusion, the WHO mapping table between ICD-10 and ICD11 works well in a “real-world” database from a large Canadian province. Nearly all ICD-10 codes could be mapped to ICD-11. We highly recommend ICD-11 mortality tabulation lists be tested and refined.

Abbreviations

ICD-11 11Th revision of the International Classification of Disease

ICD-10 10Th revision of the International Classification of Disease
 WHO World Health Organization
 CDC Centers for Disease Control and Prevention
 COD Causes of death

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12963-024-00358-6>.

Additional file 1.

Acknowledgements

The authors would like to acknowledge the support provided by Eric Fu.

Author contributions

HQ designed the Study. CD, DAS directed the study's implementation. DAS, JAK designed the analytical strategy, performed the analysis and helped to interpret the findings. CD conducted the literature review and helped prepare the draft manuscript. CD and DAS drafted the manuscript. All authors reviewed and approved of the final manuscript.

Funding

This work is supported by a Foundation Grant, led by Hude Quan, through the Canadian Institutes of Health Research.

Availability of data and materials

Due to data sharing policies and the Data Disclosure Agreement of the data custodians, the dataset is not able to be made publicly available. It may be able to be shared only to researchers in Alberta with approval from the data custodians. (<https://www.albertahealthservices.ca/research/Page16074.aspx>).

Declarations

Ethics approval and consent to participate

This study was approved by the Conjoint Health Research Ethics Board at the University of Calgary (REB20-1557). To access personally identifiable health information, a waiver of consent was provided on the conditions of Sect. 50 of the Health Information Act of Alberta, Canada.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 11 January 2023 Accepted: 8 December 2024

Published online: 18 December 2024

References

- Centres for Disease Control and Prevention. ICD-10 Cause-of-Death Lists for Tabulating Mortality Statistics (Updated October 2002 to include ICD codes for Terrorism Deaths for data year 2001 and WHO updates to ICD-10 for data year 2003) 2002 [Available from: https://www.cdc.gov/nchs/data/dvs/im9_2002.pdf.pdf, last accessed March 8, 2022]
- World Health Organization. International Statistical Classification of Diseases and Related Health Problems (ICD). [Available at: <https://www.who.int/standards/classifications/classification-of-diseases>, last accessed March 8, 2022]
- Peden AE, Franklin RC, Leggat PA. Fatal river drowning: the identification of research gaps through a systematic literature review. *Inj Prev*. 2016;22:202–9.
- Lu TH, Lunetta P, Walker S. Quality of cause-of-death reporting using ICD-10 drowning codes: a descriptive study of 69 countries. *BMC Med Res Methodol*. 2010;10:30.
- Davie G, Langley J, Samaranyaka A, et al. Accuracy of injury coding under ICD-10-AM for New Zealand public hospital discharges. *Inj Prev*. 2008;14:319–23.
- Smith GS, Langley JD. Drowning surveillance: how well do E codes identify submersion fatalities. *Inj Prev*. 1998;4:135–9.
- Finch CF, Boufous S. Do inadequacies in ICD-10-AM activity coded data lead to underestimates of the population frequency of sports/leisure injuries? *Inj Prev*. 2008;14:202–4.
- Woo JG, Zeller MH, Wilson K, et al. Obesity identified by discharge ICD-9 codes underestimates the true prevalence of obesity in hospitalized children. *J Pediatr*. 2009;154:327–31.
- Lunetta P, Penttilä A, Sajantila A. Drowning in Finland: “external cause” and “injury” codes. *Inj Prev*. 2002;8:342–4.
- Jetté N, Quan H, Hemmelgarn B, Drosler S, Maass C, Moskal L, Paoin W, Sundararajan V, Gao S, Jakob R, Ustün B, Ghali WA; IMECCHI Investigators. The development, evolution, and modifications of ICD-10: challenges to the international comparability of morbidity data. *Med Care*. 2010;48(12):1105–10.
- Kurbasic I, Pandza H, Masic I, Huseinagic S, Tandir S, Alicajic F, Toromanovic S. The advantages and limitations of international classification of diseases, injuries and causes of death from aspect of existing health care system of bosnia and herzegovina. *Acta Informatica Medica*. 2008;16(3):159–61.
- Janssens A, Devos I. The limits and possibilities of cause of death categorisation for understanding late nineteenth century mortality. *Soc His Med*. 2022;35(4):1053–63.
- Eynstone-Hinkins J, Moran L. Enhancing Australian mortality data to meet future health information demands. *Int J Environ Res Public Health*. 2022;19(1):603.
- Popescu MH, Celik C, Della Mea V, Jakob R. Preliminary validation of a rule-based system for mortality coding using ICD-11. *Stud Health Technol Inform*. 2022;25(294):679–83.
- Stausberg, J, & Vogel, U. (2024). Coding Diagnoses from the electronic death certificate with the 11th revision of the international statistical classification of diseases and related health problems: an exploratory study from Germany. *Healthcare*, 12(12).
- World Health Organization. ICD-11 for Mortality and Morbidity Statistics–ICD-10/ICD-11 Mapping Tables. [Available at: <https://icd.who.int/brows/e11/Downloads/Download?fileName=mapping.zip>, last accessed March 9, 2022].
- Centres for Disease Control and Prevention. Tenth Revision 358 Selected Causes of Death Adapted for use by DVS. [Available at: https://wonder.cdc.gov/wonder/sci_data/mort/mcmort/type_txt/mcmort04/358cause.pdf, last accessed March 8, 2022]
- Centres for Disease Control and Prevention. Tenth Revision 130 Selected Causes of Infant Death Adapted for use by DVS. [Available at: https://wonder.cdc.gov/wonder/sci_data/natal/linked/type_txt/cohort99/130Cause99.pdf, last accessed March 8, 2022]
- National Centre for Health Statistics. Instructions for Classification of Underlying and Multiple Causes of Death – Section I – 2021. [Available at: <https://www.cdc.gov/nchs/nvss/manuals/2a-sectioni-2021.htm>, last accessed Aug 30, 2024].
- World Health Organization. ICD-11 Coding Tool. [Available at: https://icd.who.int/ct/icd11_mms/en/release, last accessed Aug 30, 2024].
- Python Software Foundation. Python Language Reference, version 2.7. [Available at: <http://www.python.org>, last accessed Aug 30, 2024].
- Statistics Canada. Statistics Canada, Canadian Vital Statistics - Death database (CVSD). Available at: <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3233>, last accessed Aug 30, 2024]
- Bishop K, Moreno-Betancur M, Balogun S, Eynstone-Hinkins J, Moran L, Rao C, Banks E, Korda RJ, Gourley M, Joshy G. Quantifying cause-related mortality in Australia, incorporating multiple causes: observed patterns, trends and practical considerations. *Int J Epidemiol*. 2022;52(1):284.
- Désésquelles A, Demuru E, Egidi V, Frova L, Meslé F, Pappagallo M, Salvatore MA. Cause-specific mortality analysis: is the underlying cause of death sufficient? *Revue Quetelet/Quetelet Journal*. 2014;1(2):119–35.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.