Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200170

Logical Rules and a Preliminary Prototype for Translating Mortality Coding Rules from ICD-10 to ICD-11

Vincenzo DELLA MEA[a,1], Mihai H POPESCU[a], Francesco GRIPPO[b], Chiara ORSI[b], Friedrich HEUSER[c]
[a] *Dept. of Maths, Computer Science and Physics, University of Udine, Italy*
[b] ISTAT, Italy
[c] *Iris Institute, DIMDI, Germany*

Abstract. Iris is a system for coding multiple causes of death in ICD-10 and for the selection of the underlying cause of death, based on a knowledge base composed by a large number of rules. With the adoption of ICD-11, those rules need translation to ICD-11. A pre-project has been carried out to evaluate feasibility of transition to ICD-11, which included the analysis of the logical meta-rules needed for rule translation and development of a prototype support system for the expert that will translate the coding rules.

Keywords. Classifications, ICD-10, ICD-11, ontology mapping

1. Introduction

With its adoption by the 72th World Health Assembly in May 2019, ICD-11 (International Classification of Diseases, Revision 11) will become the new standard for coding diseases and health problems. For mortality statistics, the ICD coding is performed with the use of semi-automated coding systems, mainly Iris [1].

Iris is a system for coding multiple causes of death in ICD-10 and for the selection of the underlying cause of death. Iris is based on the international death certificate form and coding rules provided by WHO (World Health Organisation). The core component of Iris are the decision tables, currently based on ICD-10 codes. For the transition of Iris to ICD-11, it will be necessary to translate the decision tables and to allow them to include all the features of the new revision.

This transition is expected to have a big impact on the system transition and use. Periodic revision of the ICD is essential to stay abreast of advances in medical science and changes in medical terminology [2,3]. Institutionally, revision of the ICD requires an enormous investment of national resources to revise software, training, publications, edit procedures, etc [4]. For the Iris transition to ICD-11, classification and coding experts are needed. To support their work, formal procedures are needed to ensure the correctness of the transition and validation of the system.

[1] Corresponding Author, Vincenzo Della Mea, Dept. Of Mathematics, Computer Science and Physics, University of Udine, via delle Scienze 206 – 33100 Udine Italy E-mail: vincenzo.dellamea@uniud.it.

A pre-project has been carried out to evaluate feasibility of transition to ICD-11, which included the analysis of the logical meta-rules needed for rule translation and development of a prototype translation support system aimed at the expert [5]. This paper describes the logical rules identified for translating from ICD-10 to ICD-11 with their consequences from the point of view of automation and human expert intervention, with an overview of the prototype, and evaluation of efforts needed.

1.1. Decision tables

The decision tables form a knowledge base of relations between pairs of codes (representing the causes of death reported on the death certificate) that must be taken into consideration during the application of the steps for the selection of the underlying cause.

This knowledge base was first developed by the NCHS (US National Center for Health Statistics) for the ACME system [6]. Successively it has been embedded in the new automated coding system Iris and updated on the basis of the recommendations of the Mortality Reference Group, which operates in the network of the WHO Collaborating centers for the Family of International Classifications (WHO-FIC) [7].

In practical terms, the decision tables are a list of possible kind of relations between pairs of terminal code. The relation between the codes are specified as “rule type”, which are used in the different steps of selection.

1.2. Mapping between classification versions

Along with ICD-11, the WHO releases the mapping table between the ICD-10 and ICD-11 classification. The mapping table contains the mapping of both terminal codes and higher categories which can be used to add detail.

The mapping between the ICD-10 and ICD-11 entities can be specified by cardinality as: *1x1* (equivalent), *1xn*, *nx1* and *nxn* (structure of the classification change and the categories intersect each other) with *axb* where *a* is the cardinality of the ICD-10 codes and *b* is the cardinality of ICD-11 codes.

2. Methods

For this analysis the 2019 decision tables are considered, since they include the most recent ICD-10 updates. The most represented rule is the DUETO, with more than 27 million records (90.1% of the total rules). The analysis thus focused on the DUETO rules. These rules represent the vast majority of the rules and are the most critical for coding.

2.1. Translation method

From former evaluations of the transition between the ICD-9 and ICD-10, it is known that the transition to ICD-11 will have an impact and the mapping will not be enough to completely automate the transition. The mapping table gives the possibility to translate single codes, but when it comes to the rules, we need to translate relationships between codes. Due to the relation between the pair elements, we cannot translate codes separately, but we need to interpret the relation from a logical point of view and the impact that the diverse mapping cardinality has on the two codes involved.

2.2. DUETO rule-type translation

The *DUETO* relation can shortly be described as:
Code A DUETO B if B is an acceptable cause of A (according to ICD provisions); in IRIS terminology, A is called codeDef and B is called subcodeDef.

The goal of the translation is to identify all the possible translation rules that can be automated. Afterwards we define the DUETO rule-type in combination with the mapping cardinality of the two codes that can be automated, and which needs expertise support.

The decision tables are based on pair of codes, but since the DUETO rules may include many consecutive pairs, the conditions can be represented also as ranges of codes. Using ranges, it is possible to achieve better results but also to reuse part of the translation of the *subcodeDef*.

Given a rule, the basic idea is to verify whether a mapping exists between single codes for both *codeDef* and *subcodeDef* and of which kind. A rule can be translated automatically if both sides can be translated automatically. In some cases, we may need a different translation of codes for *codeDef* and *subcodeDef*, since the consequences of the rule-type could be different. So we want to always suggest the translation codes based on the mapping but, basing on the mapping type and the implications of the relation, the system also suggest which are the codes that can be accepted or need expert supervision.

Since the disease knowledge evolves, some ICD-10 codes may have no mappings to the ICD-11 classification since the concept is no longer used, or ICD-11 codes may have no mapping from the ICD-10 since they are new. Those codes are problematic and need expert supervision. Sometimes, using ranges for the translation, some of those codes could be handled. Since the classification is hierarchical, translating the higher levels could avoid manual intervention on some of the terminal codes with no mapping, since the results would be covered by the parent translation.

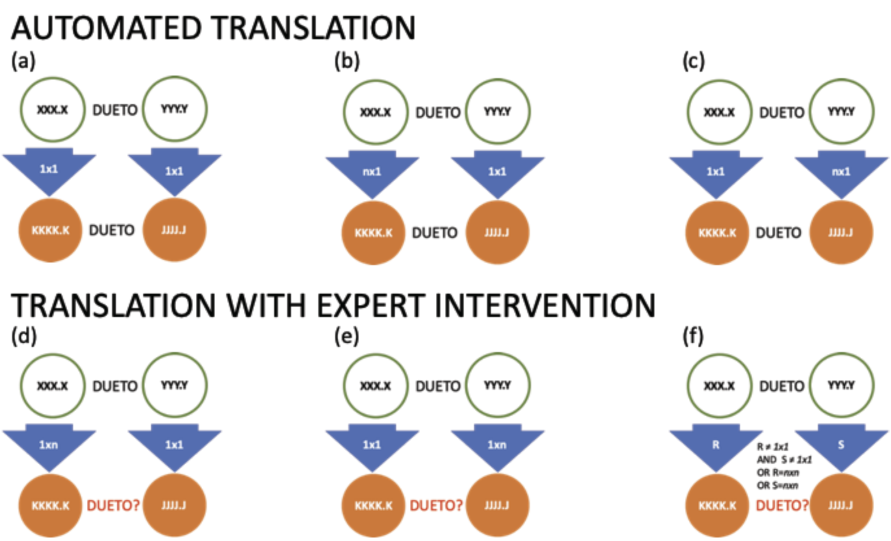

Figure 1. Summary of translation cases.

3. Results

Basing on the cardinality of the mappings on both sides, we can distinguish between manual and automatic translations.

The equivalent mapping *(1x1)*, when found on both sides of the rule, is the basic case of automation, applicable to all the rule types. On the other side, when a mapping with cardinality of *nxn* is present, translation will always be manual. Rules where mapping of type *nx1* and *1xn* appear are those that provide some chance of automated translation, as depicted in Figure 1.

The key for the DUETO rule translation is to remember that is not representing causality, but the possibility of causality. Thus, whenever aggregated conditions are present in the ICD-11 translation, such causality could not be excluded and thus the mapped rule is valid as obtained from the mappings. Let's examine two examples of automated translation in table 1.

Table 1. Automated translation examples.

case	ICD-10 rule	Mappings involved	ICD-11 rule
(a)	A01.0 (Typhoid fever) DUETO C33 (Malignant neoplasm of trachea)	A01.0 ≡ 1A07 C33 ≡ 2C24	1A07 (Typhoid fever) DUETO 2C24 (Malignant neoplasms of trachea)
(b)	F03 (Unspecified dementia) DUETO R54 (Senility) *But also:* *F01 (Vascular dementia DUETO R54 (Senility)*	F03 ⊑ 6D8Z R54 ≡ MG2A *But also:* *F00 (Dementia in Alzheimer disease) ⊑ 6D8Z* *F01 (Vascular dementia) ⊑ 6D8Z*	6D8Z (Dementia, unknown or unspecified cause) DUETO MG2A (Old age)
(c)	J96.9 (Respiratory failure, unspecified) DUETO F03	J96.9 ≡ CB41.2 F03 ⊑ 6D8Z *But also:* *F00 ⊑ 6D8Z* *F01 ⊑ 6D8Z*	CB41.2 (Respiratory failure, unspecified as acute or chronic) DUETO 6D8Z

Let's now examine the cases that need expert intervention in table 2. In this case, since ICD-11 conditions are recorded with greater detail, we cannot ensure all of them are reasonably part of a DUETO rule (but the expert will know and decide). No example is provided for mixed situations involving *nxn* because too complex to be shown here.

Table 2. Examples of rules translation needing expert intervention.

case	ICD-10 rule	Mappings	ICD-11 rule
(d)	I46.9 (Cardiac arrest, unspecified DUETO R26.3 (Immobility)	I46.9 ⊒ MC82 (Cardiac arrest) I46.9 ⊒ MC82.0 (Ventricular tachycardia and fibrillation cardiac arrest) I46.9 ⊒ MC82.1 (Bradycardic cardiac arrest) I46.9 ⊒ MC82.2 (Asystolic cardiac arrest) I46.9 ⊒ MC82.3 (Cardiac arrest with pulseless electrical activity) R26.3 (Immobility) ≡ MB44.3 (Immobility)	We cannot ensure that all types of cardiac arrest can be due to immobility
(e)	I27.9 (Pulmonary heart disease, unspecified) DUETO B44.1 (Other pulmonary aspergillosis)	I27.9 ≡ BB0Z (Pulmonary heart disease or diseases of pulmonary circulation, unspecified) B44.1 ⊒ CA82.4 (Aspergillus-induced allergic or hypersensitivity conditions) B44.1 ⊒ 1F20.12 Chronic pulmonary aspergillosis)	We cannot ensure that both CA82.4 and 1F20.12 can cause Pulmonary heart disease.

In cases like (d) and (e), the expert should be given the list of mappings that map to the same ICD-11 entity, to let him decide which ones generate valid DUETO rules.

Using this method for the translation the expert has a tool to quickly get an overview of the suggestions for the translation, but also has the possibility to focus more on the results that are suggested as manual after the validation of the method.

CodeDef and *subcodeDef* can be translated separately, and for the translation of the subcodeDef we can reuse the same translation in different rules. A further improvement is the translation reuse of a sub-range, that can be used for the translation of ranges that incorporate the translated range. Both techniques help in reducing human intervention.

Further details on the translation methodology can be found in [5].

By exploring the distribution of mapping types in the coding rules, we evaluated in about 3.2 million pairs those needing manual intervention, out of about 27 millions. With the reuse of the validated range translation we estimated a relevant reduction, where the cases that still need manual support are about 87000 codes (considering manual translation of *codeDef* and *subcodeDef*).

3.1. Prototype

The prototype is aimed at providing decision support in the translation of mortality rules from ICD-10 to ICD-11. For the implementation, we choose a web-based model where experts can work collaboratively from a different location but also for the consistency of the results. It has a full implementation of the visualization and editing of the "DUETO" translation rules of the decision tables. Rules are grouped where possible with the same subcodeDef to ease the translation and facilitate maintenance over time.

4. Conclusion

The presented method and prototype seem suitable for supporting the process of transition of Iris from ICD-10 to ICD-11, however it further needs expert validation to correctly estimate the workload needed. Adaption of results obtained on the DUETO rules is being carried out for the other rules, in particular those describing direct sequels of conditions.

References

[1] The Iris Institute. URL: www.iris-institute.org

[2] T.Boerma, J.Harrison, R.Jakob, C.Mathers, A.Schmider, S.Weber, Revising the ICD: explaining the WHO approach, *Lancet* **388** (2016), 2476-2477.

[3] The World Health Organization. *ICD-11 Implementation or Transition Guide.* WHO, 2019. URL: https://icd.who.int/docs/ICD-11%20Implementation%20or%20Transition%20Guide_v105.pdf

[4] R.N. Anderson, A.M. Miniño, D.L. Hoyert, H.M. Rosenberg, Comparability of Cause of Death Between ICD–9 and ICD–10: Preliminary Estimates, *National Vital Statistics Report* **49** (2001),1-32.

[5] The Iris Institute. *Evaluation Project on Iris and ICD-11*. URL: https://www.dimdi.de/dynamic/en/classifications/iris-institute/iris-and-icd-11/

[6] R.A. Israel, Automation of mortality data coding and processing in the United States of America, *World Health Stat Q* **43** (1990),259-62

[7] S.Navarra, M.Cappella, L.A.Johansson, L.Pelikan, L.Frova, F.Grippo, *Decision Table Editor: a web application for the management of the international tables for mortality coding*, ISTAT working papers, Rome, 6/2016. URL: https://www.istat.it/it/archivio/184113