Challenges, opportunities, and approaches in a health KOS vocabulary's revisions -- Insights from the 11th revision of the International Classification of Diseases (ICD-11)

Marcia Zeng [mzeng@kent.edu]

School of Information, Kent State University, Kent, Ohio, USA.

Yi Hong [ylphhong@yahoo.com]

Department of Professional Services, MedeAnalytics, Inc., Richardson, Texas, USA.

Aims

When a widely used and standard health knowledge organization system (KOS) needs to respond to challenges to essential health services and consider evolutions in treating health issues in the fast-advanced digital age, the KOS also faces challenges. Since a dramatic revision may call for the need to adjust its contents, structures, coding systems, and guidelines for use, this KOS must ensure semantic interoperability, data reusability, and the community's agreement. The *International Classification of Diseases* (ICD) has been representing very unique features in its revisions for more than 100 years. The most recent 11th Revision of ICD (ICD-11) is officially in implementation in 2022, nearly three decades after the release of the 10th Revision (ICD-10). A study of the ICD-11 revealed its distinctive methods of embracing digital technologies and unifying the ultimate KOS approaches used by classification, thesauri, ontologies, etc., much more than a replication of the previous edition. This presentation aims to discuss the movements of knowledge organization systems in the digital age, especially aligning with the advancements of information technologies, as demonstrated by ICD-11.

Method

This presentation is based on a study of the ICD revisions, especially the one that has been in use for nearly 30 years (ICD-10) and the newest one (ICD-11) which came in effect starting in Jan. 2022.
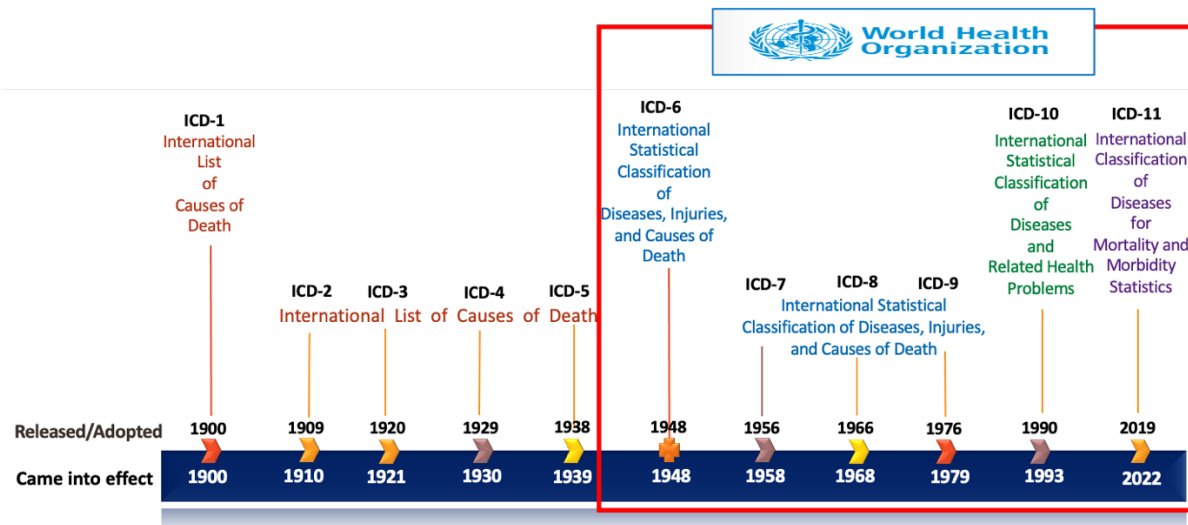
Image source: Created by Yi Hong and Marcia Zeng, based on Moriyama et al. 2011. *History of the Statistical Classification of Diseases and Causes of Death* and World Health Organization. 2022a. "History of ICD."

ICD-11 maintained all ICD-10 chapters while made a vast enlargement. The study analyzed the contents, structures, and notation systems (referred to as "code structures" in ICD's Reference Guide) of ICD-10 and ICD-11. It considered the outcomes and the challenges of the demands of new contents together with the semantic interoperability and data reusability issues with any revision.

Main Findings

The study on the contents, notation systems, and components has led us to understand its overall approaches in maximizing the functions of this comprehensive new revision, including:

a) *Having a framework beyond one classification*. In addition to the continuing content changes which reflect knowledge and perspectives on diseases and their causes in each ICD version, changes in design and structure are noteworthy in ICD-11's information framework. The framework is comprised of : 1) a semantic knowledge base (referred as the Foundation), 2) a biomedical ontology lined to the Foundation, and 3) classifications derived from the Foundation (Harrison et al. 2021)."

b) *Becoming a classification AND terminology*. As a statistical classification of diseases, the term "categories" was chosen by ICD regarding the features of ICD's components, intending to facilitate the statistical study of disease phenomena, since it must be confined to a limited number of mutually exclusive categories and be able to encompass the complete range of morbid conditions. To determine whether an entity qualifies to become a unique category, ICD has been applying a set of measures, which aims to retain the dual abilities to identify specific disease entities and to allow statistical presentation of data for broader groups to enable the attainment of useful and understandable information (World Health Organization 2022b, 1.2.1). Aligned with around 17,000 unique codes for the classification, more than 120,000 codable terms are now entirely digital and the smart coding algorithm now interprets more than 1.6 million terms. For example, when searching on COVID-19 related entries, a significant number of deprecated terms can be found, which led to the formal classification entries.

Source: Searching results captured from ICD-11
https://icd.who.int/browse11/l-m/en

c) *Providing Foundation URIs*. For the first time in ICD, each of the ICD-11 entities has its Foundation URI. The use of the URIs enables the ICD-11 entities fully language independent, while having a specific place in a hierarchy of groups, categories, and narrower terms. ICD-11 is combining the elements of classification and terminology and is designed to be linked to other terminologies that may provide additional detail or serve different purposes. Coding in ICD-11 can draw on statistical codes and on URIs (World Health Organization 2022b, 1.1.2). In this way, an international translation base facilitates translations or multilingual browsing (with over 20 languages). It will also eliminate the semantic conflicts and inconsistency issues when modifications are made to the classification, or new versions of the classification are released, after its worldwide implementation.

d) *Delivering with an Implementation Package*. Three decades after the release of ICD-10, the world is now in the digital age and metaverse. ICD-11 comes with an implementation package that includes an online coding tool, an *ICD-11 Implementation or Transition Guide*, and an ICD-API that provides web services to enable remote programmatic access to ICD-11, with enhanced capability to capture and combine clinically relevant characteristics of cases and integrated support for multiple languages. ICD-11 ensures semantic interoperability and reusability of recorded data for the different use cases beyond mere health statistics, including decision support, resource allocation, reimbursement, guidelines and more (World Health Organization 2022c). This electronic architecture allows assignment of unique identifiers to any condition listed - independently whether the condition is grouped in a statistical class or whether it represents a class of its own (World Health Organization 2022b: 1.2.1).

The insights from ICD-11 reveal that a revision of a KOS vocabulary could have unique methods of embracing digital technologies and unifying the principal KOS approaches.

References

- Moriyama, Iwao M, Ruth M. Loy, Alastair H. T. Robb-Smith, Harry M. Rosenberg, and Donna L. Hoyert. 2011. *History of the Statistical Classification of Diseases and Causes of Death*. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention (CDC).
- World Health Organization. 2022a. "History of ICD." In: "International Statistical Classification of Diseases and Related Health Problems (ICD)" homepage. https://www.who.int/standards/classifications/classification-of-diseases.
- World Health Organization. 2022b. ICD-11 Reference Guide. https://icdcdn.who.int/icd11referenceguide/en/html/index.html.
- World Health Organization. 2022c. "ICD purpose and uses." In: "International Statistical Classification of Diseases and Related Health Problems (ICD)" homepage. https://www.who.int/standards/classifications/classification-of-diseases.