SEMI-AUTOMATIC INCORPORATION OF ICD-10-PCS CONCEPTS INTO A COMPREHENSIVE PROCEDURE HIERARCHY BASED ON SNOMED-CT IN THE OMOP COMMON DATA MODEL

Dmitry Dymshyts MD[1], Mr. Eduard Korchmar[1], Christian Reich MD[2], PhD

1-Odysseus Data Services, Cambridge MA USA; 2-IQVIA, Cambridge MA USA

Introduction

We present the process of automatically placing leaf-level ICD-10-PCS concepts in SNOMED-CT hierarchy using manually matched attribute sets as part of our efforts to consolidate procedures from different vocabularies into a singular hierarchical model.

There are key differences in the construction of SNOMED-CT and ICD-10-PCS concepts. As a result, equivalence mapping was not considered an optimal way of constructing a combined hierarchical system.
Placing ICD-10-PCS leaf-level concepts into existing SNOMED-CT hierarchy was chosen instead.

Design philosophy

ICD10PCS design philosophy

1. Procedure concepts are precoordinated by combining all possible attribute combinations, even if not performed in practice.

2. Hierarchy always has 7 levels. Meaning of the procedure is kept entirely in it's attributes.

3. Terminology is strictly defined, not allowing anyambiguity. Terminology may sometimes differ from widely used terms in favour of standardization. E.g. "Diagnostic Excision" instead of "Surgical Biopsy"

4. Only core nature of the procedure performed can be encoded. Usually circumstances of the procedure or target pathology is not represented. Excision of malignant neoplasm, of infectious abscess or of heterotopic structure will all be encoded in the same way.

SNOMED CT design philosophy

1. All concepts are added manually from many sources when deemed neccessary under set of rules. Their ancestors (classification concepts) are added simultaneously.

2. Each procedure concept can have multiple parent concepts and inherits attributes from them, in most cases adding precision to attribute set.

3. Terminology is based on traditional use in clinical practice and is not guaranteed to be consistent. Terms like 'excision' and 'resection' are used interchangeably while ICD10PCS considers them different procedures.

4. Attributes and hierarchy are not strictly defining of procedure meaning. Some attributes may be missing and some aspects of the procedure may be impossible to represent, like procedure-specific approaches and techniques.

Our approach to solution

1. Not every ICD10PCS concept will have exact match among SNOMED CT procedure concepts. We will try to find place in SNOMED CT hierarchy for each ICD10PCS concept by finding closest parent concepts in existing set.

2. Closest parents for each ICD10PCS concept will be found based on having closest attribute set; we can move up the hierarchy and pick more generic targets in cases when there are no target procedures with desired attribute set available.

3. ICD10PCS attributes must be mapped to SNOMED CT concepts that serve as attributes for procedures. Some additional processing may be required, like mapping of pairs of attributes instead of individual ones.

4. Close attention must be paid to SNOMED CT concepts that may have meanings outside their attributes. These must be preprocessed to have missing attributes assigned or excluded from the analysis altogether.

References
1. "SNOMED-CT Editorial Guide" by International Health Terminology Standards Development Organisation. Available from: https://confluence.ihtsdotools.org/display/DOCEG
2. ICD-10-PCS Official Guidelines for Coding and Reporting. Available from: https://www.cms.gov/Medicare/Coding/ICD10/Downloads/2019-ICD10-Coding-Guidelines-.pdf
3. Kin Wah Fung, MD, MS, MA, Julia Xu, MD, PhD;, Filip Ameye, MD, FRCS(Eng), FACS;, Arturo Romero Gutierrez, MD, and Arabella D'Have, Msc. Achieving Logical Equivalence between SNOMED-CT and ICD-10-PCS Surgical Procedures, 2017

Algorithm description

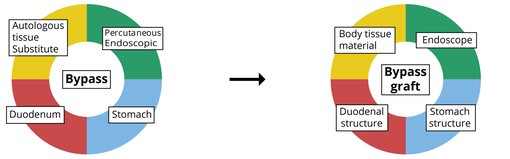
1. Represent ICD10PCS concepts following SNOMED CT attribute-value logic.
2. Create pool of possible parents among SNOMED CT concepts based on common high-level root procedure ancestor and matching attributes.
3. Find closest matching parents in the pool. Reiterate for each available attribute (see **Filtering cycle**).
4. Reiterate steps 2 and 3 with simplified rules to capture generic ancestors
5. Process all found parent concepts to eliminate redundancy; add attribute relations inherited from new-found parent concepts to keep consistency with SNOMED logic.

Attribute transition

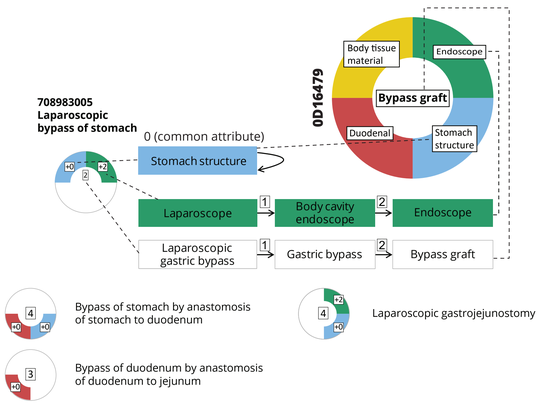
Attributes for 0D16479 "Bypass Stomach to Duodenum with Autologous Tissue Substitute, Percutaneous Endoscopic Approach"

Hierarchical representation:
1. Medical and Surgical
2. Gastrointestinal System
3. Bypass
4. Stomach
5. Percutaneous Endoscopic
6. Autologous Tissue Substitute
7. Duodenum

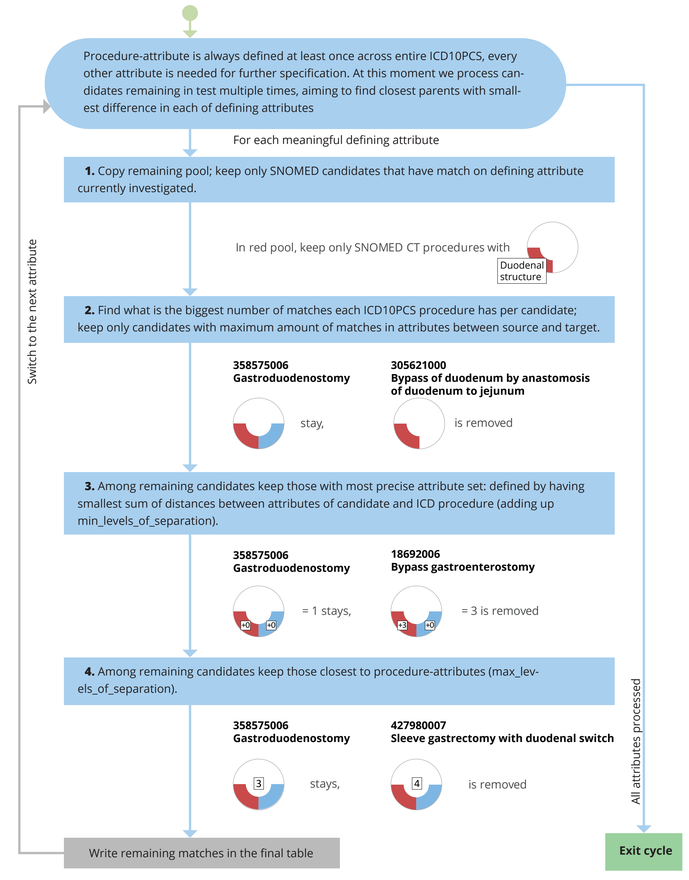
When mapped to SNOMED CT attributes:
1. Medical and Surgical - removed as direct ancestor
2. Gastrointestinal System - removed as direct ancestor
3. 48537004 | Bypass graft (procedure)
4. 69695003 | Stomach structure (body structure)
5. 37270008 | Endoscope (device)
6. 413675001 | Body tissue material (substance)
7. 38848004 | Duodenal structure (body structure)



Examples from test pool

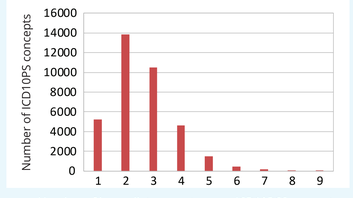


Filtering cycle

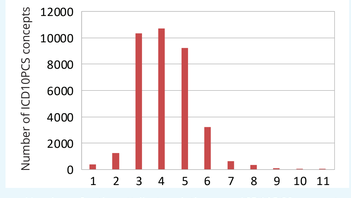


Results

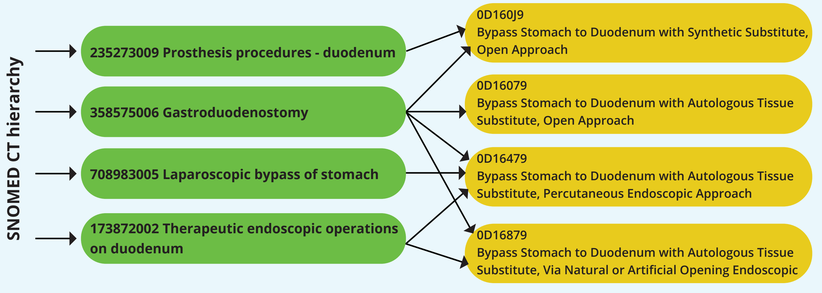
We have found direct parents for more than 40 000 leaf-level ICD-10-PCS concepts most commonly encountered in patient data. Multiple (median of 2) are found to represent each subhierarchy ICD-10-PCS concept is placed in, based on matching attributes.



Number of immediate ancestors per ICD10PCS concept



Number of active attribute relations per ICD10PCS concept



Conclusion

The SNOMED-CT model is an acceptable basis for Procedure consolidation through definition of source Procedure concepts using attributes and subsequent automated analysis to place procedures from other coding systems into the SNOMED-CT hierarchy. Leaf-level concepts from ICD-10-PCS vocabulary can be seamlessly integrated as an extension of existing concept set, complete with attributes. Similar approaches can be used to standardize other procedure vocabularies, given there is a way to adequately represent them using SNOMED-CT attribute-value model.