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.

**ICD10PCS design philosophy**

1. Procedure concepts are pre-coordinated 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 its attributes.
3. Terminology is strictly defined, allowing any ambiguity. Terminology may sometimes differ from widely used terms in favor of standardization.
4. Only core nature of the procedure performed can be encoded. Usually circumstances of the procedure or target pathology is not represented.
5. SNOMED CT design philosophy

   1. All concepts are added manually from many sources where deemed necessary under set of rules. Their ancestors (classification concepts) are added simultaneously.
   2. Each procedure concept can have multiple parent concepts and new attributes are inherited from them, in most cases adding precision to attribute set.
   3. Terminology is based on traditional use of clinical practice and is not guaranteed to be consistent. Terms like ‘excision’ and ‘resection’ are used interchangeably while ICD10PCS considers them different concepts.
   4. Attributes and hierarchy are not strictly defined for 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 parent concept 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. SNOMED CT 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.

**Algorithm description**

**Filtering cycle**

1. Find what is the biggest number of matches each ICD10PCS procedure has per candidate keeping in mind maximum amount of matches in a procedure must be limited in size.
2. Among remaining candidates keep those with most precise attribute set: defined by having smaller set of distances between attribute of candidate and ICD10PCS procedure (adding up min_levels_of_separators).
3. In red pool, keep only SNOMED CT procedures with attributes for which procedure is rightmost ancestor.
4. Among remaining candidates keep those closest to procedure attributes (maxlev). Finally, write remaining matches in the final table.

**Find closest matching parents in the pool. Reiterate for each available attribute (see Filtering cycle).**

**Find what is the biggest number of matches each ICD10PCS procedure has per candidate keeping in mind maximum amount of matches in a procedure must be limited in size.**

**Process all found parent concepts to eliminate redundancy; add attribute relations inherited from new-found parent concepts to keep consistency with SNOMED logic.**

**Create pool of possible parents among SNOMED CT concepts based on common high-level root procedure ancestor and matching attributes.**

**Represent ICD10PCS concepts following SNOMED CT attribute-value logic.**

References:
3. ‘The Integrated SNOMED CT and ICD-10-PCS Procedure Ontology’ 2017

**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.

**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.