

Usability of SNOMED CT for secondary use of data in nursing

Arabella D'Havé, Chief of Terminologie, Classification, Grouping & Audit

Belgian Federal Public Service of Health, Food Chain Safety and Environment

Terminology Center

ABSTRACT

Objective: To create an interoperable list of SNOMED CT concepts pertinent to nursing practice through cross-mapping of seven standardized nursing languages recognized as interface nursing terminologies to support secondary use of data.

Design: The UMLS Metathesaurus was queried in order to retrieve all SNOMED CT concepts where the concept was also represented within seven standardized nursing languages recognized both nationally and internationally for use as an interface terminology in paper-based as well as electronic health records. A candidate list of SNOMED CT concepts was retrieved and presented to a project team in the University Hospitals Leuven in order to validate the SNOMED CT concepts that are present or missing from this list to accommodate with the mandatory Minimum Nursing Dataset within the Minimum Hospital Dataset, which is mainly used for hospital financing and nurse staffing allocation.

Results: An initial query was performed against seven standardized nursing terminologies to retrieve the corresponding UMLS Concept Unique Identifiers (CUIs). Through the retrieved CUIs, any linked SNOMED CT concept was selected in a list. The list was cleaned to remove duplicate concepts resulting in a final list of unique SNOMED CT concepts.

Conclusion: SNOMED CT allows for the use of multiple standardized nursing languages within an electronic health record. Also, with SNOMED CT standardized nursing languages can be mapped to accommodate with mandatory Minimum Nursing Datasets developed for hospital policy, thus supporting secondary use of data. This list can be used to develop a rule-based map allowing semi-automatic retrieval of the mandatory Minimum Nursing Dataset.

METHODS

Step 1 Determine the sources

There are 12 terminology sets that support nursing practice and are recognized by the American Nursing Association (ANA) as interface terminologies for nursing. These include Clinical Care Classification (CCC; formerly Home Health Care Classification), International Classification for Nursing Practice (ICNP), NANDA-International (NANDA-I; Nursing Diagnosis Classification), Nursing Intervention Classification (NIC), Nursing Outcome Classification (NOC), Omaha System, Perioperative Nursing Data Set (PNDS), ABC Coding Solutions, Logical Observation Identifiers Names and Codes (LOINC), Systematic Nomenclature of Medicine Clinical Terms (SNOMED CT), Nursing Minimum Data Sets (NMDS), and Nursing Minimum Data Sets (NMDS). The last two terminologies have been excluded from this study.

Step 2 Extract a dataset of nursing concepts for each source

The UMLS Metathesaurus was the authoritative source for all terminology sampling (2014AA). The UMLS relational files were loaded into a local MySQL database and accessed through SQL queries. Repeated queries for nursing concepts were executed against the first nine nursing terminologies recognized by the American Nursing Association as interface terminologies for nursing.

Queries were executed against MRCONSO taking into account dat every string or concept name in the Metathesaurus appears in this file, connected to its language, source vocabularies and its concept identifier.

Step 3 Create a combined dataset based on unique CUIs

The dataset is reduced to include unique CUIs. Regardless of the source, if concepts are semantically identical, the same concept unique identifier (CUI) is assigned while maintaining the unique code from the source vocabulary. A combined dataset can thus be created based on the unique CUIs resulting from the intersection of the datasets made for each source.

Step 4 Create an initial SNOMED CT subset

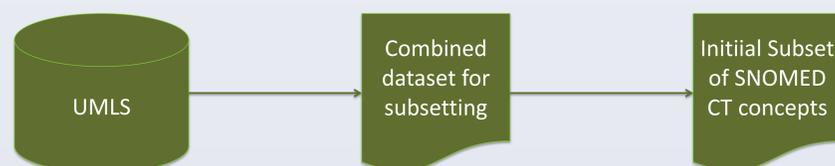
The combined dataset was reduced to include those SNOMED CT concepts that had a corresponding CUI. For the CUIs withheld within the combined dataset, the ones assigned to a SNOMED CT concept were retained, resulting in a subset of SNOMED CT concepts for nursing.

RESULTS

According to the purpose of the study, a total number of 405,131 nursing concepts were extracted from UMLS. For each interface terminology of nursing, the number of unique concepts was extracted, 410 concepts from the Clinical Care Classification (CCC; formerly Home Health Care Classification), 1,689 concepts from the International Classification for Nursing Practice (ICNP), 237 concepts from the NANDA-International (NANDA-I; Nursing Diagnosis Classification), 11,255 from the Nursing Intervention Classification (NIC), 5,707 from the Nursing Outcome Classification (NOC), 554 from the Omaha System, 199 from the Perioperative Nursing Data Set (PNDS), 9,065 from the ABC Coding Solutions system, and 376,015 from the Logical Observation Identifiers Names and Codes (LOINC).

The next step was to create a combined dataset regardless of the source, meaning that if concepts are semantically identical, the same concept unique identifier (CUI) is assigned. This resulted in a combined dataset based on unique CUIs consisting of 156,926 concepts.

Finally, for every CUI in the combined dataset, those corresponding with a SNOMED CT concept were extracted (n=32,981).



Dataset extraction

- CCC (n=410)
- ICNP (n=1,689)
- NANDA-I (n=237)
- NIC (n=11,255)
- NOC (n=5,707)
- Omaha (n=554)
- PNDS (n=199)
- ABC (n=9,065)
- LOINC (n=376,015)
- Total (n=405,131)

Combined dataset based on unique CUIs (n=156,926)

Initial SNOMED CT subset for nursing (n=32,981)

CONCLUSIONS

The initial SNOMED CT subset for nursing now has to be reviewed and cleaned.

- Navigational concepts need to be removed.
- The concepts need to be classified using the components of the nursing process. It needs to be checked if the concepts with regards to the nursing process, including nursing diagnoses, interventions, actions, outcomes, assessments, problem identification, and goal development present are well represented.
- It is expected that the use of this initial subset will save effort in the terminology development for nursing in SNOMED CT.
- It needs to be checked if the subset meets the needs of the ISO reference terminology for nursing.
- It is important to determine whether the language used in the EHR can be represented using the initial subset. The subset may be expanded or reduced through the evaluation of nursing concepts stored within the EHR and everyday concepts used by nurses.

REFERENCES

1. American Nurses Association. ANA recognized terminologies that support nursing practice. ANA, Washington, DC; 2012. <http://www.nursingworld.org/Terminologies> [last accessed 07.05.14].
2. National Library of Medicine. UMLS reference manual. Bethesda (MD): NLM; 2009. <http://www.ncbi.nlm.nih.gov/books/NBK9676/> [last accessed 07.05.14].
3. Fung, K. W., & Bodenreider, O. (2005). Utilizing the UMLS for semantic mapping between terminologies. *AMIA Annu Symp Proc*, 266-270.
4. Kim, T. Y., Coenen, A., & Hardiker, N. (2012). Semantic mappings and locality of nursing diagnostic concepts in UMLS. *J Biomed Inform*, 45(1), 93-100.
5. Matney, Susan A., Warren, Judith J., Evans, Jonathan L., Kim, Tae Youn, Coenen, Amy, & Auld, Vivian A. (2012). Development of the nursing problem list subset of SNOMED CT®. *J Biomed Inform*, 45(4), 683-688.

