# Applying the SNOMED CT concept model to represent value sets for head and neck cancer documentation

Højen, Anne Randorff; Brønnum, Dorthe; Gøeg, Kirstine Rosenbeck; Elberg, Pia Britt

*Published in:*
Exploring Complexity in Health: An Interdisciplinary Systems Approach

*DOI (link to publication from Publisher):*
10.3233/978-1-61499-678-1-436



*Publication date:*
2016

*Document Version*
Publisher's PDF, also known as Version of record



*Citation for published version (APA):*
Højen, A. R., Brønnum, D., Gøeg, K. R., & Elberg, P. B. (2016). Applying the SNOMED CT concept model to represent value sets for head and neck cancer documentation. In *Exploring Complexity in Health: An Interdisciplinary Systems Approach: Proceedings of MIE2016 at HEC2016, 28 August-2 September 2016, Munich, Germany* (pp. 436-440). IOS Press. https://doi.org/10.3233/978-1-61499-678-1-436

*Exploring Complexity in Health: An Interdisciplinary Systems Approach*
*A. Hoerbst et al. (Eds.)*

*doi:10.3233/978-1-61499-678-1-436*

# Applying the SNOMED CT Concept Model to Represent Value Sets for Head and Neck Cancer Documentation

Anne Randorff HØJEN[a,1], Dorthe BRØNNUM[b], Kirstine Rosenbeck GØEG[a] and Pia Britt ELBERG[a]
[a]*Department of Health Science and Technology, Aalborg University, Denmark*
[b]*Department of Oncology, Aalborg University Hospital, Denmark*

**Abstract.** This paper presents an analysis of the extent to which SNOMED CT is suitable for representing data within the domain of head and neck cancer. In this analysis we assess whether the concept model of SNOMED CT comply with the documentation needed within this clinical domain. Attributes from the follow-up template of the clinical quality registry for Danish Head and Neck Cancer, and their respective value sets were mapped to SNOMED CT using existing mapping guidelines. Results show that post-coordination is important to represent specific types of value sets, such as absence of findings and severities. The concept model of SNOMED CT was found suitable for representing the value sets of this material. We argue for the development of further mapping guidelines for consistent post-coordination and for initiatives that demonstrate use of this important terminological feature in actual SNOMED CT implementations.

**Keywords.** Clinical terminology, SNOMED CT, mapping, Clinical informatics, Hospital information systems, computerized medical record, Research infrastructure and EHR data reuse.

## 1. Introduction

The use of clinical information systems in health care holds the promise of better use of data. An essential prerequisite for the exchange of clinical data between local health records and national registries for clinical quality is structured and standardized data representation [1]. The clinical terminology SNOMED CT is the world's most comprehensive clinical terminology, aimed at supporting machine-readable, consistent and meaningful representation of clinical data. An important feature of SNOMED CT is post-coordination, which enables composition of concepts to form new meanings. The concept model of SNOMED CT supports meaningful post-coordination, by specifying how concepts can be combined, refined and qualified [2]. However, post-coordination in SNOMED CT is challenging [3]. For example, allowing post-coordination increases the likelihood of representing identical clinical ideas inconsistently i.e. the inter-rater disagreement increases. In addition, the editorial guidelines associated with post-coordination are even by experts perceived as difficult to learn as they do not provide enough guidance to ensure consistent representation. If

[1] Corresponding Author: arra@hst.aau.dk

post-coordinated expressions are to be used routinely, it requires a realistic view on this aspect in real life mapping projects. Secondly, it requires strong guidance preferably with reusable semantic patterns.

To explore the use of post-coordinated expressions, we took our point of departure in the national clinical quality registry for Danish Head and Neck Cancer (DAHANCA), which consists of a set of forms for capturing information related to treatment of head and neck cancer. This information is used for both quality monitoring and research. An important part of the DAHANCA-documentation is the follow-up template, which contains information on tumor control and late effects of radiation therapy.

## 2. Methods

### *2.1. Material*

The follow-up template from DAHANCA consists of a total of 19 attributes with a total of 56 values. IHTSDO's SNOMED CT browser was used to conduct the map (v20150731) of the attributes as well as the related values.

Table 1 provides an overview of 19 attributes in the follow-up template and the type of values related to each attribute. The Attribute term is translated from Danish.

**Table 1.** L= List, B= Boolean (yes/no), T= Text field, S=Severity scale, S*= context-specific severity scale

| Attribute term | Value set | Attribute term | Value set | Attribute term | Value set |
|---|---|---|---|---|---|
| Tumor status | 7 (L) | Hoarseness | 5 (S*) | Neurological symptoms | 2 (B) |
| Follow-up status | 4 (L) | Laryngeal edema | 4 (S) | Osteoradionecrosis | 2 (B) |
| Smoking | 2 (B) | Atrophic mucosa | 4 (S) | Tracheostomy | 2 (B) |
| Weight | 1 (T) | Subcutaneous induration/fibrosis | 4 (S*) | Laryngectomy | 2 (B) |
| Probe | 2 (B) | Perichondritis | 2 (B) | Other late reaction | 1 (T) |
| Dysphagia | 5 (S*) | Increased caries | 2 (B) | Location and histology of other cancer | 1 (T) |
| Dryness of mouth/throat | 4 (S) | | | | |

### *2.2. Mapping approach*

The mapping guidelines described in [4] was applied. Consequently, an important part of the mapping strategy was to ensure high mapping coherency to support retrieval based on the subtype-hierarchy of SNOMED CT. This was done by ensuring close subtype relations between an attribute and its values.

The mapping was performed by a single subject matter expert experienced within DAHANCA documentation. The subject matter expert had also received SNOMED CT training prior to this study (2nd author). The mapping was subsequently reviewed by a SNOMED CT expert (1st author).

## 3. Results

All 19 attributes were mapped to SNOMED CT: 17 were represented by pre-coordinated concepts and 2 were represented by post-coordinated expressions.

51 items of 56 (91%) were mapped to expressions in SNOMED CT. Of the 51 items, 30 items (59%) were represented by post-coordinated expressions and 21 items (41%) were represented by pre-coordinated concepts. Table 1 shows the mapping results divided by the type of value set, and presenting the number and type of post-coordination applied. The few items which could not be mapped, we represented as local codes to be added to a local SNOMED CT extension.

### *3.1. Result of mapping List Items*

**Table 2.** Results of mapping the DAHANCA follow-up template. Mapping results are presented by value-set type and types of representation (pre- vs. postcoordination). In situations of postcoordination the involved attribute and range is presented.

| Attribute (model) term | Coverage | Pre-coordination | Post-coordination/ Type of post-coordination | |
|---|---|---|---|---|
| **List items** | | | | |
| Tumor status | 57% (4/7) | 3 | 1 | By combination |
| Follow-up status | 75% (3/4) | 3 | 0 | N/A |
| **Boolean values** | | | | |
| Smoking | 100 % (2/2) | 2 | 0 | |
| Neurological symptoms | 100% (2/2) | 1 | 1 | Finding context, known absent |
| Probe | 100% (2/2) | 1 | 1 | Finding context, known absent |
| Osteoradionecrosis | 100% (2/2) | 1 | 1 | Finding context, known absent |
| Tracheostomy | 100%( 2/2) | 1 | 1 | Finding context, known absent |
| Laryngectomy | 100 % (2/2) | 2 | 0 | |
| Perichondritis | 100% (2/2) | 1 | 1 | Finding context, known absent |
| Increased caries | 100% (2/2) | 1 | 1 | Finding context, known absent |
| **Severity Scales** | | | | |
| Dysphagia* | 100% (5/5) | 0 | 5 | Finding context, known absent (1)<br>Severity (4) |
| Dryness of mouth/throat | 100% (4/4) | 0 | 4 | Finding context, known absent (1)<br>Severity (3) |
| Hoarseness* | 80% (4/5) | 3 | 1 | Course (1) |
| Laryngeal edema | 100% (4/4) | 0 | 4 | Finding context, known absent (1)<br>Severity (3) |
| Atrophic mucosa | 100% (4/4) | 0 | 4 | Finding context, known absent (1)<br>Severity (3) |
| Subcutaneous induration/fibrosis* | 100% (4/4) | 0 | 4 | Finding site (4)<br>Finding context, known absent (1)<br>Severity (3) |
| **Text Fields** | | | | |
| Weight | 100% (1/1) | 1 | 0 | N/A |
| Other late reaction | 100% (1/1) | 1 | 0 | N/A |
| Location and histology of other cancer | 100% (1/1) | 0 | 1 | Finding site<br>Associated with |

### 3.2. Post-coordination patterns

Of the 8 Boolean value sets, which are present in the material, 2 situations were represented by pre-coordinated SNOMED CT concepts. In 6 situations a postcoordination was applied to obtain a full semantic map. In all of these cases the attribute | Finding context | and the destination concept | Known absent | was applied, which comply with the IHTSDO specification for representing negations.

The material includes 6 examples of Severity Scales, and in 5 of these the same pattern for post-coordination was applied, hence using two attributes and associated range concepts. Table 3 shows the typical mapping pattern for severity scales included in this material. Both types of attribute relationships comply with the SNOMED CT concept model. The sixth example was the value set related to the term/attribute Hoarseness, which consisted of 5 items which included a set of context-specific severity values. None of these could be interpreted as any of the general severities available in SNOMED CT.

**Table 3.** A typical pattern for mapping of severity scales includes an item for the absence of the finding and a set of severities. * this value was only used in a single situation.

| Attribute | Destination concept |
|---|---|
| \| Finding context \| | \| Known absent \| |
| \| Severity \| | \| Mild \| |
| \| Severity \| | \| Moderate \| |
| \| Severity \| | \| Moderate to severe \|* |
| \| Severity \| | \| Severe \| |

### 3.3. Non-matches

A total of 5 items could not be mapped. As shown in Table 4 the common characteristics of these items is that they include content parts which are context-specific (see underlined phrases).

**Table 4.** Overview of terms, which could not be mapped to SNOMED CT. The parts of the content which complicates mapping are underlined.

| Model term | Value term – not map-able |
|---|---|
| Tumor status | Never tumor-free, <u>despite treatment</u> |
| | Treatment for recurrence, <u>active disease</u> |
| | Treated for recurrence. <u>No current sign of</u> recurrence |
| Follow-up status | Control <u>elsewhere</u> |
| Hoarseness | <u>Occasional</u> hoarseness |

## 4. Discussion

The overall mapping guidelines and principles for mapping described in [4] were found useful as point of departure for this study. However, as revealed by the results of this

study the guidelines can be refined to include guidance on consistent use of post-coordination for different types of value sets.

For SNOMED CT to represent value sets, such as Boolean value sets and Severity Scales, this study demonstrates the importance of post-coordination. This result is similar to a study conducted on data entry forms in vasculitis research, which also showed the need for post-coordination to represent the semantics of a complex clinical expression or for contextual qualification [5]. However, in a survey of SNOMED CT implementations from 2013, it was found that the use of post-coordination within SNOMED CT implementations was limited [6]. Of 13 SNOMED CT implementations 6 did not support any use of post-coordination, two supported post-coordination for the terminology team only, four used post-coordination to enable limited qualification for end-users, and only one academic implementation project supported post-coordination for both refinement and qualification for end users. The study also emphasizes the issue of post-coordination as a challenge for two reasons: The design of user interfaces and the terminological challenge of implementing features that only enables terminological meaningful expressions in accordance with the SNOMED CT concept model.

## 5. Conclusion

With the acknowledgement of SNOMED CT as a core clinical terminology worldwide and an increased awareness of the potential and importance of the compositional nature of the terminology, we want to emphasize the need for software that utilizes the terminology to deliver enhanced software services to support clinical as well as administrative purposes. Therefore, we encourage future SNOMED CT implementations to explore options for enhanced use of post-coordination from the perspective of interface design as well as with respect to enhanced storage models and retrieval features.

## References

[1] D. Kalra and B. Blobel, Semantic interoperability of EHR systems, *Studies in health technology and informatics*, **127** (2007): 231-245.

[2] R. Cornet, Definitions and qualifiers in SNOMED CT, *Methods of information in medicine*, **48**.2 (2009),178-183.

[3] R. Cornet, M. Nyström and D. Karlsson, User-directed coordination in SNOMED CT, *MedInfo*. **192** (2013), 72-76.

[4] A. R. Højen and K. R. Gøeg, Snomed CT implementation. Mapping guidelines facilitating reuse of data, *Methods Inf. Med*., **51** (2012), 529–538.

[5] R. L. Richesson, J. E. Andrews, and J. P. Krischer. Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research, *Journal of the American Medical Informatics Association*, **13**.5 (2006), 536-546.

[6] D. Lee, A survey of SNOMED CT implementations, *Journal of biomedical informatics*, **46**.1 (2013), 87-96.