

Analysis of SNOMED ‘bleeding’ concepts & terms

Jonathan Bona¹, Selja Seppälä², Werner Ceusters¹

¹Department of Biomedical Informatics, University at Buffalo, Buffalo, NY

²Department of Health Outcomes and Policy, University of Florida, Gainesville, FL

Abstract— We present an analysis of SNOMED CT ‘bleeding’ concepts – those concepts with descriptions that include ‘hematoma’, ‘hemorrhage’, or ‘bleeding’; or that are descended from ‘Bleeding (finding)’ in the Is-a hierarchy; or that have Hematomas or Hemorrhages as their associated morphology – to assess how consistently they are used in the ontology.

Keywords— SNOMED CT, terminology, ontology

I. INTRODUCTION

SNOMED CT is increasingly used as reference terminology to encode clinical records in electronic health record systems. In order to realize their potential and, crucially, avoid causing lethal mistakes, such systems must support the creation and use of records that represent, as fully and accurately as possible, the state of the world as it pertains to the patient’s health. So long as they do not, clinicians and other humans using these systems will be unable to access detailed, accurate, and up-to-date information that may be essential to preserving patients’ health and lives. Further, the systems themselves will be unable to provide truly useful decision support.

It is essential that SNOMED CT be properly understood by its users and software vendors. As noted by He *et al.* [1], many such users do not have the expertise required to make proper use of the terminology. SNOMED distinguishes between concepts described with seemingly identical terms, which may lead non-expert users to use the wrong concept when encoding medical records. Other issues, such as errors in the synonyms associated with each concept may lead to unwanted concept selection with potentially harmful consequences. To prevent such errors it is important to assess SNOMED’s quality and provide automated methods to address potential terminological issues and discrepancies with respect to the underlying logical definitions.

This work follows previous studies proposing quality review and auditing methods of SNOMED CT’s concepts with a focus on their descriptions. Nash [2] examined “concepts in the Procedure hierarchy of SNOMED CT” to reveal nonsynonymous synonyms. He *et al.* [1] “evaluate and categorize aspects of concept descriptor issues across SCT from a practical use perspective.” Bodenreider *et al.* [3] “apply lexical knowledge to the analysis of biomedical terminologies, with the aim of assessing the consistency of a terminology.” We present the results of a preliminary analysis focused on “bleeding” concepts – those with descriptions that include the terms ‘hemorrhage’, ‘bleeding’, and ‘hematoma’, or that are related in the ontology to one or more key ‘bleeding’ concepts. Our goal is to assess how consistently the descriptions composing these terms match the logical specifications of the corresponding concepts.

II. SNOMED CLINICAL TERMS (SNOMED CT)

SNOMED CT is a reference terminology for the clinical domain centered around concepts [4]. As a reference terminology, it includes machine-readable description logic definitions of its concepts that can be used for logical inference, for example, in decision-support systems. SNOMED concepts are organized into a hierarchy of ‘Is-a’ relations. Each concept is associated with one or more short textual descriptions, including a primary description known as the Fully Specified Name (FSN) and synonyms (SYNs). Each FSN ends with a *semantic tag* in parentheses to help disambiguate it from other concepts with similar descriptions. Many concepts are linked through non-hierarchical associative relations such as ‘has finding site’. For example, ‘**Gastrointestinal hemorrhage (disorder)**’ has as its finding site ‘**Gastrointestinal tract structure (body structure)**’, and has as its associated morphology ‘**Hemorrhage (morphologic abnormality)**’.

III. ANALYSIS

We analyzed SNOMED CT concepts whose descriptions use the terms ‘hematoma’, ‘hemorrhage’, and ‘bleeding’ to see how the terminological use of these words and their lexical variants in concept descriptions reflects real ontological distinctions in the concept model. We assembled a set of SNOMED CT concepts (“*bleeding-concepts*”) and terms related to bleeding, focusing on *findings* and *disorders*, and excluding other types of concepts (e.g. *procedures*). For each bleeding-concept, we also gathered any body structure concept (descendant of ‘**Body structure (body structure)**’) asserted to be its finding site, along with any associated morphology concepts descended from ‘**Hematoma (morphologic abnormality)**’ or ‘**Hemorrhage (morphologic abnormality)**’. This set of concepts and terms was organized into two tables, described below.

The TERMS table was constructed using the SNOMED CT Descriptions and Relationships tables, by collecting (1) all terms for every concept that has a term with ‘bleeding’, ‘hematoma’, or ‘hemorrhage’ as a substring; (2) all terms for all concepts that have either ‘**Hematoma (morphologic abnormality)**’ or ‘**Hemorrhage (morphologic abnormality)**’ as their associated morphology; and (3) all terms for all concepts that are descendants of ‘**Bleeding (finding)**’ in the Is-a hierarchy.

The TERMS table has a row for each of the items collected in (1) – (3). Every entry in the table has values in columns for the *concept ID* and *term*, as well as columns indicating whether the *concept is active*, and whether the *description is active*. Every row that is specifically about a description has entries for its *semantic tag*, as well as information about which, if any, of our three seed terms (‘*bleeding*’, ‘*hematoma*’, ‘*hemorrhage*’) – or their variants – appears in the description. For this purpose, we manually curated a list of variants with in-use alternative

spellings or abbreviations for our seed terms ('hemorrhage', 'haemorrhage', 'haemorrhage', etc.). Rows for associated morphologies have columns for the associated morphology concept ID, its FSN and semantic tag, and an indication of whether it is a descendant of either 'Hemorrhage' or 'Hematoma'. Rows for finding sites have the ID, FSN, and semantic tag of the finding site concept. The CONCEPTS table, derived from the TERMS table, has one row for each concept in TERMS with columns for the concept ID, description, semantic tag, concept active/inactive, and whether that concept is marked as duplicate. It includes fields with information about which descriptions for each concept match our seed terms or variants. For each concept, there are associated morphology counts for descendants of both 'Hemorrhage' and 'Hematoma', and similar values for finding sites associated with the concept.

IV. RESULTS

Using these tables we assessed whether the terminological use of the words 'hematoma', 'hemorrhage', and 'bleeding' – or any lexical variant thereof in SNOMED CT terms – correlates with the ontological structure according to which the concepts described by means of these terms are organized. To that end, we classified the concepts into eight classes depending on which of the three words were used, alone or in combinations, in any description for the concept. The classes, accounting for all possible combinations, are labeled in Table 1 as follows: *N*: None; *B*: 'bleeding'; *T*: 'hematoma'; *R*: 'hemorrhage'; *BT*: 'bleeding' and 'hematoma', etc.

We also classified the concepts into four other categories depending on whether they are asserted in SNOMED CT as being descendants of (1) 'Hematoma', (2) 'Hemorrhage', (3) 'Bleeding', or (4) are related to a descendant of 'Bleeding' by means of any associative relation. The results in Table 1 are restricted to those concepts whose FSNs have semantic tags 'disorder', 'finding', or 'morphologic abnormality'.

Table 1 shows that the terms 'bleeding' and 'hemorrhage', synonyms at first sight, do not behave as such in SNOMED CT, as they appear either in isolation or in combination in many cases (see columns (3)-No-Act/Inact and (4)-Yes-Inact for *B*, *R*, and *BR*). Surprisingly, many concepts in which these terms appear are not bleedings. However, that is because the terms can be used in negations, such as 'gastric ulcer without bleeding', or in procedures that involve bleeding control, such as 'Control of postoperative hemorrhage of bladder (procedure)'.

From Table 1, it is also clear that the term 'hematoma' (classes *T*, *TR*) is used in a very precise way. In only three cases out of the 213 in which the term is used, the corresponding concept is not a descendant of the class 'Hematoma'. Every concept that is a descendant of that class includes the term 'hematoma' in one or more of its descriptions (see column (1)-Yes-Act for *T* and *TR*). Furthermore, only eight hematoma concepts also include the term 'hemorrhage' in their descriptions, including three that are not hematomas ontologically. For instance, the concept 'Hemorrhage into ovary (disorder)', a descendant of 'Bleeding (finding)', includes as an active synonym 'Ovarian hematoma'. The other two non-hematomas with hematoma synonyms are the concepts 'Epidural intracranial hemorrhage (disorder)' ('*EDH - Extradural haematoma*'), and 'Retrolbulbar hemorrhage (disorder)' ('*Retrolbulbar haematoma*'). In sum, this analysis reveals discrepancies between the terminological descriptions and the ontological structure.

V. CONCLUSIONS

We have collected a set of SNOMED CT concepts with terms related to bleeding, along with information about each of their descriptions, their placement in the concept hierarchy relative to high-level concepts like 'Bleeding (disorder)' and 'Hematoma (disorder)', and related finding sites and associated morphologies. A preliminary analysis provides some insight into how these words are systematically used as parts of descriptions, and reveals a small number of aberrations that might contribute to the creation of misleading health record entries. Further analysis based on this work is ongoing.

REFERENCES

- [1] He, Zhe, Michael Halper, Yehoshua Perl, and Gai Elhanan. "Clinical clarity versus terminological order: the readiness of SNOMED CT concept descriptors for primary care." In *Proceedings of the 2nd international workshop on managing interoperability and complexity in health systems*, pp. 1-6. ACM, 2012.
- [2] Nash, Shelly K. "Nonsynonymous Synonyms: Correcting and Improving SNOMED CT®." In *AMIA Annual Symposium Proceedings*, vol. 2003, p. 949. American Medical Informatics Association, 2003.
- [3] Bodenreider, Olivier, Anita Burgun, and Thomas C. Rindfleisch. "Assessing the consistency of a biomedical terminology through lexical knowledge." *International journal of medical informatics* 67, no. 1 (2002): 85-95.
- [4] Donnelly, Kevin. "SNOMED-CT: The advanced terminology and coding system for eHealth." *Studies in health technology and informatics* 121 (2006): 279.

TABLE I. CLASSIFICATION OF CONCEPTS WITH 'HEMORRHAGE', 'BLEEDING', 'HEMATOMA', IN THEIR DESCRIPTIONS. (ACT = ACTIVE; INACT = INACTIVE)

Class	isa_hema (1)				isa_hemorr (2)				isa_bleeding (3)				assoc_bleeding (4)			
	Yes		No		Yes		No		Yes		No		Yes		No	
	Act	Inact	Act	Inact	Act	Inact	Act	Inact	Act	Inact	Act	Inact	Act	Inact	Act	Inact
N	0	0	245	0	245	0	0	0	245	0	0	0	0	0	0	0
B	0	0	74	0	74	0	0	0	74	0	57	82	0	58	0	0
T	205	0	0	0	205	0	0	0	205	0	16	60	0	49	0	0
R	0	0	633	0	633	0	0	0	633	0	128	483	0	330	0	0
BR	0	0	41	0	41	0	0	0	41	0	7	26	0	22	0	0
TR	5	0	3	0	8	0	0	0	8	0	0	6	0	6	0	0
Totals	210	0	996	0	1206	0	0	0	1206	0	208	657	0	465	0	0