Multidisciplinary Modelling of Symptoms and Signs with Archetypes and SNOMED-CT for Clinical Decision Support

Luis MARCO-RUIZ\(^a,1\), J.Alberto MALDONADO\(^b\), Randi KARLSEN\(^c\), Johan G. BELLKA\(^a\)

\(^a\)Norwegian Centre for Integrated Care and Telemedicine
\(^b\)Instituto ITACA, Universitat Politècnica de València
\(^c\)Department of Computer Science, Faculty of Science and Technology, University of Tromsø

Abstract. Clinical Decision Support Systems (CDSS) help to improve health care and reduce costs. However, the lack of knowledge management and modelling hampers their maintenance and reuse. Current EHR standards and terminologies can allow the semantic representation of the data and knowledge of CDSS systems boosting their interoperability, reuse and maintenance. This paper presents the modelling process of respiratory conditions’ symptoms and signs by a multidisciplinary team of clinicians and information architects with the help of openEHR, SNOMED and clinical information modelling tools for a CDSS. The information model of the CDSS was defined by means of an archetype and the knowledge model was implemented by means of an SNOMED-CT based ontology.

Keywords. Clinical Decision Support, clinical modelling, archetypes, ontology, SNOMED-CT

Introduction

CDSS are software applications that help clinicians or patients when delivering health care. Previous works have documented how they improve health by reducing errors and guiding clinicians through complex care processes[1]. At this time, most modern Electronic Health Records (EHR) take advantage of CDSS and embed decision support functionalities. However, two main challenges have been identified in their adoption. The first is the ‘impedance mismatch’ among the EHR and CDSS information models [2]. The second is the knowledge management (KM) to incorporate and maintain new medical knowledge [3]. The impedance mismatch has been confronted by several works elsewhere [4]–[6]. Advances have been done in KM of computer interpretable guidelines [7] in the so called knowledge-based CDSS [3] by keeping track of the rules and knowledge artefacts that code medical knowledge. However, limitations like non-shareability, lack of standardization and lack of extensibility are still present [8]. To overcome these limitations, knowledge models should express clinical knowledge, allow extensibility and deal with constraints and...
relationships [8]. Defining the information and conceptual models using EHR standards and ontologies may allow to keep a structured annotated set of data and a meaningful representation of the concepts overcoming some of the current limitations. In [9] we described the meta-architecture for archetype-based CDSS. In the context of that project (archetype-based CDSS for patients) the concepts and information of the domain of respiratory symptoms and signs have been modelled with archetypes and ontologies. The research question we answer in this paper is how a multidisciplinary team can leverage archetypes and ontologies to model the information and knowledge of a CDSS to power its reuse and maintenance. It is important to clarify that we identify knowledge from an ontological perspective; i.e. all those clinical concepts and relationships among them identifying a meaningful dataset present during the input, execution or output of the CDSS.

1. Methods

1.1. OpenEHR and Clinical Knowledge Management

OpenEHR is an open standard for the representation and communication of EHR data. Its main feature is the dual model paradigm in which formal computable clinical models, called archetypes, are defined by constraining a generic reference model (RM). Semantics can be attached to archetypes via linkage with clinical terminologies creating a bridge among clinical information data structures and clinical concepts. Clinical information modelling and governance is carried out via collaborative tools where information architects and clinicians discuss and coordinate the development of archetypes. Currently several openEHR Clinical Knowledge Managers (CKM) exist. Some like the international openEHR CKM [10] are operating internationally generating the reference clinical concepts that are, in many cases, used as input of national CKMs defining the computable clinical concepts for their health systems. An example of national CKM is the Norwegian CKM [11] used as clinical concept development tool in the Norwegian national strategy to define the interoperable set of clinical concepts.

1.2. Information and conceptual modelling

Information and domain modelling were performed in 2 different stages. On the one hand, the information modelling aimed to define a general symptom archetype not constrained to any particular domain. On the other hand, the knowledge domain modelling was performed for the specific domain of symptoms and signs. Information modelling was performed by means of archetypes whereas the particular knowledge domain of respiratory symptoms and signs was modelled by means of SNOMED-CT and the ontology language OWL. We performed the archetype development as a collaborative effort in the Norwegian CKM departing from the international CKM Symptom.v1 archetype. We coordinated a team of reviewers composed by 5 practitioners (3 belonging to the Norwegian Healthcare System and 2 to the international CKM) and 5 information architects. Our team carried out iterative revisions over the original draft archetype where changes were recommended. During the editing process changes involved the time management, simplification and capabilities to specify several locations for the symptom were applied ending with a
maximum data set for the information structure for a general clinical use in Norway. Finally, for the CDS particular use case, the archetype was pruned, simplifying it and making some fields rely in SNOMED-CT for concepts specification rather than links to other archetypes.

To implement the ontology, the clinical literature related to general respiratory conditions was first reviewed identifying each of the terms used to indicate general symptoms or signs. As a result a list of terms was produced. In a second stage, the terms were searched in SNOMED-CT using the software Snow-Owl to find their equivalent in the standard terminology. During this search taxonomic relations, superterms and subterms referred to combinations of the identified terms (e.g. “sputum with color green” linked to “green sputum, code 277908001”) were identified and ordered according to the SNOMED-CT hierarchy. When the taxonomy was stable, it was validated by a team of 2 information architects and 1 practitioner to verify that no terms were missing. Finally, the ontology representing a maximum conceptual set of respiratory symptoms and signs was implemented with Protégé.

2. Results

2.1. The symptom archetype

The archetype defined in the Norwegian CKM was intended for a general clinical use at a national scope. After being analysed, several changes were determined to be needed to use it for the particular case of our CDSS. The changes applied have been: to rely on SNOMED-CT to specify the anatomical location of the symptom as a list of codes (see the node “Sites”); include it inside an OBSERVATION class to be able to specify the temporal nature of the symptom using the RM HISTORY class linked to it and, finally, delete associated symptoms and track them at a ontology level. The CDSS adapted archetype is depicted in Figure 1.

2.2. Respiratory symptoms and signs ontology for CDSS

A total of 126 classes have been defined in the ontology representing the maximum concepts set for respiratory symptoms and signs that can be used in our CDS
tasks, 67 of which are leaf concepts. The Most important sub taxonomies are cough, sputum, chest pain, and symptom findings related to time. For cough, 10 subtypes where identified to be representative. For chest pain, 4 subcategories (central, chest, musculoskeletal and upper chest pain) were identified, of which wall pain was specialized into 8 subcategories more. For sputum, subcategories for the color, consistency and volume were created. To specify the symptom finding related to time 6 subterms where selected to specify when the symptom debuted; i.e. days ago, months ago, years ago etc. Respiratory tract hemorrhages were also included and specialized into hemoptysis, epistaxis and hematemesis. Some specializations as cough or headache contained more terms in SNOMED-CT that those selected for the ontology. Pruning of those terms was applied considering the importance given to them in clinical practice and the capacity to represent pruned concepts by other existing concepts.

3. Discussion

In this paper we have presented an openEHR archetype and a SNOMED-based ontology modelling, the information and knowledge in a CDSS, respectively. The approach based in two detached models intends to allow information and knowledge to evolve at different rates. Thus, as the archetype is a generic information schema, the extensibility of the system relies mainly in the ontology powering scalability. For the archetype definition, a similar strategy as the proposed by Garde et al. [12] was followed in the frame of the Norwegian CKM. Practitioners from several disciplines and information architects collaborated to define a general symptom information structure and an ontology of respiratory symptoms and signs related to respiratory conditions. Finally, the archetype, originally designed for clinical use, could be adapted to the CDS needs and aligned with the ontology. Some challenges appeared when defining the strategy to align the archetype with the ontological model. A mismatch was found between the fine grained information that the archetype provides and the high abstract concepts contained in the ontology. For the conversion from one model into the other we rely on the SNOMED-CT code set in the node “Symptom name”. The allowed codes in that node are general symptom terms such as “Chest pain” or “Cough” that are inferred into more specialized concepts of the ontology by considering the values in the other nodes of the archetype; e.g. “respiratory track hemorrhage” coming from the lung can be inferred into a “hemoptysis”. Following the classification of Peleg et al. in [6] we found the following types of abstractions:

a) Classification hierarchy: the anatomical location of the symptom leads to a more specific symptom, e.g. “chest pain” in the “upper chest” can be inferred into the more specialized concept “upper chest pain”;
b) Temporal abstraction: the time pattern of appearance of a concept leads to a more specific concept, e.g. a symptom registered as being present since 15 days ago can be inferred into the concept “symptom started weeks ago”;
c) Definition of abstract terms: the presence of two concepts in the same encounter lead to one single concept in the ontology. E.g. an instance of “cough” registered with an instance of “sputum” is inferred into a “productive cough”

Another important challenge appeared when determining the level of specialization of certain concepts; e.g. chest wall pain has 8 possible subconcepts. The team tried to keep a pragmatic view based in the medical advisor’s opinion and prioritize general
concepts which may be specialized in the future updating the ontology model. Finally, we conclude that despite challenges in models alignment and mapping are still present in clinical modelling for CDSS, EHR standards and terminologies can be effectively used by multidisciplinary teams to create information and knowledge models to power KM, reuse and interoperability of CDS. In this models archetypes can provide a generally agreed information model while the ontology allows to count on a logic structure allowing long-term consistency and maintainability [13].

Acknowledgements

This work was supported by Helse Nord [grant number HST1121-13]. E. Fosse, H. Lærum, H. Leslie, I. McNicoll, N. Kolstrup and S. Bakke contributed in the review of the Symptom.v1 archetype in the Norwegian CKM. The draft archetype taken as starting point of the symptom information modelling was authored by Tony Shannon.

References