Towards Semantic Interoperability for Cardiovascular Risk Stratification into the Electronic Health Records Using Archetypes and SNOMED-CT

Alfonso Sánchez-Caro[1], Cristina Soguero-Ruiz[1], Inmaculada Mora-Jiménez[1], Luis Lechuga[3], Javier Ramos-López[1], Arcadi García-Alberola[4], Pablo Serrano-Balazote[3], José L Rojo-Álvarez[1,2]

[1] Department of Signal Theory and Communications, University Rey Juan Carlos, Madrid, Spain
[2] Prometeo, Electric and Electronic Department, Universidad de las Fuerzas Armadas ESPE, Ecuador
[3] University Hospital of Fuenlabrada, Fuenlabrada, Spain
[4] Unit of Arrhythmias, University Hospital Virgen de la Arrixaca, Murcia, Spain

Abstract

Clinical data exchange among different organizations can be of great value in the field of Cardiovascular Risk Stratification (CVRS) research. Semantic interoperability is an essential key in order to integrate and exchange medical records and to automate the clinician workflow. We present a proposal to smooth out a way towards the exchange of CVRS data, prototyped and focused on Heart Rate Turbulence (HRT), by using archetypes and ontologies. Starting from a previously developed prototype of HRT ontology, based on the conceptual model of SNOMED-CT in the electronic health records (EHR), an HRT archetype has been created, yielding an agreed, formal, and interoperable specification for representing clinical entities within the EHR. The built archetype binder with the HRT ontology achieves: (1) a standardized HRT data structure, involving domain specialists; (2) interoperability in HRT data exchange; (3) structured HRT recordings for a simple follow up by medical societies, and with statistical, research and educational purposes. We also propose a simple system to bind the HRT ontology with the archetype, hence enabling clinicians the continuos improvement and clinical research in this domain.

1. Introduction

Cardiovascular risk stratification (CVRS) constitutes a patient classification technique widely used in clinical practice, allowing cardiologists to focus resources on patients with higher cardiac morbidity and mortality risk [1]. In recent years, several cardiac risk estimators based on signal processing techniques using electrocardiogram (ECG) recordings have been proposed[2–5]. However, not all these proposals are used in clinical practice, since they have not been validated by the scientific community yet, due to the difficulty of conducting large studies, and because obtained results are mainly based on ECG records without taking into account other patient health records. Structured data from Electronic Health Records (EHR) following formal model for recording clinical information improve the CVRS decision support.

However, in the real world, different hospitals and health systems have heterogeneous information systems. Therefore, it is difficult to perform multi-center studies in which the CVRS proposed methods can be validated properly. For this reason, it is needed achieving semantic interoperability, so that the exchange of clinical information is performed with full meaning. To achieve widespread interoperability, we need a common data model, an agreed definition of complex domain concepts, and a shared vocabulary. Archetypes allow semantic interoperability when representing and storing data in EHR [6, 7].

Starting from a previously developed prototype of HRT ontology in the EHR, based on the conceptual model of SNOMED-CT [1], an HRT archetype is created here, yielding an agreed, formal, and interoperable specification for representing a given clinical entity within the EHR. This archetype compiles the ECG signal processing information of 24h Holter recordings, in order to have a standard data structure to infer the CVRS in terms of HRT parameters. Furthermore, the archetype supports the recording of risk factors such as hypertension, diabetes, age, gender or smoking, in order to provide with a complete CVRS research repository. Interoperability between heterogeneous systems in Holter recordings for HRT analysis can be obtained by designing a HRT data structure (the built archetype), where specialists from different fields work together.

The paper is organized as follows. Section 2 summarizes the state of art and presents the fundamentals about ontologies, SNOMED-CT, archetypes and HRT. In Section 3, the HRT archetype is described with detail. Conclusions are finally stated in Section 4.

2. Background

We next present the fundamentals of ontologies, SNOMED-CT, archetypes, and HRT.

2.1. Ontologies and SNOMED-CT

SNOMED-CT is the most comprehensive, multilingual clinical healthcare terminology in the world (`www.ihtsdo.org`). Its main advantage is the reference to terminologies of different health professions, hence facilitating the communication within diversified health teams and in searching health information. SNOMED-CT contributes to the improvement of patient care by underpinning the development of EHR that records clinical information in ways that enable meaning-based retrieval.

SNOMED-CT consists of a structured collection of health care terms, which are attached to concept codes with multiple definitions per code. A concept is a clinical meaning identified by a unique identifier (ConceptID) that never changes. Although it includes more than 311,000 concepts since 2011, this number cannot be enough for representing many clinical expert domains. Hence, local and national extensions can be created. Many EHR and HIS are adopting SNOMED-CT as their standard for interoperability, and some authors agree that its conceptual model is in fact an ontology. In [8], SNOMED-CT is described as ontology-orientated, which means that it represents a network of medical phenomena that is based on logic.

2.2. Clinical Archetypes

The CEN/ISO EN13606 standard (`www.en13606.org`) has been developed by CEN/TC251, the technical committee responsible for developing standards in the field of Health Information and Communications Technology in Europe, and it has also been adopted as an ISO standard. Its aim is to define an information architecture for communicating EHR between heterogeneous systems. It is based on a dual model that defines separation between information and knowledge. The information is supported by a Reference Model containing the basic entities for representing any information of a specific domain. The knowledge is supported by an Archetype Model, a structured and constrained combination of entities of a Reference Model that represents a particular clinical concept. The main advantage of this dual approach is that knowledge is upgraded when it changes, whereas Reference Model (information) remains unaltered.

OpenEHR is an international nonprofit organization that describes specifications, open source software and tools to help in the development of EHR systems. Like the CEN/ISO EN13606 standard, OpenEHR proposed a dual model. Both standards may be compatible, in fact, CEN/ISO EN13606 is a subset of the full OpenEHR specification [9].

Archetypes have been used in previous works. In [10], a system performed risk classification using medical image analysis and semantically structured information of patient data, by means of MeSH thesaurus. Improved interoperability was achieved in [11] with an ontology for chronic patients built on SNOMED-CT. In [12], EHR archetypes were integrated into a IHE XDS environment, which is an architecture specification for the implementation of shared EHR systems. Authors in [12] analyzed the information needed by health-care providers, focusing on the treatment of diabetes, and then designed 128 ISO/EN 13606 archetypes with that information. The feasibility of applying the openEHR archetype approach to model clinical information about donors and sample relate information were investigated in [13].

2.3. Basic Concepts on HRT

HRT is the phenomenon of short-term fluctuation in sinus cycle length over about 20 beats following a Ventricular Premature Complex (VPC) [14]. HRT is usually assessed from 24-h ECG signals. From such recordings, a VPC tachogram is constructed by aligning and averaging the RR interval sequences around isolated VPCs, according to the guidelines [14]. HRT is quantified using two parameters, namely turbulence onset (TO) and turbulence slope (TS). TO reflects the amount of sinus acceleration following a VPC, whereas TS reflects the rate of sinus decelerations after sinus acceleration.

3. CVRS Archetype and Binding

This section presents the proposed CVRS archetype based on HRT parameters. First, we summarize the HRT Ontology presented in [1]. Second, we explain the two archetypes needed to build the CVRS archetype. Note that terms used in both ontologies and archetypes have been validated by an expert group of clinicians. Last, we explain how to bind the HRT ontology in the archetype.

3.1. HRT Ontology

As a previous work to the creation of the HRT archetype, we developed a HRT ontology using SNOMED-CT [1]. Notwithstanding the extension of the SNOMED-CT terminology, we had to create local extensions to cover the whole HRT domain. Thus, from the 308 concepts of the HRT ontology, 19 concepts were created as local extensions.

The ontology was created using OWL language (`www.w3.org/TR/owl-features`), obtaining a prototype

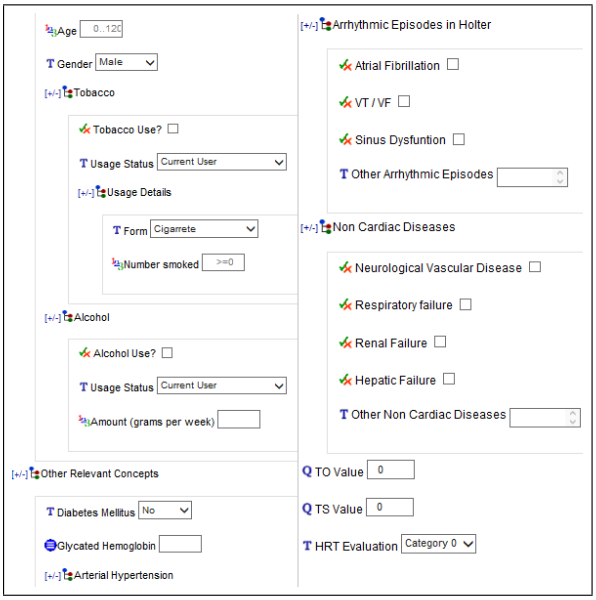

Figure 1. User interfaces for Patient Summary (left) and HRT (right) archetypes. Archetypes were created with *Archetype Editor*, and user interfaces with *Template Designer* software.

for medical records at regional and national scientific societies level. Since this prototype was not semantically interoperable, in the present work we build an archetype for sharing and exchanging HRT information and knowledge among heterogenous systems.

3.2. CVRS Archetype based on HRT

There are several software tools available for archetype modeling. Two of the most popular ones are *Archetype Editor* (`www.openehr.org`) and *LinkEHR* (`www.linkehr.com`). We explored both editors to build our CVRS archetype based on two archetypes, namely, Patient Summary Archetype and HRT Archetype.

On the one hand, the Patient Summary archetype is used to record health information useful for CVRS. It includes factors such as hypertension, diabetes, age, gender, tobacco and alcohol use, or previous heart diseases. Free text is only allowed in fields "other drugs" and "other heart disease", so that the clinician can include additional information. Figure 1 (left) shows a screenshot of the user interface for the Patient Summary archetype.

On the other hand, the HRT archetype compiles the information of 24h Holter recordings and non cardiac diseases, in order to infer the CVRS in terms of HRT parameters. Free text is only allowed in fields "other arrhythmic episodes" and "other non cardiac diseases". With TS and TO values, it is possible to infer the CVRS. Figure 1 (right) shows a screenshot of the HRT archetype user interface.

3.3. SNOMED-CT Binding

To make the archetypes built in this work semantically interoperable, the use of a terminology or an ontology is needed. The binding process associates the concepts of the archetype with standard terminology concepts. Several methods were proposed for automatic or semi-automatic binding. In [15], authors proposed an archetype editor with supports manual or semi-automatic creation of bindings between archetypes and terminology systems. However, this archetype editor has not been updated since 2008. In [8], authors proposed a method oriented to archetypes repositories. They measure the clinical content coverage of archetypes using terminology systems as a metric, in this case the SNOMED-CT terminology, to obtain information about the coverage of archetypes regarding terminology.

We propose here a solution to help clinicians to work with a previous defined ontology, so that no terminology server is needed. We build a simple system to support the binding among SNOMED-CT concepts and nodes of the HRT archetype. For this purpose, we transformed the concepts of the HRT ontology based on SNOMED-CT to an XML file which was consistent and readable by *LinkEHR*. Pre-processing of data was required, but this method makes easier to find the ConceptID (Code) for a specific concept. An example for the concept *heart rate turbulence* is presented in Fig. 2. In the bottom, the subset of concept (one concept corresponds to a "Name" in Fig. 2) associated with the HRT ontology was loaded using the built XML terminology file (SNOMED-CT HRT_subset). This provides the SNOMED-CT code for each concept, so we can copy and paste them in the binding window to link the concept with a node of the archetype without exploring the whole terminology.

4. Discussion and Conclusion

A CVRS archetype based on HRT parameters has been built to achieve interoperability among heterogenous EHR and HIS. A simple, yet helpful way of binding the HRT ontology in the archetype has been proposed. This approach facilitates clinicians the search of SNOMED-CT concepts associated with the archetypes nodes. However, our binding proposal is limited by the fact that archetype nodes that correspond with coded phrases in *dv_coded_text* object (data type containing multiple coded phrases) cannot be directly assigned to SNOMED-CT concepts in LinkEHR software. Therefore, we propose a server-based ontology system as a future line of work.

Oncoming work is devoted to apply this archetype in the daily practice to automate and streamline the clinicians workflow, with the ability to generate a complete record of a clinical patient encounter.

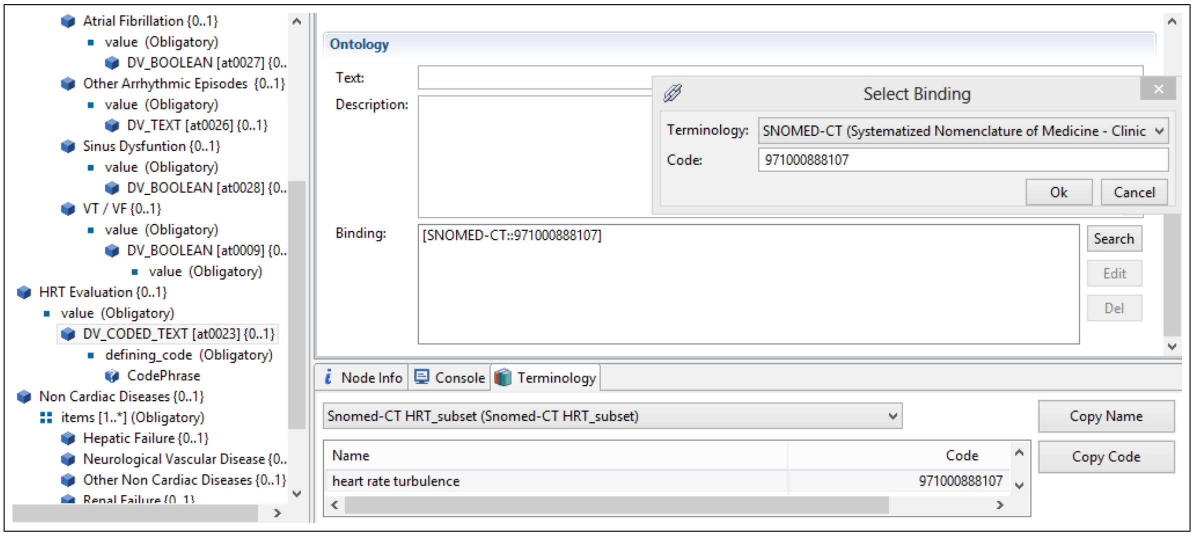

Figure 2. HRT ontology binding: screenshot of LinkEHR software. The terminology used for binding was SNOMED-CT.

Acknowledgements

This work has been partly supported by Spanish Research Projects TEC2010-19263, PRIN13_IYA12. TEC2013-48439-C4-1-R and by the Prometeo Project of the Secretariat for Higher Education, Science, Technology and Innovation of the Republic of Ecuador. CSR is supported by FPU grant AP2012-4225.

References

[1] Soguero-Ruiz C, Lechuga-Suárez L, Mora-Jiménez I, et al. Ontology for heart rate turbulence domain from the conceptual model of snomed-ct. IEEE Trans on Biomed Eng 2013; 60(7):1825–1833.

[2] Grimm W, Glaveris C, Hoffmann J, et al. Arrhythmia risk stratification in idiopathic dilated cardiomyopathy based on echocardiography and 12-lead, signal-averaged, and 24-hour holter electrocardiography. American Heart Journal 2000;140(1):43–51.

[3] Adachi K, Ohnishi Y, Yokoyama M. Risk stratification for sudden cardiac death in dilated cardiomyopathy using microvolt-level T-wave alternans. Japanese Circulation Journal 2001;65(2):76–80.

[4] Goldberger J, Cain M, Hohnloser S, et al. Heart rhythm society scientific statement on noninvasive risk stratification techniques for identifying patients at risk for sudden cardiac death. Circulation 2008;118:14971518.

[5] Klingenheben T, Ptaszynski P, Hohnloser S. Heart rate turbulence and other autonomic risk markers for arrhythmia risk stratification in dilated cardiomyopathy. J of Electrocardiology 2008;41:306–311.

[6] Martínez Costa C, Menárguez-Tortosa M, Fernández-Breis JT. Clinical data interoperability based on archetype transformation. J of Biomedical Informatics 2011;44(5):869–80.

[7] Tapuria A. Kalra D. KS. Contribution of clinical archetypes, and the challenges, towards achieving semantic interoperability for EHRs. Healthcare Informatics Research 2013;19(4):286–292.

[8] Sheng Yu Damon Berry JB. Clinical coverage of an archetype repository over SNOMED-CT. J of biomedical informatics 2012;45:408–418.

[9] Schloeffel P, Beale T, Hayworth G, Heard S, Lesliel H. The relationship between CEN 13606, HL7, and openEHR. In HIC 2006 bridging the digital divide: clinician, consumer and computer. Washington, USA: Seattle, 2006; 1–4.

[10] Doulaverakis C, Papadogiorgaki M, Vasileios Mezaris AB, et. al. Ivus image processing and semantic analysis for cardiovascular diseases risk prediction. Intl J Biomedical Engineering and Technology 2010;3:349–374.

[11] Sampalli T, Shepherd M, Duffy J. A patient profile ontology in the heterogeneous domain of complex and chronic health conditions. In Int Conf on System Sciences. Kauai, HI, 2011; 1–10.

[12] Duftschmid G, Rinner C, Kohler M, et. al. The EHR-ARCHE project: Satisfying clinical information needs in a shared Electronic Health Record System based on IHE XDS and archetypes. International journal of Medical Informatics 2013;82(12):1195–1207.

[13] J. SMBG. Applying the archetype approach to the database of a biobank information management system. International journal of Medical Informatics 2011;80(3):205–226.

[14] Bauer A, Malik M, Schmidt G, et. al. Heart rate turbulence: Standards of measurement, physiological interpretation, and clinical use: (ishne consensus). J Am Coll Cardiol 2008;52(17):1353–1365.

[15] Sundvall E, Qamar R, Nystrom M, et. al. Integration of tools for binding archetypes to SNOMED CT. BMC Medical Informatics and Decision Making 2008;8(Suppl 1):S7.