Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200156

Feasibility of Using EN 13606 Clinical Archetypes for Defining Computable Phenotypes

Archana TAPURIA,[a] Dipak KALRA,[b] and Vasa CURCIN[c]
[a] *Research Fellow, King's College London (KCL)*
[b] *Professor in Health Informatics, University College London*
[c] *Lecturer in Health Informatics, KCL*

Abstract. Introduction: Computable phenotypes are gaining importance as structured and reproducible method of using electronic health data to identify people with certain clinical conditions. A formal standard is not available for defining and formally representing phenotyping algorithms. In this paper, we have tried to build a formal representation of such phenotyping algorithm. **Methods:** We built EN 13606 EHR standard for building clinical archetypes to represent the computable phenotyping algorithm for 'diagnosis of cardiac failure'. As part of this work, we created a set of new clinical archetypes for defining 'cardiac failure diagnosis'. The EN13606 editor called Object Dictionary Client was used which was in-house developed by University College London. We evaluated the ability of EN 13606 to provide clinical archetypes to define EHR phenotyping algorithms using the predefined desiderata for the purpose [Mo et al]. **Results:** EN 13606 archetypes could represent phenotype components grouped and nested based on their logical meaning. It was possible to build the EHR phenotyping algorithm with the clinical elements and their interrelationships along with hierarchical structure and temporal criteria. But the specific mathematical calculation and temporal relations involved in the algorithm was difficult to incorporate. These will need to be coded and integrated within the clinical information system. These archetypes can be mapped for comparison with the openEHR models. Binding to external clinical terminology is fully supported. However, it does not satisfy all the desiderata defined by Mo et al. A possible way could be an approach using phenotype ontologies and its architectural representation integrated with ISO interoperability. **Conclusion**: The EN13606 archetypes can be used to define the phenotype algorithm that basically identifies patients by a set of clinical characteristics in their records. Phenotype representations defined in EN 13606 do not satisfy all the desiderata proposed by Mo et al. and thus currently has a limited ability to define the computable phenotyping algorithms. Further work is required to make the EN13606 standard to fully support the objective.

Keywords. computable phenotypes, EN13606, archetypes, EHRs, algorithms.

1. Introduction

A computable phenotype refers to a clinical condition or characteristic that can be determined with a computerised query to an Electronic Health Record (EHR) system or clinical data repository using a defined set of data elements and logical expressions (https://rethinkingclinicaltrials.org/resources/ehr-phenotyping/). Standardised

computable phenotypes can promote efficient and reliable recruitment in clinical trials across various clinical faculties and health systems.

A formal standard is not available for defining and formally representing phenotyping algorithms in a standardized way. Hence the automatic or computerized translation of phenotyping algorithms to computational implementation code is not possible. In this paper, we have tried to build a formal representation of such phenotyping algorithm using EN 13606 EHR standard. Similar work has been done to evaluate openEHR standard for building computable phenotyping algorithms [1]. This attempt shall add to the learning process.

1.1 CEN/ISO EN13606 and Clinical archetypes:

The CEN/ISO EN13606 is a European norm from the European Committee for Standardization also approved as an international ISO standard [2] intended to support the interoperability of systems and components that need to communicate EHR data: EN13606 follows an innovative Dual Model architecture. The former is structured through a reference model (RM) [3] that is an object-oriented model used to represent the generic and stable properties of health record information and the later is based on clinical archetypes [4,5]. Figure 1 shows the building blocks or classes of EN13606 RM, such as Folder, Composition, Section, Entry, Cluster and Element.

An archetype is a formal and standard way of representing the clinical data structures and their inter-relationships for a given clinical domain. They are content specifications expressed in terms of constraints on a specific reference model [6]. Archetypes specify EHR clinical data hierarchies and the kinds of data values used within each entry [7].

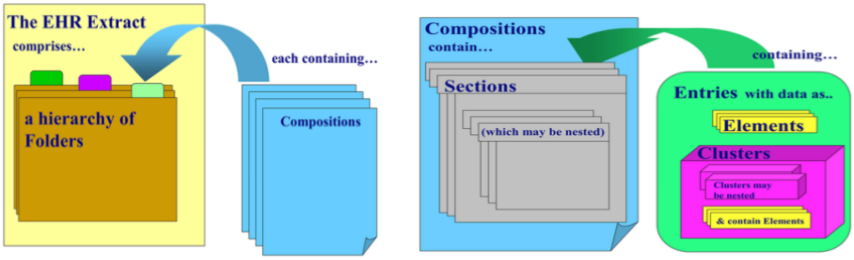

Figure 1: Structure of the reference model of EN13606 and EHR extract hierarchy [8]

2. Methods

2.1. Clinical condition/characteristic and its algorithm chosen for phenotyping:

We chose ‘cardiac failure’ as the clinical condition, diagnosis of which would need a computable phenotyping algorithm. We used a previously validated EHR phenotyping algorithm for identifying and diagnosing patients with cardiac failure. (https://www.mdcalc.com/framingham-heart-failure-diagnostic-criteria) The algorithm as shown in Fig. 2, represents clinical information that contributes towards the diagnosis

of cardiac failure. The algorithm was designed to identify patients with cardiac failure. Identification of various types of cardiac failure is not taken into consideration here.

Figure 2. Cardiac failure EHR phenotyping algorithm for the diagnosis of Heart failure

2.2. Building Clinical Archetypes:

Archetypes relevant to 'cardiac failure diagnosis' were not found in the publicly available archetype repositories or libraries like the 'Clinical Knowledge Manager' by the openEHR foundation (https://www.openehr.org/ckm/|) or 'Ocean Informatics' (https://www.openehr.org/industry_partners/ocean_informatics). As part of this work, we created a set of new archetypes for defining 'cardiac failure diagnosis' underpinned by the EHR standard EN 13606. The EN13606 editor called Object Dictionary Client (ODC) was used which was in-house developed by University College London, which is recently published, and the information models built with it are open source for use [9].

The EN 13606 Archetype describes the definitive semantic model of archetypes. It defines data types, constraints, and a reference mechanism allowing one archetype to reference another (slot). Archetypes are usually built and implemented with the help of templates, that help to combine the various archetypes together in a logical manner. EN 13606 uses the Archetype Definition Language (ADL) for building the archetypes and Archetype Query Language (AQL) for querying the data.

2.3. Desiderata for assessing clinical archetypes:

We used the desiderata published by Mo et al [10] to evaluate the ability of EN 13606 to build clinical archetypes to define EHR phenotyping algorithms. Mo et al. reviewed common features for multisite phenotype algorithms and proposed 10 desired characteristics for a flexible, computable Phenotype Record Model.

3. RESULTS

a) Archetypes created for computable phenotyping algorithm: Archetypes created were organized into a single composition with subgroups composed of number of elements (Fig. 3). Each node represents a data element, which is bound to external clinical terminologies (Read Codes, ICD-10 and SNOMED CT). Additional constraints are added to each data element including, the cardinality, optionality, datatype, value range and units for a physiological quantity, internal

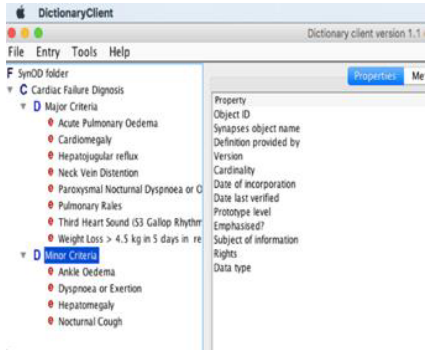

Figure 3 Clinical archetype built in ODC, an EN 13606 archetype editor

codes, etc. With EN 13606 archetypes, the components of a phenotype algorithm could be grouped or nested based on their clinical semantics into Symptoms, Signs, Investigation Results, Diagnoses, Treatment or Procedures. It can very well define the hierarchy of the data elements and the interrelationships between the data elements with the help of its Record Model (RM) properties along with the archetype slot mechanism. Still, it provides a very limited way to express the actual logic of the phenotype algorithm (especially the relational algebraic operation).

b) Dessiderata for assessing clinical archetypes: (here we have partly followed the pattern used by Papez et. Al, to represent the desiderata for assessing clinical archetypes for phenotyping algorithms) [1] Here we represent the desired desiderata [8] and where EN 13606 standard stands to satisfy it, to define phenotyping algorithms.

Table1: Summary of the feasibility of EN13606 archetypes in in defining EHR phenotyping

Desiderata by Mo. et al	**Evaluation of EN 13606 EHR standard in defining computable phenotype algorithms**
1. Human-readable and computable representation	***Fully supported***: Clinical Archetypes are defined in ADL and templates are stored as XML. EN 13606 editors can transform the archetypes (in ADL) into human-readable format such as Hypertext Markup Language (HTML) and vice versa.
2. Set operations and relational algebra	***Not supported***: Need to be coded and integrated with the EN 13606 clinical information model/system
3. Structured rules	***Not supported***: These need to be coded and integrated into the clinical information system
4. Temporal phenotype criteria	***Partially supported:*** Archetypes support temporal criteria using points in time or time intervals. Temporal relations need to be coded within the clinical information system
5. Standardized nomenclature	***Fully supported***: Supports binding of the names/values of the data elements with external standardized controlled clinical terminology or classification systems (e.g. ICD-9, ICD-10, Read Codes v2/v3, SNOMED CT)
6. External interfacing	***Partially supported***: The entire EN13606 model can be viewed in the AWB RM schema and Class tools, and comparisons can be made with classes in the openEHR reference model.
7. Backward compatibility	***Fully supported***: The tracking of changes during their development and evaluation is feasible and older versions can be referred to.

4. Discussion

Since EN 13606 archetypes define the characteristics/attributes of a set of data elements and phenotyping algorithms are primarily based on identifying patients satisfying a list of criteria (e.g. signs, symptoms, investigation results, diagnosis, etc.), archetypes could be used for defining the computable phenotyping algorithms especially for their content, hierarchical structure, interrelationships of data, and temporal criteria. EN 13606 however does not satisfy all the desiderata defined by Mo et al in defining the EHR phenotyping algorithms, especially the algebraic operations and temporal relations. Natural language processing is not provided by EN 13606 and external interfacing is limited. The example of ‘cardiac failure diagnosis’ is fairly simple but more complex diagnostic criteria for various diseases would require algebraic and temporal relations to be operated and some would demand external interfacing as well.

A possible way could be an approach using phenotype ontologies and its architectural representation. The most recent published version of ISO 13606 (late 2019) refers to this approach by integrating the ISO Interoperability and Integration Reference Architecture [2] that formalizes the relationship between clinical domains and their information representation. Without that formalization, there is a risk that archetypes will proliferate to formalize pieces of a domain in disconnected and *ad hoc* ways. Ontologies can successfully formalize the representation of a domain, and therefore be an appropriate semantic anchor for the development of archetypes. With that starting point, an implementation of computable and easily searchable phenotype definitions implemented as FHIR resources is a feasible way. Furthermore, a phenotype ontology also offers rich logical relationships including the consideration of relevant context., and the desiderata defined by Mo et al. such as the delivery of structured rules, temporal criteria, etc. could be better met by the described alternative approach. By deploying the Interoperability Reference Architecture part of ISO/EN 13606, transformation between the different representation styles including archetypes can be easily performed.

5. Conclusions

EN 13606 archetypes could be used for defining the computable phenotyping algorithms especially for their contents, hierarchical structure, interrelationships of data, and temporal criteria to identify patients with a set of clinical criteria. However, it does not satisfy all the desiderata defined by Mo et al. and hence further attempts are required to achieve the objective, especially for more complex phenotyping algorithms. Clinical archetype editors underpinned by EN13606 EHR standard might need to be used along with some other programmes to make it completely feasible to represent the phenotyping algorithms. Thus, EN 13606 currently has an ability to define the computable phenotyping algorithms but has its limitations.

References

[1] Papež V, Denaxas S, Hemingway H. Evaluating openEHR for storing computable representations of EHR phenotyping algorithms. In: 2017 IEEE 30th Int Symp Comput Med Syst. 2017. p509–514.
[2] ISO EN 13606 Standard. http://www.en13606.org/the-ceniso-en13606-standard (accessed 1 Dec 2019).
[3] ISO. BS ISO EN 13606 part 1 (No. EN 13606) (eds D Kalra & D Lloyd) 2007; 1: 1–103. Google Scholar
[4] Austin, T, Lim, YS, Nguyen, D Design of an electronic healthcare record server based on Part 1 of ISO EN13606. J Healthc Eng 2011; 2: 143–160. http://multiscience.metapress.com/content/121507.
[5] ISO. BS ISO EN 13606 part 3 (No. EN 13606) (eds D Kalra & D Lloyd) 2008; 3: 1–47. Google Scholar
[6] Health informatics – Electronic health record communication – Part 1: Reference model (ISO 13606-1:2008). https://www.iso.org/standard/40784.html (accessed 1 April 2018). Google Scholar
[7] Tapuria, A., Kalra, D., & Kobayashi, S. (2013). Contribution of Clinical Archetypes, and the Challenges, towards Achieving Semantic Interoperability for EHRs. Healthcare informatics research, 19(4), 286-92.
[8] Tapuria, A., Bruland, P., Delaney, B., Kalra, D., & Curcin, V. (2018). Comparison and transformation between CDISC ODM and EN13606 EHR standards in connecting EHR data with clinical trial research data. DIGITAL HEALTH. https://doi.org/10.1177/2055207618777676
[9] Austin, T, Sun, S, Lea, N Patterns: a simple but expressive data modelling formalism. Int J Knowledge Eng Data Mining 2016; 4(1): 74–92. Google Scholar, Crossref
[10] H. Mo, W. Thompson, L. Rasmussen, et. al, "Desiderata for clinical archetypes of EHR-driven phenotype algorithms". JAMIA, vol. 22, 2015, pp. 1220-30, doi: 10.1093/jamia/ocv112

Acknowledgement: This research is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London.