*Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth*
*A. Ugon et al. (Eds.)*


*doi:10.3233/978-1-61499-852-5-835*

# Challenges in Design and Creation of Genetic OpenEHR-Archetype

Priscila A. MARANHÃO[a] and Gustavo BACELAR-SILVA[a,b] and Duarte GONÇALVES-FERREIRA[a] and Pedro VIEIRA-MARQUES[a] and Ricardo CRUZ-CORREIA[a,b,1]
[a] *CINTESIS, Porto, Portugal*
[b] *VirtualCare, Portugal*

**Abstract.** Since the Human Genomic Project discovered the sequencing of human genome, the interest about genome content in clinical practice has increased. Genetic information has become a key point to understand diseases or improve treatments, for example, the nutrigenomic and nutrigenetics. However, the huge amount of data generated raises the need for Electronic Health Record (EHR) improvements as it becomes increasingly necessary that it includes more specific genetic information. Thus, we aim to propose standard genetic archetypes (in openEHR) and describe our main challenges in this context. We assessed 2 bibliographical databases (Pubmed and Web of science) to determine the main clinical statements needed to create the archetypes. The clinical statements were organized in archetype-concepts, and they were created in openEHR archetype editor. One archetype - genetic test results - was created from a set of genetic data and submitted to CKM repository for review. Based on the modeled archetypes, an openEHR template can be created from the proposed archetype, mainly in the nutrigenomic area, genetic labs and others related to genetic.

**Keywords.** Nutrition, Genetics, openEHR, Electronic Health Record, Archetype

## Introduction

The Electronic health records (EHR) plays a fundamental role in health care systems. When compared to paper-based records, EHRs present innumerable advantages such as completeness, timeliness, availability, legibility and accuracy [1, 2]

Several efforts have been developed to address issues regarding EHR modeling and its implementation allowing the development of future-proof EHR systems, mainly related with interoperability of health information [3]. One of the outcomes of these efforts is openEHR [4]. OpenEHR is a set of open specifications for EHR development based on two-level modelling [5]: (1) a reference model (RM) and (2) an archetype model (AM) [6]. The purpose of its design is to enable the semantic interoperability of health information between EHR systems [7]. This approach also avoids vendor lock-in [7] and differs from single-level, which IT developers represent clinical concepts by themselves [5]. In openEHR, all clinical knowledge is represented by archetypes [5]. Bacelar-Silva and Cruz-Correia (2015) [8] describes archetype as "a computable

[1] Corresponding Author: Ricardo Cruz-Correia, Porto, Portugal; E-mail: rcorreia@med.up.pt

electronic model of a concept-structured and detailed as completely as possible." The openEHR community have been modeling archetypes to represent several essential and specific contents [9-11]. The openEHR-archetypes approach allows a complete and adequate clinical practice, improving the Clinical Decision support (CDS) [11], and its increasing adoption demonstrates that more and more specifics archetypes are needed, e.g. genetic archetypes.

Since 2001, when the Human Genomic Project sequenced the human genome, the interest of this genetic content in EHRs has been increasing [12]. The discoveries into genetic testing technologies, new biomarkers test, gene-based diagnostic, and services for both rare and diseases have also increased [13, 14]. In a general way, this demand is related to personalized information, mainly with the focus on nutrigenomic, pharmacogenomics, among other genetic results [14, 15]. These new concepts promise personalized care, facilitate decision making between health professional and the patient [13], and more effective treatments, mainly to chronic disease [16]. They also aim at health outcomes for more accurate diagnosis and consequently targeted therapies, improve not just individual health but also public health policies [13].

However, Scheuner et al. (2009) [13] concluded that although the EHR have potential to enable clinical integration of genetic medicine and improve delivery of personalized care; the structured and standardized data elements, and functionality requirements are needed. Therefore, there is currently an interest to propose genetic archetypes for EHR to improve data quality [17].

This study aims to describe the process, challenges and resulting archetypes of creating an openEHR genetic test result archetype.

## 1. Methods

This work was developed in the following 5 stages.

**Review articles** - This study started with a literature review aiming to identify the main clinical concepts used in genomic studies. The literature review was carried out using Pubmed and Web of Science database to determine the articles published in this area. The search in the sources mentioned above was carried having as search query: "genetics" AND/OR "Electronic health record" AND/OR "archetype". This research was complemented with books research. A total of 51 articles were found and 26 were considered.

**Archetype-friendly concept identification** - After analyzing the articles, the authors created a set of clinical concepts, and they organized them in possible archetypes-friendly concepts.

**The Clinical Knowledge Manager (CKM) analysis** - The Clinical Knowledge Manager (CKM) is an international repository of clinical knowledge [16, 18]. It is a library of openEHR archetypes and templates. The first important point to CKM is to avoid duplicate archetypes, so, we have searched all archetypes-friendly concepts in the repository. Hence, it was needed to be included the concept name and core data items. If the archetype concept was not available in CKM, the authors could create it.

**Archetype modeling in openEHR** - The authors used Ocean Archetype Editor Software®, which is available on the openEHR site [10]. Each clinical concept modeling to archetype was discussed by the authors to the archetype to contain complete information. Finally, the new archetype was recorded in Archetype Definition Language (ADL), a formal language to express openEHR archetype [18].

**The Clinical Knowledge Manager submission for review** - The archetypes were submitted to CKM on 6th August 2017. After the submission, the openEHR editors assess if the archetypes proposed are necessary for CKM repository. After editors' approval, the archetypes are reviewed by the international clinical community, and after approval, they are made available in CKM.

## 2. Results

After a literature review was possible to identify 11 clinical statements. Our research concluded that genetic content was not available in CKM repository, and so our team developed an archetype named "Genetic Test Results" that included all statements. This archetype was modelled in English as an observation class. It aims to record genetic test results of an individual, besides presenting significant data. Figure 1 presents the proposed archetype. The data section includes:

- sample detail;
- test results;
- genetic scope;
- genetic sequence – genetic sequence file;
- genetic sequence location;
- clinical information provided;
- comment.

In Protocol section, we added the following content:

- Responsible laboratory [cluster];
- Laboratory test identifier;
- Test device details [cluster];
- Test request details;
- Laboratory scope and Extension [cluster].

In "person state section", we included confounding factors and patient state details as a cluster.

## 3. Discussion

The archetype "genetic test results" aimed to record the results and interpretation of genetic tests in an individual. It can make a significant contribution mainly in the way that health professionals will conduct care practice, which is related to have the whole information needed and organized in an objective manner. However, the clinical information should not be defined to a specific EHR, but as comprehensively as possible and that facilitates the exchange of data between different information systems [11]. It explains why we choose openEHR-archetypes; what brings a substantial benefit that is semantic interoperability, which influences the exchange of information between EHRs [4]. In this context, the advantages in modeled openEHR-archetypes are well described in the literature [19]. However, we had important challenges in this genetic modeling process as: 1- Genetic information stored does not follow any standard protocol [20]; 2 - Genetic data stored in EHR is extensive and complex, and the EHR is unable to save all genome data and analyze it [21, 22] ; 3 - Much of an EHR content is collected in free text notes, which increases the variability of the data [21].

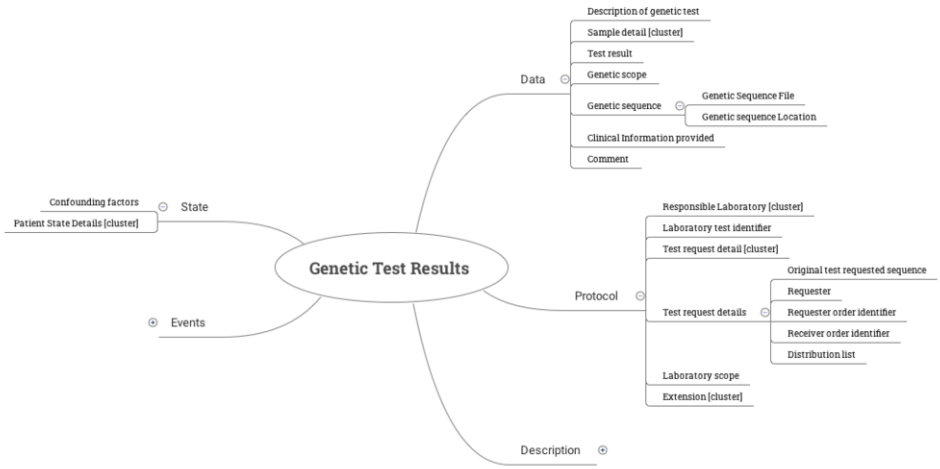


**Figure 1**. Genetic Test Result Archetype Mindmap

For this reason, one of our challenges was to analyze the data, and, only those indispensables were created in free text notes; 4 - Marsolo et al. (2013) [21] describe that quantity and quality of phenotypic data in the EHR will depend on both types of EHR and the level of customization. This point is crucial because EHR implementation decision will directly influence a data collection, mainly related to the category of EHR. These points are significant challenges which increase the difficulty to create the archetype. Very recently, Mascia et al. (2017) [17] have proposed the following genetic archetypes: ``genetic findings'', “sequence variation” among others. However, the “genetic test results” archetype is still in development. In our opinion, some data in this archetype are not related to “results” as the data of “clinical information provided”, “interpretation summary” and “Recommendations”. Besides that, “laboratory identifier"' is not enough and requires a “laboratory scope” data, to determine whether it is a diagnostic or research laboratory. In our opinion, the archetype proposed by Mascia at al. (2017) still presents points to be improved.

## 4. Conclusion

The presented paper describes the process of creating genetic openEHR-archetype, from the analysis of currently review literature. Hence, the new archetype can be incorporated in different EHR systems and help the clinicians to store genetic data and improve data quality; consequently, optimize clinical practice. As future work, we are creating a patient record the includes our archetype to test in a real-case scenario. We reinforce the importance of health professionals being involved in EHR creation together with the IT professionals, resulting in a complete and structured EHR system.

## Acknowledgement

The authors would like to acknowledge the project NanoSTIMA(NORTE-01-0145-FEDER-000016), which is financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF).

## References

[1] C. Delpierre, L. Cuzin, J. Fillaux, M. Alvarez, P. Massip, T. Lang, A systematic review of computer-based patient record systems and quality of care: more randomized clinical trials or a broader approach? Int J Qual Health Care **16** (2004), 407-416.
[2] C.J. McDonald, The barriers to electronic medical record systems and how to overcome them, J Am Med Inform Assoc **4** (1997), 213-221.
[3] Wang, L. Min, X. Lu, H. Duan, Archetype Relational Mapping Based on openEHR, Chin J Biomed Eng **33** (2014).
[4] openEHR Foundation: openEHR. An open domain-driven platform for developing flexible e-health systems. Welcome to openEHR - Homepage.http://www.openehr.org/.
[5] L. Heather, Archetypes 101, (2006).
[6] T. Beale, D. Kalra, D. Lloyd, P. Schloeffel, Introducing openEHR. In.: openEHR, (2006).
[7] L. Heather, openEHR - the world´s record, Pulse + IT **6** (2007), 5.
[8] G. Bacelar-Silva, R. Cruz-Correia, Manual de introdução à norma openEHR, 2015.
[9] G. Bacelar-Silva, R. Cruz-Correia, From clinical guideline to openehr: converting JNC7 into archetypes and template. In: XIII Congresso Brasileiro em Informática em Saúde – CBIS, (2012), 1-6.
[10] M. Braun, A.U. Brandt, S. Schulz, M.Boeker, Validating archetypes for the Multiple Sclerosis Functional Composite, BMC Med Inform Decis Mak **14** (2014), 64.
[11] D.N.S. Ronchi, D. Spigolon, D.Garcia, P.E. Cicogna, H. Bylegon, C. M. C. Moro, Desafios no desenvolvimento de prontuários eletrônicos baseados em arquétipos: avaliação fisioterapêutica funcional, Fisioter Mov **25** (2012), 497-506.
[12] D. Castle, R. DeBusk, The electronic health record, genetic information, and patient privacy, J Acad Nutr Diet **108** (2008), 1372-1374.
[13] M.T. Scheuner, H. de Vries, B. Kim, R.C. Meili, S.H. Olmstead, S. Teleki, Are electronic health records ready for genomic medicine? Genet Med **11** (2009), 510-517.
[14] D. Seo, G.S. Ginsburg, Genomic medicine: bringing biomarkers to clinical medicine, Curr Opin Chem Biol **9** (2005), 381-386.
[15] J. Hesketh, Personalised nutrition: how far has nutrigenomics progressed? Eur J Clin Nutr **67(5)** (2013), 430-435.
[16] Y.T. Joffe, C.A. Houghton, A Novel Approach to the Nutrigenetics and Nutrigenomics of Obesity and Weight Management, Curr Oncol Rep **18** (2016), 43.
[17] [C. Mascia, P. Uva, S. Leo, G. Zanetti, OpenEHR modeling for genomics in clinical practice, bioRxiv (2017).
[18] L.R. Ferguson, R. De Caterina, U. Gorman, H. Allayee, M. Kohlmeier, C. Prasad, M.S. Choi, R. Curi, D.A. de Luis, A. Gil et al, Guide and Position of the International Society of Nutrigenetics/Nutrigenomics on Personalised Nutrition: Part 1 - Fields of Precision Nutrition, J Nutrigenet Nutrigenomics **9** (2016), 12-27.
[19] B. Purin, C. Eccher, S. Forti, A real application of a concept-based Electronic Medical Record, AMIA Annual Symposium proceedings AMIA Symposium (2003), 977.
[20] D. Bosca, L. Marco, V. Burriel, T. Jaijo, J.M. Millan, A. Levin, O. Pastor, M. Robles, J.A. Maldonado, Genetic testing information standardization in HL7 CDA and ISO13606, Stud Health Technol Inform **192 (**2013), 338-342.
[21] K. Marsolo, S.A. Spooner, Clinical genomics in the world of the electronic health record, Genet Med **15** (2013), 786-791.
[22] A.N. Kho, L.V. Rasmussen, J.J. Connolly, P.L. Peissig, J. Starren, H. Hakonarson, M.G. Hayes, Practical challenges in integrating genomic data into the electronic health record, Genet Med **15** (2013), 772-778.