

Interoperability Standards for Data Sharing as a Basis to Fill in a Tailored EHR for Undiagnosed Rare Diseases

Norbert MAGGI^{a,b}, Ariam BOAGLIO^a, Carmelina RUGGIERO^{a,c},
Roberto FANCELLU^{d,e}, Francesco COCCHIARA^{e,f}, Domenico COVIELLO^b,
Deborah CAPANNA^g and Mauro GIACOMINI^{a,c1}

^a*Department of Informatics, Bioengineering, Robotics and Information Systems,
University of Genoa, Genoa, Italy*

^b*Laboratory of Human Genetics, IRCCS Giannina Gaslini, Genoa, Italy*
^c*Healthropy S.r.l., Savona, Italy*

^d*Unit of Neurology, IRCCS Ospedale Policlinico San Martino, Genoa, Italy*

^e*Clinical Centre for Orphan Diagnosis Patients, IRCCS Ospedale Policlinico San
Martino, Genoa, Italy*

^f*Unit of Endocrinology, IRCCS Ospedale Policlinico San Martino, Genoa, Italy*
^g*Comitato i Malati Invisibili, Genoa, Italy*

Abstract. Undiagnosed rare diseases include diseases with a well-characterised phenotype, diseases with unknown molecular causes or due to non-genetic factors, and pathological condition that cannot be named. Several initiatives have been launched for healthcare of patients with undiagnosed rare diseases. A project for development of medical records with special reference to the HL7 standards is being carried out in Genoa (Italy), taking into account regional and national regulations. The project is based on the integration of functionality related to patient diagnostics, taking into account omic sciences for disease prevention and risk assessment. Considering the evolution of standards, the use of FHIR is being considered in order to increase the elasticity of the system also in view of foreseeable adoption of this standard by the Italian healthcare system.

Keywords. Undiagnosed Rare Diseases, interoperability, standards.

1. Introduction

The Undiagnosed Rare Diseases (URDs) are conditions that describes people with a range of disorders and/or disabilities, probably caused by a genetic cause or genetic predisposition, that has not been yet identified. For these people, the lack of a definite diagnosis causes significant physiological and social consequences, with considerable diagnostic and therapeutic delays.

URDs may also include diseases with a well-characterised and described phenotype, or pathological conditions but that cannot be classified by a name, and they have an unknown molecular cause or they are due to epigenetic factors that interact with

¹ Corresponding Author, Mauro Giacomini, via all'Opera Pia, 13 – 16145 Genova, Italy; E-mail: mauro.giacomini@dibris.unige.it.

environmental factors. Several initiatives at the international level have been launched, culminating in the creation in 2014 of the NIH Undiagnosed Diseases Network, which with an interdisciplinary network of seven clinical sites has begun to make a significant impact on patients with undiagnosed rare diseases [1]. Other initiatives have been launched in Italy, such as the "Malattie senza diagnosi" (undiagnosed diseases) programme launched in 2016 by the Telethon Institute of Genetics and Medicine in Pozzuoli (Tigem) and the Clinical Centre for Orphaned Diagnosis Patients set up at the IRCSS Ospedale Policlinico San Martino in Genoa, in 2017 following an agreement with the Comitato I Malati Invisibili (Invisible Patients Association). Specifically, the latter was set up with the aim of reducing social distress, compromising quality of life and increasing co-morbidity in patients whose diagnosis is uncertain, limiting the phenomenon of diagnostic 'nomadism', and reducing the high costs to the national health system of repeatedly prescribing various types of investigation in the absence of proper coordination and critical evaluation. In order to achieve these goals, an operational protocol (clinical pathway) was developed which intends to implement IT-based communication and methodological tools for patient health management. In this respect an innovative electronic clinical record was designed, focusing on the integration of different dataset from different sources.

The requirement to use different data sources makes the use of standards essential to ensure proper interoperability [2-4]. This paper aims to describe the project for the development of a medical record with special reference to the standards that allows the use of data shared by different sources to be collected in the proposed architecture. The choice of standards will also be delimited with respect to the regulations produced by regional and national legislators [5] concerning healthcare facilities and the national/regional Health Information Infrastructure (HII).

2. Methods

The rationale behind the design of the information system dedicated to the clinical centre in Genoa is based on the integration of functionalities. In order to adequately develop the functionalities related to the clinical section and to the patient diagnostics, it will be necessary to adopt standards that will allow the complete interoperability of the medical record that will be developed with respect to the Hospital Information System (HIS), the Laboratory Information System (LIS) and Radiology Information System (RIS), also with different hospital in regional panorama.

The development of omics sciences and the increased availability of specific molecular medicine data combined with innovative DNA sequencing analysis by NGS (Next Generation Sequencing) for WES (Whole Exome Sequencing) and WGS (Whole Genome Sequencing) can improve disease prevention and risk assessment. In this respect, the use and inclusion of individual omics data and common clinical data within a tailored electronic health record (EHR) will improve disease prevention, risk assessment, treatment and diagnosis. Our approach is aimed to constitute a 'patient-centric translational EHR', relevant to the deployment of translational medicine [6], and speed up the identification of the disease. This will be based on standard classifications and terminologies such as ICD, SNOMED-CT, Orphanet, HPO and other for the development of a tool capable of identifying correlations between genotypic and phenotypic information collected and described in a relational database. This integration, together with self-learning analytical systems, may also lead to results in the

development of support mechanisms for disease identification and improved accuracy of diagnosis.

Moreover, in order to achieve a better and more accurate diagnosis in a shorter time, it is envisaged that the web-based tool will allow access to specialist doctors outside the structure both for patient's clinical history consulting and to input data and observations relating to the patient. This characteristic requires data management relating to the privacy and consent of the patients, an issue that becomes even more important considering the integration of the patient's omics data [7]. In this respect, also in compliance with the provisions of the EU regulation on the processing of personal data [8], we intend to develop a privacy by design and privacy by default approach, providing from the outset the tools and the correct settings to protect personal data (Role based access control RBAC) [9]. A decision is currently being taken in coordination with the hospital authorities on whether to host the system at the hospital's data centre or to develop a cloud-based system, possibly using blockchain technology to verify the sources of and access to data [10; 11].

According to the decision of Italian legislator [5], we decide to centre our data collection system on the HL7 Clinical Document Architecture (CDA).

The use of standards will make it possible to obtain data already available in the data centres of healthcare facilities, thus filling in the data required by the physicians (Fig 1) who have commissioned this specific tailor-made EHR, achieving the goal of obtaining an early diagnosis and treatment of the patients.

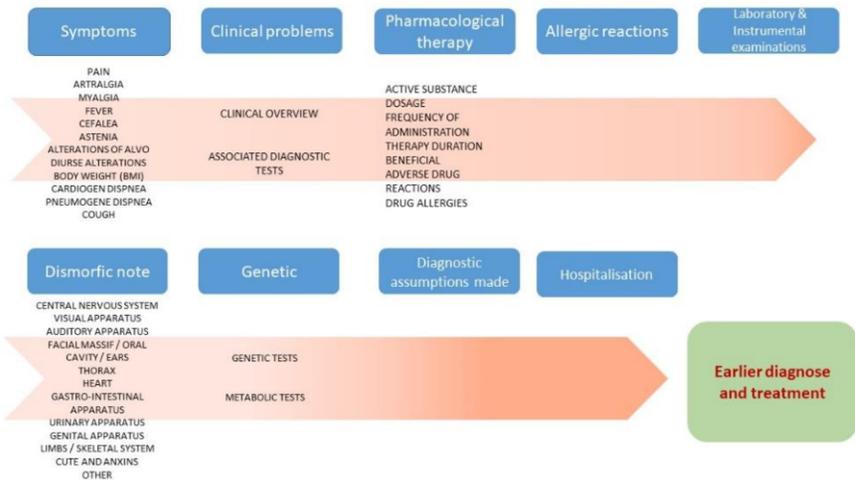


Figure 1. Patients data and processing flow for earlier diagnose EHR

3. Results

A schematic diagram of the proposed architecture is shown in figure 1, encompassing different scenarios that may arise. In a first case (a) the hospital provides the possibility of obtaining data through the use of web services, but these are not fully standardised for

the production of CDA compliant documents. In this perspective a client application can be used for translation into the necessary format to ensure interoperability between the other parts of the system. In case (b), a series of views can be obtained by means of agreements with the hospital structure. An extensive discussion with the data centre managers is necessary in order to understand the logical schema with which the data is stored in the HIS. It is also necessary to interpose a standardisation client for the generation of a CDA compliant document. In the case of external specialists (c) it will be necessary to digitise the paper documentation and to generate specific meta-data describing the documents and their contents. Most of Italian regions adopted the Retrieve Locate and Update Service (RLUS) interface [2], to allow authorised entities to feed and retrieve data to and from regional HII. The presented system has been authorised to interact with HII to extract previous data and to update it.

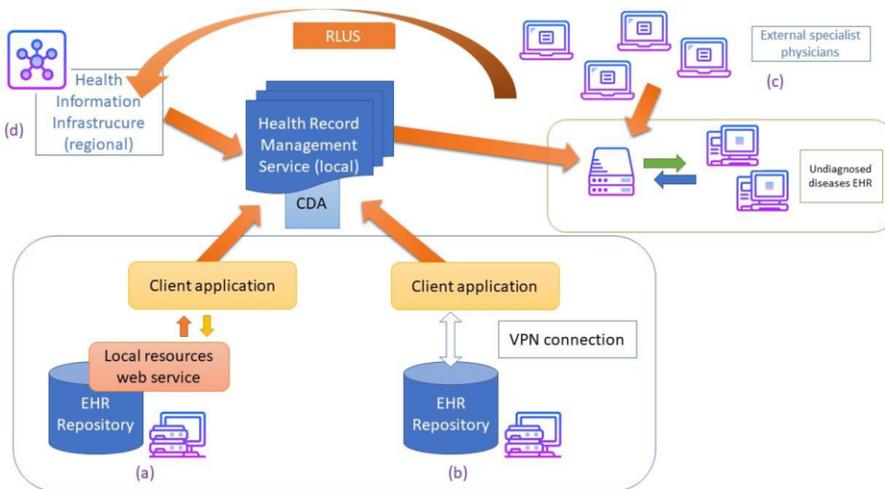


Figure 2. Outline of the proposed architecture

Moreover, since the patient will have interest in providing consent to access the data in the HII, the system will be able to access through the CDA the documents available at HII, if they contain analytically significant elements (d).

Considering the evolution of the standards, the use of FHIR will also be evaluated to increase the elasticity of the system, also in view of the adoption of this standard by the Italian legislator. In this respect, it is worth highlighting the introduction by HL7 of the genetic profile to the FHIR Observation resource, which will be taken into account in the development of the system [12] resulting in a future CDA and FHIR based architecture.

4. Discussion and Conclusions

Standards are an essential basis for ensuring interoperability between different information systems. The adoption of the various standards will allow the new system being developed to enable both management cooperation – different applications can

interact to exchange requests and results (authentication of users with hospital credentials (LDAP)), prescriptions for drugs and analyses, discharge letters (CDA v2), booking of services through unified booking centres (HL7 v2) – and clinical cooperation. Clinical information stored in applications, even remotely managed by other healthcare professionals, can be accessed promptly at the time of need, for improved patient care.

This application could be extended later with an intensive use of Natural Language Processing (NLP) to also integrate the numerous patients' documents expressed in natural language into the information generation chain.

References

- [1] Taruscio D, Baynam G, Cederroth H, Groft SC, Klee EW, Kosaki K, Lasko P, Melegh B, Riess O, Salvatore M, and Gahl WA. The Undiagnosed Diseases Network International: Five years and more!, *Molecular Genetics and Metabolism*. 2020; 129: 243-254.
- [2] Gazzarata R, Giannini B, and Giacomini M. A SOA-Based Platform to Support Clinical Data Sharing, *Journal of Healthcare Engineering*. 2017;2017:1-24.
- [3] Chervitz SA, Deutsch EW, Field D, Parkinson H, Quackenbush J, Rocca-Serra P, Sansone SA, Stoeckert CJ, Taylor CF, Taylor R, Ball CA. Data standards for Omics data: the basis of data sharing and reuse. *Bioinformatics for Omics Data*. 2011:31-69.
- [4] D'Amore JD, McCrary LK, Denson J, Li C, Vitale CJ, Tokachichu P, Sittig DF, McCoy AB, and Wright A. Clinical data sharing improves quality measurement and patient safety. *Journal of the American Medical Informatics Association* 2021;28:1534-1542.
- [5] DECRETO DEL PRESIDENTE DEL CONSIGLIO DEI MINISTRI 29 settembre 2015, n. 178. Regolamento in materia di fascicolo sanitario elettronico, in: *Gazzetta Ufficiale Serie Generale n.263 del 11-11-2015*, 2015.
- [6] Shabo A, The patient-centric translational health record, *Pharmacogenomics*, 2013 14, 349-352.
- [7] Tantoso E, Wong W-C, Tay WH, Lee J, Sinha S, Eisenhaber B, and Eisenhaber F. Hypocrisy Around Medical Patient Data: Issues of Access for Biomedical Research, Data Quality, Usefulness for the Purpose and Omics Data as Game Changer. *Asian Bioethics Review*. 2019;11:189-207.
- [8] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), in, 2016.
- [9] Gazzarata G, Gazzarata R, and Giacomini M. A standardized SOA based solution to guarantee the secure access to HER. *Conference on Enterprise Information Systems/International Conference on Project Management/Conference on Health and Social Care Information Systems and Technologies, Centeris/Projman / Hcist 2015*. 2015;64: 1124-1129.
- [10] Mehta S, Grant K, and Ackery A. Future of blockchain in healthcare: potential to improve the accessibility, security and interoperability of electronic health records. *BMJ Health & Care Informatics*. 2020:27.
- [11] Sharma Y and Balamurugan B. Preserving the Privacy of Electronic Health Records using Blockchain. *Procedia Computer Science*. 2020;173:171-180.
- [12] Alterovitz G, Heale B, Jones J, Kreda D, Lin F, Liu L, Liu X, Mandl KD, Poloway DW, Ramoni R, Wagner A, and Warner JL. FHIR Genomics: enabling standardization for precision medicine use cases. *npj Genomic Medicine*. 2020:5.