

From OpenEHR to FHIR and OMOP Data Model for Microbiology Findings

Eugenia RINALDI^{a,1} and Sylvia THUN^a

^a Charité – Universitätsmedizin Berlin, Germany

Abstract. HiGHmed is a German Consortium where eight University Hospitals have agreed to the cross-institutional data exchange through novel medical informatics solutions. The HiGHmed Use Case Infection Control group has modelled a set of infection-related data in the openEHR format. In order to establish interoperability with the other German Consortia belonging to the same national initiative, we mapped the openEHR information to the Fast Healthcare Interoperability Resources (FHIR) format recommended within the initiative. FHIR enables fast exchange of data thanks to the discrete and independent data elements into which information is organized. Furthermore, to explore the possibility of maximizing analysis capabilities for our data set, we subsequently mapped the FHIR elements to the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM). The OMOP data model is designed to support the conduct of research to identify and evaluate associations between interventions and outcomes caused by these interventions. Mapping across standard allows to exploit their peculiarities while establishing and/or maintaining interoperability. This article provides an overview of our experience in mapping infection control related data across three different standards openEHR, FHIR and OMOP CDM.

Keywords. interoperability, standard, data model, mapping, microbiology, FHIR, OMOP CDM, openEHR

1. Introduction

Under the umbrella of the Medical Informatics (MI) Initiative, the HiGHmed Consortium seeks to enhance the efficiency of clinical research and improve patient care through novel medical informatics solutions and cross-institutional data exchange [1]. In particular, Charité together with seven other German universities that are members of HiGHmed, have been involved in the use case Infection Control whose primary aim is to merge all necessary pathogen-related data and information to establish a smart infection control system.

While infection surveillance is already in place in hospitals and at national scale, it frequently suffers from a lack of data standardization and, consequently, from limited data integration, and limited availability of relevant data [2].

Regardless of the local systems used to store patient information in hospitals, standards ensure a level of interoperability that all centers can interface with. openEHR and FHIR are the most robust and complete healthcare data persistence and exchange

¹ Corresponding Author, Eugenia Rinaldi, Berlin Institute of Health (BIH), Charité – Universitätsmedizin Anna-Louisa-Karsch-Straße 2, 10178 Berlin, Germany; E-mail: eugenia.rinaldi@charite.de.

specifications that support full semantic interoperability [3]. Both standards model the clinical and administrative data based on reusable patterns that describe the medical information. These patterns are called “Archetypes” in openEHR and “Resources” in FHIR. Archetypes are maximal data sets for a given single clinical concept and are expected to contain all the clinical information [4]. Archetypes are designed to be used in “Templates” that define specific use cases.

From its conception, it was planned that HiGHmed would make use of openEHR to model the infection control-related information. On the other hand, the over-arching MI Initiative requires the use of FHIR as exchange standard among all MII Consortia. Thus, to pursue interoperability within MII, we mapped the openEHR template to the FHIR data model.

FHIR provides efficiency in information exchange, allowing access to granular patient health data along with cross-references to other related information. A consequence of the optimization for data exchange is that the data format is not designed for storage and analysis like in traditional relational databases. The nested structure of information in openEHR is also denormalized and not optimized for exploration and analysis of data. Therefore, to explore further possibilities for research analysis for our use case, we have considered the use of OMOP CDM.

The OMOP CDM consists of a collection of table schemas and offers an optimized access to information for the sharing of health research data. Each table schema represents a particular OMOP domain, which is analogous to the resources as defined in FHIR [5].

Within the activities of the HiGHmed Interoperability Work Package, we have mapped Archetypes to Resources and Resources to OMOP tables.

2. Methods

For the use case “Infection Control,” the HiGHmed Consortium members have agreed upon a minimal dataset. The dataset contains the most relevant information that should be exchanged among institutions concerning infections. It contains a selection of administrative, patient movement, and microbiology data.

The openEHR modelling group within the Consortium has modelled the agreed data set using existing and new Archetypes. In some cases, local modifications of international Archetypes were required to adapt them to the health care culture and the definition of health-related concepts in Germany [6]. The relevant Archetypes were then combined in an openEHR Microbiology Finding Template. The template was published on the Clinical Knowledge Manager, the web platform for collaborative development, management and publishing of openEHR assets.

The standard currently used at Charité to exchange laboratory data is HL7 v.2, which is also the most widely implemented healthcare standard for electronic data exchange in the clinical domain [7]. Therefore, before proceeding with our mapping of the openEHR template to FHIR, we decided to take an intermediate step and map the template to HL7 v2. Thus, we had the opportunity to better compare the data model to its application in the real environment. Moreover, FHIR represents an evolution of the HL7 v2 standard. Therefore, in the FHIR specifications, it is possible to find several suggestions or points for consideration for mapping between HL7 v2 and FHIR v4 [8]. However, most suggestions typically offer mapping to the latest version HL2.7, whereas, within Charité, version 2.3 is still widely used. Therefore, we could not base our mapping

solely on the FHIR specifications but we needed to adjust it to our specific context and to HL7 version 2.3. To do this we used tools such as Caristix [9] which documents all versions of the HL7 v2 standard.

Mapping from FHIR v4 to OMOP CDM v6 brings in new challenges as FHIR resources usually define more information for each clinical concept. The mapping was performed using the documentation found on the OHDSI website [10] where a detailed description of the OMOP tables can be found.

Mapping to OMOP tables is strictly related to the terminology used for that concept. It is therefore very important to have defined FHIR profiles with terminology bindings that can be used for the mapping to OMOP. Tools such as Athena [11] or the Implementation Guide for Common Data Models Harmonization [12] offered great support in mapping the FHIR concept to the right OMOP table.

In the microbiology finding data set, the examinations performed on samples have quite a relevant role. In openEHR these can be found nested inside different Archetypes. In FHIR all the information relevant to the examination is modelled with the Observation resource. In OMOP we first need to distinguish between standardized tests, typically laboratory test expecting a numerical result, which should be mapped to the table *Measurement*, and a general examination without a standardized test, belonging to the OMOP Observation table.

In FHIR, all the laboratory test codes to investigate the microorganisms are defined through the element *Observation.code* which was mapped to the field *Measurement_concept_id* of the OMOP Measurement table. The code/id itself is provided by the selected terminology system, for example LOINC, which both standards support.

3. Results

FHIR profiles for the microbiology dataset were modelled and made available on the Simplifier platform [13].

Culture	HL7 v.2	openEHR	FHIR 4.0	OMOP CDM 6.0
Organism presence	OBX.8	openEHR-EHR-CLUSTER.laboratory_test_analyte.v1/items[at0024]	Observation.interpretation	MEASUREMENT value as concept_id
Organism name	OBX-2,OBX-5	openEHR-EHR-CLUSTER.laboratory_test_analyte.v1/items[at0001]	Observation.value	MEASUREMENT value as concept_id
Colony count	OBX-2,OBX-5, OBX-6	openEHR-EHR-CLUSTER.erregerdetails.v1/items[at0035]	Observation.value	MEASUREMENT value as number
Colony frequency	OBX.8	openEHR-EHR-CLUSTER.erregerdetails.v1/items[at0003]	Observation.interpretation	MEASUREMENT value as concept_id
Virulence factor	OBX-2,OBX-5	openEHR-EHR-CLUSTER.erregerdetails.v1/items[at0016]	Observation.value	OBSERVATION value as string
MRO class	OBX-2, OBX-5	openEHR-EHR-CLUSTER.erregerdetails.v1/items[at0018]	Observation.value	OBSERVATION value as string
Mechanism of resistance	OBX-2, OBX-5	openEHR-EHR-CLUSTER.erregerdetails.v1/items[at0057]	Observation.value	OBSERVATION value as string

Figure. 1. Mapping Example between openEHR, FHIR, OMOP CDM

Most clinical concepts have their own detailed definition in openEHR while in FHIR we have a model for the broad concept that can be profiled according to needs. In fig.1 we can see an example of the data from the openEHR template for microbiology finding that has been mapped to FHIR and OMOP.

Archetypes are maximal data sets covering all aspects of a clinical concept, while FHIR resources offer a minimum data set which can be extended to include more information in case of need. However, for the dataset analyzed, extensions were not necessary: the correspondence of elements between openEHR Archetypes and FHIR Resources for the microbiology data analyzed had a coverage of 100%. This was a very

important result as inconsistencies between the models could pose a significant challenge for data interoperability [14]. Direct matches were also found between FHIR and OMOP for laboratory tests and Specimen information. However, we noticed that some of the challenges going from one model to the other is the cardinality of elements. For example, in OMOP the date of the investigation *Measurement_date* is mandatory whereas in FHIR *Observation.effective* is not mandatory unless otherwise specified in the profile as it was the case for infection control.

The FHIR resource for specimen has an element, *Specimen.parent*, which refers to the specimen from which this specimen originated. This parental association in OMOP was obtained using the *Fact_Relationship* table which contains records about the relationships between facts stored as records in any table of the CDM. All observations for which a LOINC code was identified were mapped to the *Measurement* table, whereas observations where no LOINC code was available, such as the German classification for Multi resistant Organisms or Mechanism of Resistance which expects a string value as result, were mapped to *Observation*. All commentary fields in OMOP are stored in a separate Note table. *DiagnosticReport* and *ServiceRequest* elements which contain laboratory report and some more organizational information did not have an actual correspondence in OMOP. This was expected as OMOP focuses on research and population level analysis and is not designed to describe laboratory workflow details. The OMOP Tables *Observation* and *Measurement* together with *Fact_relationship*, however, offered the possibility to cover also these fields as shown in fig.2.

openEHR	FHIR	OMOP CDM
COMPOSITION.report-result	DiagnosticReport	Observation
CLUSTER.case_identification.v0	Encounter	Visit_occurrence
OBSERVATION.laboratory_test_result	DiagnosticReport, ServiceRequest	Observation, Measurement
CLUSTER.specimen	Specimen	Specimen
CLUSTER.anatomical_location	Specimen	Specimen
CLUSTER.laboratory_test_analyte	Observation	Specimen, Measurement
CLUSTER.erregerdetails	Observation	Observation, Measurement
CLUSTER.laboratory_test_panel	Observation	Observation

Figure 2. openEHR-FHIR-OMOP Mapping

4. Discussion

We were able to map between the data between the three standards. However, the dataset analyzed was limited to the microbiology finding of our use case and there might be harder challenges when we consider bigger datasets. In order to fit the requirements, which in our case were defined by the openEHR template, FHIR offers the possibility to modify its resources into specific profiles or even add extensions to fit other data elements. On the other hand, FHIR resources can contain more information than the OMOP CDM supports. FHIR concentrates on the patient level and on the simplicity of the queries to fetch specific data, while OMOP focuses on the population level analysis. The information that cannot be matched is either lost or could for example be hosted in a supplementary *ad hoc* FHIR table. However, the library of standard analytic routines offered by OHDSI, is of course based only on CDM tables. There are increasing efforts to cross-reference standards and to make the mapping process automatic. LinkEHR [15] for example, is described as a tool with the functionality to transform openEHR

archetypes to FHIR R4 Observation resources. There are also initiatives such as The FHIR Project at Georgia Tech [16] for mapping between FHIR and OMOP. However, relying completely on automatic tools or existing mapping is not advisable and further reviewing is always recommended. Such tools as well as existing mapping efforts are also often bound to specific versions and might become obsolete as standards evolve. Additionally, the diversity of implementation of standards in different contexts or in different institutions remains indeed a big challenge when mapping.

5. Conclusion

For the use case Infection Control of the HiGHmed Project, mapping the data between the standards openEHR, FHIR and OMOP proved to be feasible without particular issues. This was possible also because all the above standards support the use of the standard terminologies SNOMED CT and LOINC.

References

- [1] HiGHmed | Medical Informatics Initiative, Available at: <https://www.medizininformatik-initiative.de/en/konsortien/highmed>, Accessed November 15, 2019.
- [2] Haarbrandt B, Schreiweis B, Rey S, Sax U, Scheithauer S, Rienhoff O, et al. HiGHmed – An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. *Methods Inf Med.* 2018 Jul;57(Suppl 1):e66–81
- [3] Allwell-Brown E. A Comparative Analysis of HL7 FHIR and openEHR, Available at: https://ki.se/sites/default/files/migrate/eneimi_allwell_brown_a_comparative.pdf, Accessed November 15, 2019.
- [4] Bosca D, Moner D, Maldonado JA, Robles M. Combining Archetypes with Fast Health Interoperability Resources in Future-proof Health Information Systems. *Stud Health Technol Inform.* 2015;210:180–4.
- [5] Jiang G, Kiefer R, Sharma D, Prud'hommeaux E, and Solbrig R. A Consensus-based Approach for Harmonizing the OHDSI Common Data Model with HL7 FHIR. *Stud Health Technol Inform.* 2017; 245: 887–891.
- [6] Wulff A, Sommer KK, Ballout S, HiGHmed Consortium, Haarbrandt B, Gietzelt M. A Report on Archetype Modelling in a Nationwide Data Infrastructure Project. *Stud Health Technol Inform.* 2019;258:146–50
- [7] HL7 Standards Product Brief - HL7 Version 2 Product Suite | HL7 International, Available at: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=185, Accessed November 18, 2019.
- [8] Comparison-v2 - FHIR v4.1.0, Available at: <https://build.fhir.org/comparison-v2.html>, Accessed November 28, 2019.
- [9] Caristix HL7-Definition V2 Home (HL7 v2.5.1), Available at: <https://hl7-definition.caristix.com/v2/HL7v2.5.1>, Accessed March 4, 2021.
- [10] cdm60.utf8, Available at: <https://ohdsi.github.io/CommonDataModel/cdm60.html>, Accessed March 4, 2021.
- [11] Athena, Available at: <https://athena.ohdsi.org/search-terms/start>, Accessed March 4, 2021.
- [12] Common Data Models Harmonization, Available at: <http://build.fhir.org/ig/HL7/cdmh/profiles.html#omop-to-fhir-mappings>, Accessed March 4, 2021.
- [13] Medizininformatik Initiative - HiGHmed - IC - SIMPLIFIER.NET, Available at: <https://simplifier.net/medizininformatikinitiative-highmed-ic/>, Accessed March 4, 2021.
- [14] Topaz M, Seger DL, Goss F, Lai K, Slight SP, Lau JJ, et al. Standard Information Models for Representing Adverse Sensitivity Information in Clinical Documents. *Methods Inf Med.* 2016;55(2):151–7.
- [15] LinkEHR, Available at: <https://linkehr.veratech.es/>, Accessed March 4, 2021.
- [16] The FHIR Project at Georgia Tech. GT-FHIR2, the OMOP on FHIR Project, Available at: <http://omoponfhir.org/>, Accessed March 4, 2021.