Applied Medical Informatics
Original Research Vol. 46, No. Suppl. 2 /2024, pp: S25-S28

Maximizing Research on Long COVID using FHIR and OMOP

Eugenia RINALDI[1,*], Lorenzo CANZIANI[2], Salvatore CAUTADELLA[3], Chiara DELLACASA[3], Anna GORSKA[2], Juan Mata NARANJO[3], Thomas OSMO[4], Miroslav PUSKARIC[5], Elisa ROSSI[3], and Sylvia THUN[1]

[1] Berlin Institute of Health at Charité-Universitaetsmedizin Berlin, Luisenstr. 65 10117 Berlin, Germany
[2] University of Verona, Via S. Francesco, 22, Verona, Italy
[3] Cineca Consorzio Interuniversitario, Bologna, Via Magnanelli, 6/3, 40033 Casalecchio di Reno BO, Italy
[4] Centre Informatique National de l'Enseignement Supérieur, 950 Rue de St - Priest, 34000 Montpellier, France
[5] High-Performance Computing Center, Nobelstraße 19, 70569 Stuttgart, Germany
E-mails: eugenia.rinaldi@bih-charite.de (*); lorenzomaria.canziani@univr.it; s.cautadella@cineca.it; c.dellacasa@cineca.it; anna.gorska@univr.it; j.naranjo@cineca.it; osmo@cines.fr; miroslav.puskaric@hlrs.de; e.rossi@cineca.it; s.thun@bih-charite.de

* Author to whom correspondence should be addressed

Received: 28 June 2024 /Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
In 2020 The European commission funded the ORCHESTRA project with the aim to join the efforts of several European research centers in the research around the COVID-19 disease. One of the main challenges was to harmonize data across the different cohorts and countries. The introduction of standard terminologies such as SNOMED CT or LOINC helped establish a common language within the project. Over 3500 variables from several information categories were mapped to international codes from standard terminologies. After four years since the start of the pandemic, the study of long COVID seems to be of particular relevance due to the long-term effects that some people keep experiencing even after the infection has disappeared. To facilitate this research, we selected the ORCHESTRA variables that concerned long COVID and mapped them to the standards FHIR and OMOP to possibly support further data exchange with other research organizations.

Keywords: FHIR; OMOP CDM; Long COVID; Standard; Interoperability

Introduction

The project ORCHESTRA started in 2020 in the emergency of the COVID-19 pandemic. It reunites partners from several European and non-European countries with the purpose of developing new knowledge on the novel SARS-CoV-2 virus. One of the main objectives of the project was the harmonization of data across different protocols to be able to perform analysis across the different cohorts. A dedicated group within the project worked on semantic interoperability and mapped COVID-related variables to international codes stemming from standard terminologies [1,2] such as SNOMED CT, LOINC, NCIt, ICD and ATC. Four years after the start of the pandemics, international research is focusing on the long-term effects of the COVID-19 disease, the so called long COVID [3]. To facilitate research, we selected the ORCHESTRA variables relevant for long COVID and mapped them to standards such as FHIR [4] and OMOP [5] which are commonly used for data exchange.

Materials and Methods

We started from the data dictionary of the ORCHESTRA project [6], where over 3500 variables and their answer options have been mapped to international standard terminologies. The data dictionary is a file where the variables used for data collection are listed together with their answer options, and additionally, the information category assigned by the research group to the variable can be retrieved. Starting from this file, we selected the variables from the information categories "long COVID", "new medical events", "COVID severity". Additionally, we also extracted variables relating to the category "demographics" which included information on comorbidities. We removed cohort-specific information (e.g. "record Id") as well as doubles, and started mapping the selected variables to FHIR.

For some of the variables, such as for those relating to problems that persist in time after the infection, mapping was possible to both the FHIR resources "Observation" and "Condition". In these cases, we opted for the resource "Observation".

To proceed with the second step, and map the variables also to the OMOP standard, we utilized the online instrument ATHENA offered by OHDSI which helps identify the correct OMOP tables based on the terminology concepts. For this procedure, we utilized the mapping to the international standard terminology codes available within the ORCHESTRA project.

Results

The first result was the selection of the variables that could be particularly relevant for long COVID and that should be considered for the mapping to FHIR and to OMOP. We obtained a total of 104 variables across 5 different information categories. The 104 elements were mapped to the appropriate FHIR resources which turned out to be Observation and Condition as shown in Table 1.

Table 1. Distribution of the variables in the mapping to the FHIR resources Condition and Observation

Information category	FHIT Condition	FHIR Observation	Total
Comorbidities	2	7	9
COVID-19 severity	0	6	6
Demographics	0	7	7
Long COVID	5	18	23
New medical event	46	13	59
Total	***53***	***51***	***104***

The 104 variables were further mapped to the OMOP tables. As shown in Table 2, the tables resulted to be Condition, Observation and Measurement.

Table 2. Distribution of the variables in the mapping to the OMOP tables Condition_occurence, Measurement and Observation

Information category	OMOP Condition_occurence	OMOP Measurement	OMOP Observation	Total
Comorbidities	5	1	3	9
COVID-19 severity	0	0	6	6
Demographics	0	0	7	7
Long COVID	16	0	7	23
New medical event	47	4	8	59
Total	***68***	***5***	***31***	***104***

Discussion

Most of the selected ORCHESTRA variables concern examinations, symptoms and diseases and were mapped to the FHIR resources "Observation" and "Condition". However, there are cases where both resources could have been used. For example, symptoms, which are usually temporary problems, are generally modelled with the FHIR resource "Observation". However, in the case of long COVID, symptoms persist in time and could be handled as long-lasting problems and therefore mapped to the FHIR resource "Condition". However, we decided to map them to "Observation" to facilitate comparison with initial symptoms. OMOP distinguishes between a systematic examination with a standardized procedure (Measurement) and a general evaluation of fact (Observation) and this is why scores and indexes were mapped to the OMOP table "Measurement". Interesting is the case of the Child- Pugh score where some items belong to the OMOP domain Observations and others to the Measurement domain due to the different types of evaluation involved. All symptoms are mapped to the domain "Condition" in OMOP and this results in a higher number of variables mapped to the corresponding table as shown in Table 2.

The selection of relevant ORCHESTRA variables is not exhaustive and could certainly be extended to include, for example, further information about comorbidities or the questionnaires about the psychological effects of the pandemic. This study represents only an initial attempt to show how data exchange for long COVID could be supported using interoperability standards.

Conclusions

The need to efficiently exchange research information at international level to advance research has been highlighted by the COVID-19 pandemic. The broad adoption of interoperability standards for terminology and format of data in research communities would strongly support this goal. In Europe, the European Health Data Space [7] is expected to use FHIR to smooth the flow of data from primary care to secondary use.

List of Abbreviations: Not applicable.

Author Contributions: E.R. drafted the first version and following revisions of the manuscript under ST's supervision. LC and AG contributed to the selection of variables. All other authors contributed additional content, edits, and references. All authors approved the final version.

Funding: This research was funded by European Union's Horizon 2020 research and innovation program under Grant agreement no. 101016167, 6 November 2020

Data Availability Statement: The list of variables and relative mapping can be found at https://cloud.orchestra-cohort.eu/s/j2Fe4AxpHLysdJt

Acknowledgments: We would like to thank the entire ORCHESTRA Project Group for their support.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Bodenreider O, Cornet R, Vreeman DJ. Recent Developments in Clinical Terminologies - SNOMED CT, LOINC, and RxNorm. Yearb Med Inform. 2018;27(1):129-139. doi: 10.1055/s-0038-1667077.
2. International S. GlobeNewswire News Room. 2020. SNOMED International Joins COVID-19 Interoperability Alliance to Support Access to Terminology and Health Data Resources. Available from: https://www.globenewswire.com/en/news-release/2020/05/22/2037552/0/en/SNOMED-International-Joins-COVID-19-Interoperability-Alliance-to-Support-Access-to-Terminology-and-Health-Data-Resources.html
3. Mahase E. Covid-19: What do we know about "long covid"? BMJ. 2020;370:m2815. doi: 10.1136/bmj.m2815.

4. Benson T, Grieve G. FHIR Resources. In: Benson T, Grieve G, editors. Principles of Health Interoperability: SNOMED CT, HL7 and FHIR [Internet]. Cham: Springer International Publishing; 2016. p. 361–79. (Health Information Technology Standards). Available from: https://doi.org/10.1007/978-3-319-30370-3_20
5. Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The Usage of OHDSI OMOP - A Scoping Review. Stud Health Technol Inform. 2021;283:95-103. doi: 10.3233/SHTI210546.
6. Rinaldi E, Stellmach C, Rajkumar NMR, Caroccia N, Dellacasa C, Giannella M, et al. Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic. NPJ Digit Med. 2022;5(1):75. doi: 10.1038/s41746-022-00620-x.
7. Marcus JS, Martens B, Carugati C, Bucher A, Godlovitch I. The European Health Data Space (December 12, 2022). IPOL | Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament Policy Department studies [Internet]. 2022 [cited 2024 Feb 16]. Available from: https://www.ssrn.com/abstract=4300393