Converting HL7 CDA Based Nationwide Austrian Medication Data to OMOP CDM

Florian KATSCH[a,b,1], Rada HUSSEIN[a],
Raffael KORNTHEUER[b] and Georg DUFTSCHMID[b]
[a] *Ludwig Boltzmann Institute for Digital Health and Prevention, Salzburg, Austria*
[b] *Center for Medical Data Science, Medical University of Vienna, Vienna, Austria*

Abstract. Austria's national Electronic Health Record (EHR) system holds information on medication prescriptions and dispenses in highly structured HL7 Clinical Document Architecture (CDA) documents. Making these data accessible for research is desirable due to their volume and completeness. This work describes our approach of transforming the HL7 CDA data into Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) and highlights a key challenge, namely mapping the Austrian drug terminology to OMOP standard concepts.

Keywords. Electronic Health Record, Electronic Prescription, HL7 CDA, OMOP CDM, Health Information Interoperability

1. Introduction

Since 2019 Austria has been operating a national health information exchange (HIE) system for medication data. All prescriptions from outpatient care providers and all dispenses in pharmacies are included in structured form, based on the HL7 Clinical Document Architecture (CDA) standard. Currently, about 97% of the 9 million Austrian inhabitants participate in the HIE system. Roughly 10 million medication prescriptions and dispenses are registered to the system monthly [1].

The OMOP CDM is a well-established standardized representation of clinical data that is broadly used in inter-institutional research projects. In this work we describe our project for transforming Austrian's CDA-formatted medication data to the OMOP CDM, supported by the European Health Data and Evidence Network (EHDEN).

2. Methods

In the HL7 CDA data model, prescriptions and dispenses are represented via classes *SubstanceAdministration* and *Supply*. Both reference class *ManufacturedProduct*, which holds the drug to be consumed. The CDA data model is customized to the settings of concrete HIE use cases via CDA templates. The Austrian CDA medication templates [2] are conformant to the corresponding IHE Patient Care Coordination templates [3] to achieve international interoperability. In contrast to [4], who used absolute XPaths to

---

[1] Corresponding Author: Florian Katsch, E-mail: florian.katsch@dhp.lbg.ac.at.

refer to CDA content, we apply relative XPaths based on template identifiers within our extract-transform-load (ETL) specification. Besides several other CDM tables, *DRUG_EXPOSURE* represents the main target of the CDA medication data. As CDA medication classes are highly structured, 17 from 23 *DRUG_EXPOSURE* attributes can be fed. Attributes *drug_concept_id* and *route_concept_id* are supplied from data coded in Austrian-specific terminologies and thus need to be mapped to CDM standard concepts. In our ETL activities we are supported by an EHDEN certified SME.

3. Results

We have completed the EHDEN ETL definition document that specifies the mappings. Currently, we are working on the ETL implementation by means of a Python project. After setting up prerequisite components (such as the database connectors and test infrastructure) we continued with supplying the commonly used CDM tables (e.g. *PERSON*, *CARE_SITE*). Meanwhile we have started integrating the clinical fact tables.

4. Discussion and Conclusion

A key challenge of our work is to map Austrian drug codes (PZN) to the CDM standard code system RxNorm. Each PZN is associated with at least one ATC code and OMOP features a mapping from ATC to RxNorm. However, ATC codes only allow mapping to the rather coarse RxNorm *Ingredient* level. Thus we lose information about drug strength and dose form. Furthermore, due to different possible indications for a drug administration, a PZN may have multiple ATC codes assigned, which introduces additional ambiguities. A solution could be to make use of the "Austria-Codex" drug ontology that includes, amongst others, active substances, drug strength, and dose form. As these attributes are also available in OMOP for the more granular levels of RxNorm, an ontology alignment seems possible. We plan to explore this alignment with clinical domain experts by focusing on a limited set of diabetes-specific drugs that are particularly relevant in a current research project. Our work will pave the way for Austrian medication data to be usable for research on a national scale. Since our CDA medication templates are conformant to the IHE Patient Care Coordination medication templates, our work has potential for international reuse.

References

[1]   ELGA GmbH. ELGA: Zahlen-Daten-Fakten [Internet]. 2022 [cited 2023 Mar 2]. Available from: https://www.elga.gv.at/elga/zahlen-daten-fakten/

[2]   ELGA GmbH. Implementierungsleitfaden E-Medikation (Version_2.06) [Internet]. 2015 [cited 2023 Mar 2]. Available from: https://www.gesundheit.gv.at/dam/jcr:5abf0398-6785-433b-a3d4-d0e0972c6b03/Implementierungsleitfaden_E-Medikation_(Version_2.06).pdf

[3]   IHE. IHE Patient Care Coordination (PCC) 5 Technical Framework [Internet]. 2021 [cited 2023 Mar 2]. Available from: https://www.ihe.net/uploadedFiles/Documents/PCC/IHE_PCC_TF_Vol2.pdf

[4]   Ji H, Kim S, Yi S, Hwang H, Kim J-W, Yoo S. Converting clinical document architecture documents to the common data model for incorporating health information exchange data in observational health studies: CDA to CDM. J Biomed Inform. 2020 Jul;107:103459.