

Leveraging FHIR for a generic EHR-to-OMOP ETL Can we make an ETL process reusable?

PRESENTER: **Sebastiaan van Sandijk**

INTRO

- Harmonization of EHR data can be cumbersome. The speed, effort, and quality of a conversion of EHR data to OMOP heavily depend on factors that are not always transparent in the resulting CDM.
- One way to improve data harmonization and to reduce the effort involved in OMOPing EHR data, is to (re)use a generic ETL.
- There are currently **no detailed guidelines** for a generic process.
- We explore how EHR data sets extracted according to standard, **FHIR-based specifications** can provide a **standardized basis for a generic OMOP-ETL**.

METHODS

1. Analysis of the specifications of the Dutch BGZ 'Basisgegevensset Zorg', European Patient Summary (EPS) and International Patient Summary (IPS).
2. Preliminary comparison of test data created according to these specifications. A comparison of extractions from three different EHR systems is underway.
3. Discussions with European standardization experts and projects about evolving developments (IPS in particular)

RESULTS

- Data elements included in **FHIR-based** data sets are **more clearly specified** and could provide a useful basis for a generic ETL.
- The BGZ is a structured data set extracted in separate queries. The IPS is a **document with structured data elements**. The latter will improve harmonization of extracted data from different EHR systems.



FHIR-based standards like the International Patient Summary (IPS) can be a useful basis for a generic OMOP-ETL. Metadata extensions are needed to accommodate research requirements.

- However, also the IPS is a 'snapshot' of what is in the EHR system, and the specifications do not require metadata about context, scope and completeness of the extracted data. This would result in a lack of transparency about the (observation)period covered and completeness of the data in the summary.
- Further exploration of data sets that are extracted from various EHR systems (conformant to BGZ and/or IPS specifications) is essential.
- So far, we anticipate different vendors to implement the standards in more or less slightly different manners.
- Because these are FHIR-based specifications, however, we expect that the extracted data sets still can be processed in generic data profiling, mapping, and transformation processes. This would include semantic mappings, making use of the specified (inter)national coding systems.

CONCLUSION

- Standardized FHIR-based data sets may well prove to form a useful basis for a reusable ETL process.
- The specifications currently show important gaps, especially regarding metadata that are important for research and researchers.
- Detailed comparison of practical extractions from different EHR systems is still needed.
- We work towards recommendations for (metadata) extensions in the IPS specification.

Sebastiaan van Sandijk^{1,3}
Renske Los^{2,3}

¹ Odysseus Data Services, ² Erasmus Medical Center, ³ OHDSI NL

*Sebastiaan.van.Sandijk@odysseusinc.com



ODYSSEUS
DATA SERVICES INC

