

Making PROMs Work in OHDSI: A Framework for Interoperable Collection and Mapping of PROMs Using the EQ-5D-5L as a Worked Example

Mario Lawes^{1,2}, Alexia Sampri^{1,2}, Afreen Sikandara³, David Weese¹, Rok Kos¹, Ben Illigens^{1,2,4}

¹ Data4Life, Berlin, Germany ² Hasso Plattner Institute, Potsdam, Germany ³ Data4Life Asia, Singapore

⁴ Dresden International University, Germany

Background

Patient-reported outcome measures (PROMs) capture health status, treatment response, and quality of life directly from patients, making them a valuable complement to routinely collected clinical data. While some PROMs are embedded directly in electronic health record systems, many are collected through dedicated research or patient engagement platforms and must subsequently be integrated with existing clinical data for secondary research. Integrating such externally collected PROMs with OMOP data in an interoperable way remains a significant challenge. Key barriers include fragmented collection systems, inconsistent exchange formats, and the lack of a systematic mapping approach. While HL7 Fast Healthcare Interoperability Resources (FHIR; 1) has emerged as the dominant open standard for exchanging questionnaire data, its intentional flexibility means that it does not prescribe how questionnaires and their responses should be encoded. Bridging FHIR and OMOP for PROMs therefore requires deliberate technical decisions that are rarely documented systematically. Using the EQ-5D-5L (2), a widely used standardized PROM for measuring health-related quality of life, as a worked example, we developed and implemented an end-to-end pipeline spanning digital administration, interoperable exchange, and OMOP transformation.

Methods

In our implementation we built upon the Structured Data Capture Implementation Guide (3), the HL7 Vulcan FHIR-to-OMOP implementation guide (4), and ongoing work of the OHDSI CDM Survey Subgroup on representing survey data in the OMOP CDM. Participants receive the EQ-5D-5L from a FHIR Questionnaire resource and their answers are returned as a FHIR QuestionnaireResponse.

We specified the transformation logic in a FHIR StructureMap for OMOP CDM v5.4 and executed it within a Data2Evidence ETL flow. To encode the questions, we used existing SNOMED CT codes, which resolve directly to standard OMOP concepts. Response values are encoded numerically, alongside with standard concept IDs where these exist. For table assignment, we followed the conventions of the OHDSI CDM Survey Subgroup and mapped responses that serve as the basis for a computed score to OBSERVATION and values that are directly interpretable as measures to MEASUREMENT. Each EQ-5D-5L response yields seven pieces of information, grouped into three types of data:

Item-level responses. Patients are asked to rate their current health in five items with respect to mobility, self-care, usual activities, pain/discomfort, and anxiety/depression on a five-level severity scale. The responses to these five items serve as the basis for computing the index value and are mapped to OBSERVATION, one row per dimension and response, with the response level stored in `value_as_number` and the corresponding OMOP concept ID in `value_as_concept_id`.

EQ VAS. The second type is a patient rating of overall health on a visual analogue scale (range 0 to 100), which is directly interpretable as a measure of overall health. Each rating is mapped to MEASUREMENT as `value_as_number`, one row per response.

EQ-5D index. The third type is a weighted score derived from the five item-level responses using country-specific preference weights that reflect the health-state preferences of a general population. This index value is used in health economics, most prominently for calculating quality-adjusted life years (QALYs) and is written to MEASUREMENT as value_as_number, one row per response. The index is derived by a separate Data2Evidence job that reads the OBSERVATION records and applies a configurable set of preference weights, running outside the ETL flow. This keeps the item responses unchanged as raw input and allows index values to be recalculated without re-running the full ETL. Because preference weights are country-specific, we propose that each CDM instance apply a single set of weights and record the applied weights and the derivation method in the METADATA table, so that index values within a database remain comparable.

The StructureMap and ETL flow are openly available on the GitHub page of Data2Evidence: <https://github.com/data2evidence/templates>.

Results

The pipeline builds on D4L Collect (5), a platform for PROM and sensor data collection designed for health research, and Data2Evidence (6,7), an open-source platform for OMOP-based research that can ingest FHIR resources, store them natively in an integrated FHIR server and orchestrate an ETL pipeline that transforms FHIR data into OMOP, enabling use of the standard OHDSI tools for observational research (Figure 1).

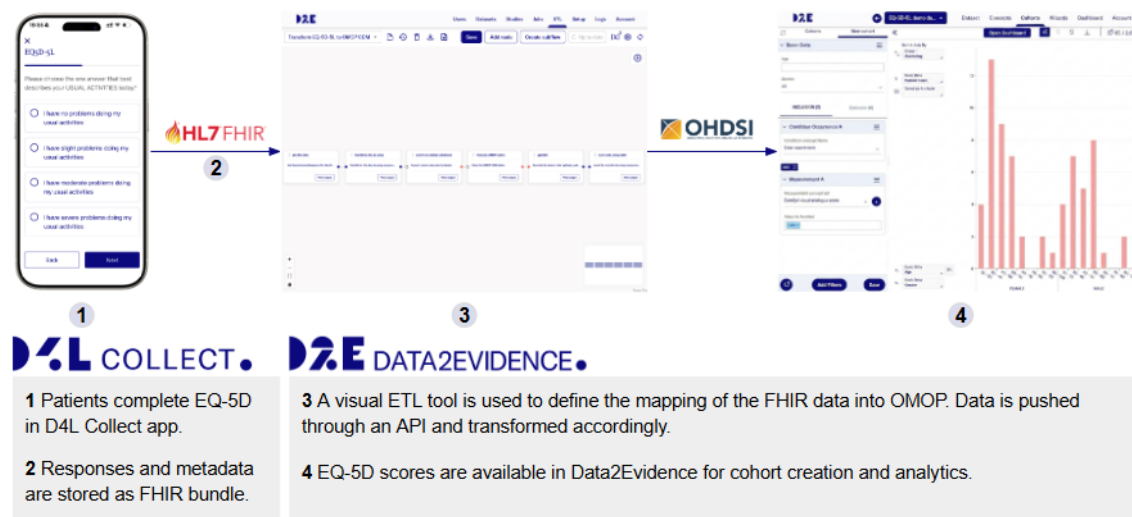


Figure 1. End-to-end pipeline from digital collection of the EQ-5D-5L to analysis-ready OMOP data using the D4L Collect and Data2Evidence platforms.

We exercised the pipeline on synthetic EQ-5D-5L responses generated in D4L Collect (25 QuestionnaireResponses from 25 simulated participants). No real patient data were used. Following the approach to FHIR-to-OMOP lineage tracking described by Berk et al. (8), each transformation writes a lineage record linking every source QuestionnaireResponse to the OMOP rows it produced. We use these records to verify the mapping against an expected pattern: five OBSERVATION rows per response for the item-level answers, one MEASUREMENT row for the EQ VAS rating, and one further MEASUREMENT row per response once the index job is run. No row should exist without a traceable source. Results for the finalized mapping are reported on the accompanying poster. Validation of the index derivation job is ongoing.

Conclusion

PROM integration into OMOP involves a series of implicit decisions, from FHIR questionnaire design and response encoding to OMOP table mappings and score provenance. The EQ-5D-5L provides a rich and instructive example for surfacing these decisions. The work presented here is a working implementation, with an openly available StructureMap and ETL flow that other researchers can use directly. The index derivation job will be published alongside them.

As a next step we are seeking collaborations with applied researchers to demonstrate the analytical value of the resulting data, such as drug effectiveness and safety studies with the EQ-5D-5L as a patient-centered endpoint, health economic analyses using index scores for QALY estimation, and rehabilitation research linking longitudinal EQ-5D-5L trajectories to routinely collected data.

In parallel, we are writing a practical blueprint for the broader OHDSI community interested in digitally collecting PROMs using FHIR and integrating them with existing OMOP data, setting out the decisions at each step and how to document them. For this we will again use the EQ-5D-5L as a worked example and discuss how these decisions apply to other PROMs. We believe that documenting and sharing these decision pathways systematically is essential for making PROMs data truly interoperable and unlocking its full potential in large-scale observational research.

References

1. HL7 International. HL7 FHIR Release 4 (R4) [Internet]. 2019. Available from: <https://hl7.org/fhir/R4/>
2. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011 Dec;20(10):1727–36. doi:10.1007/s11136-011-9903-x
3. Health Level Seven International (HL7). Structured Data Capture [Internet]. Structured Data Capture. Available from: <http://hl7.org/fhir/uv/sdc/STU4>
4. Health Level Seven International (HL7). FHIR to OMOP FHIR IG [Internet]. [cited 2026 Jul 31]. FHIR to OMOP CDM Implementation Guide. Available from: <https://hl7.org/fhir/uv/omop/INFORMATIVE1/en/>
5. Weese D, Alsaeedi S, Helmchen F, Hillebrand L, Holst T, Huenges J, et al. D4L Collect: a multimodal data collection and study management platform for digital health research. in prep.
6. Data4Life. Data2Evidence [Internet]. GitHub; 2026. Available from: <https://github.com/OHDSI/Data2Evidence>
7. Guerrero P, Ernebjerg M, Holst T, Weese D, DiBello H, Ibing S, et al. The AIR-MS data platform for artificial intelligence in healthcare. *JAMIA Open.* 2025 Nov 3;8(6):ooaf145. doi:10.1093/jamiaopen/ooaf145
8. Berk B, Benzie M, Bolisetty B, Favre S, Fortune J, Goslin J, et al. Bridging FHIR and OMOP: Data Lineage for Observational Data Conversion – OHDSI. In. 2025. Available from: <https://www.ohdsi.org/2025showcase-110/>