FHIR-to-OMOP Cookbook

Livne, G., Terry, M., Yang, Q.

[1 Introduction 2](#_Toc180574115)

[1.1 Purpose of the Guidebook 2](#_Toc180574116)

[1.2 Target Audience 3](#_Toc180574117)

[1.3 Scope and Assumptions 3](#_Toc180574118)

[1.4 Background 3](#_Toc180574119)

[1.4.1 Conventions 4](#_Toc180574120)

[2 Mapping Principles and Methodology 4](#_Toc180574121)

[2.1 Holistic Approach to Mapping 4](#_Toc180574122)

[3 Lessons learned from actual FHIR to OMOP transformations 6](#_Toc180574123)

[4 Practical Examples and Case Studies 6](#_Toc180574124)

[4.1 General Patterns 6](#_Toc180574125)

[4.1.1 What not to map 6](#_Toc180574126)

[4.1.2 The FHIR concept map is equivalent to the source-to-concept-map table in OMOP CDM. 7](#_Toc180574127)

[4.1.3 Aligning FHIR and OMOP statuses 9](#_Toc180574128)

[4.1.4 Aligning Pre- and Post-coordinated concepts 10](#_Toc180574129)

[4.1.5 Capturing and maintaining new concepts which do not exist today in any standard 11](#_Toc180574130)

[4.1.6 Certain constructs do not exist in OMOP but are required by FHIR. 11](#_Toc180574131)

[4.1.7 Mapping FHIR extensions 11](#_Toc180574132)

[4.2 FHIR Patient, Person, and Practitioner Patterns 11](#_Toc180574133)

[4.2.1 Pattern: Handling person names 12](#_Toc180574134)

[4.2.2 FHIR.human-name exists 12](#_Toc180574135)

[4.3 FHIR Condition Patterns <WIP to be continued> 12](#_Toc180574136)

[4.3.1 Condition.category 12](#_Toc180574137)

[4.3.2 Condition.clinicalStatus 13](#_Toc180574138)

[4.3.3 Patient-stated Conditions 13](#_Toc180574139)

[4.4 FHIR Observation Patterns <WIP to be continued> 14](#_Toc180574140)

[4.4.1 Map by Category 14](#_Toc180574141)

[4.4.2 Grouping of FHIR Observation components into OMOP 15](#_Toc180574142)

[4.5 FHIR Medication Patterns <WIP to be continued> 16](#_Toc180574143)

[4.6 FHIR ServiceRequest Patterns <WIP to be continued> 16](#_Toc180574144)

[4.7 FHIR Procedure Patterns <WIP to be continued> 16](#_Toc180574145)

[5 Advanced Mapping Challenges 16](#_Toc180574146)

[5.1 Handling complex mapping scenarios 16](#_Toc180574147)

[5.1.1 Bidirectional Transformations 16](#_Toc180574148)

[5.1.2 Rare Data Elements 17](#_Toc180574149)

[5.2 Managing discrepancies <WIP to be continued> 17](#_Toc180574150)

[5.3 Maintaining data fidelity <WIP to be continued> 17](#_Toc180574151)

[6 Ongoing Updates and Community Contributions 17](#_Toc180574152)

[7 Contact Information 17](#_Toc180574153)

[8 Appendix A: Abbreviations 17](#_Toc180574154)

[9 Appendix B: Mapping recommendations for specific scenarios 18](#_Toc180574155)

[10 Appendix C: Comparing FHIR and OMOP Frameworks 19](#_Toc180574156)

[10.1 Modeling in FHIR 19](#_Toc180574157)

[10.2 Framework Design Assumptions 19](#_Toc180574158)

[10.3 General Design Principles - OMOP CDM 20](#_Toc180574159)

[10.4 General Design Principles - FHIR Framework 21](#_Toc180574160)

[11 Appendix D: Resources and References for further reading 22](#_Toc180574161)

[11.1 FHIR References 22](#_Toc180574162)

[11.2 OHDSI References 22](#_Toc180574163)

# Introduction

## Purpose of the Guidebook

This is a starter guide for implementers seeking to convert Health Level 7 (HL7®) Fast Healthcare Interoperability Resources (FHIR®) resources to the Observational Health Data Sciences and Informatics (OHDSI) Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). The recommendations cover not just mapping elements, but a methodology on how to approach the mappings.

This is a ***living document*** which will be updated based on new findings and best practices from the community. We welcome the community to join us in contributing to this work. Reference the *Contact Information* section of this document for further details.

It was written in the style of [*The Book of OHDSI*](https://ohdsi.github.io/TheBookOfOhdsi/)but with the ambitious goal of addressing the needs of both the HL7 FHIR and OHDSI communities who seek to integrate both data models and tools. References are made to relevant sections in the Book of OHDSI where appropriate.

## Target Audience

The target audience is comprised of:

* OHDSI experts who want to better understand how to integrate with FHIR.
* FHIR experts who want to better understand how to integrate with OHDSI.

This document is intended for but not limited to the following:

* Informaticians
* Extract, Transform, Load (ETL) / data engineers
* Software developers
* Data scientists

We highly recommend that cookbook readers be at least familiar with the foundational principles of both FHIR and OMOP. Reference Appendix D for a list of references as well as a summary of foundational design differences among both frameworks.

## Scope and Assumptions

We recognize that FHIR profiles may extend elements for a given purpose. Some references to the use of extensions may be covered in this document, however our emphasis is on the foundational FHIR elements.

The cookbook is agnostic of a specific clinical domain or country, although some country-specific references are made for illustrative purposes.

## Background

The HL7/OHDSI partnership was announced in March 2021 with the goal that organizations will “align their standards to capture data in a clearly defined way into a single common data model.”[[1]](#footnote-1) Specifically, the partnership focused on the alignment between two information models - HL7 FHIR resources and the OMOP CDM. Four sub-groups were identified. Two of the four were general - FHIR-OMOP Core Model and FHIR-OMOP Vocabulary. The other two were domain-specific - Data Quality Measures, and Oncology. As an initial use case, the Oncology sub-group chose to map a specific FHIR implementation guide (IG) from the minimum Common Oncology Data Elements (mCODE) to the OMOP CDM, and loosely referencing the OMOP Oncology Extension where possible. Several FHIR profiles have been mapped and prototyped at the May 2022 FHIR Connectathon. Learnings that apply to the HL7/OHDSI initiative overall were captured in the Connectathon Track report-out and shared with the other groups.

In the process of mapping, analysis, and prototyping, the Oncology sub-group identified the need to create general axioms, design principles, and overarching recommendations which previously did not exist or have not yet been formalized. As such, the sub-group initiated guidance that may apply for any implementation looking to translate from FHIR to OMOP.

Other FHIR-to-OMOP standards development and consensus groups have formed since the initial release of this cookbook in November 2022. Specifically, the HL7 Vulcan FHIR Accelerator [FHIR-to-OMOP project](https://confluence.hl7.org/display/VA/FHIR+to+OMOP) sub-group gathered several implementers who have mapped from FHIR to OMOP. Each contributor shared their approaches and mappings with the aim of harmonizing all these approaches to map from OMOP CDM v5.4 to the [FHIR International Patient Summary (IPS) FHIR IG](https://build.fhir.org/ig/HL7/fhir-ips/) with the goal of creating a more formalized [FHIR-to-OMOP Implementation Guide](https://build.fhir.org/ig/HL7/fhir-omop-ig/).

The FHIR-to-OMOP Cookbook is intended to be complementary to the Vulcan effort. The cookbook adopts a more pragmatic approach to mapping among both standards. It proposes a methodology to mapping FHIR representation patterns found in a FHIR IG, and options for mapping to OMOP based on the FHIR resource representation patterns. It assumes that FHIR is a *framework* of resources, APIs, and tools that allow for the creation of implementation guides with complex extensions and computable constraints (e.g.: invariants).

### Conventions

* The FHIR notion “value set” is equivalent to the OMOP “concept set”. Both terms may be used interchangeably.
* The FHIR notion “code system” is equivalent to the OMOP “vocabulary”. Both terms may be used interchangeably.

# Mapping Principles and Methodology

Mapping principles and strategies were determined based on framework design principles and assumptions further described in Appendix C.

## Holistic Approach to Mapping

We propose a holistic methodology in approaching a FHIR-to-OMOP mapping.

1. Not all FHIR metadata is relevant for OHDSI research studies or mapping the OMOP CDM
2. Define the relevant data in the FHIR source to map to OMOP
   1. Map only complete and verified information
3. Identify the *main* FHIR element that lead the mapping to OMOP
4. Identify the FHIR elementDefinitions that align with the main element.
5. Analyze the corpus between the content inventory FHIR codeableConcept with the OMOP Ontology for gaps and misaligned domains.
6. Identify or generate FHIR profiles which align with the intended study or data-lake to ensure translation among multiple data sources to the OMOP CDM.
7. Map the relevant FHIR resources / profile elements to OMOP CDM.
8. Populate the OMOP CDM records at the atomic/record level.
9. Preserve the FHIR relationships (or provenance) from the original resource where possible (meas\_event\_id, observation\_event\_id, fact\_relationship, etc.)
10. For OHDSI Network Studies, test data quality and consistency in FHIR-OMOP data translation using a common test data set relevant to the network study.

Practical steps on the methodology for mapping common FHIR resources to OMOP structures is further illustrated in the flowchart diagram below:

Diagram

Description automatically generated

The flowchart consist of 4 main categories of tasks:

1. Identifying and scoping the most relevant FHIR resource elements to map to OMOP.
2. Mapping the defining FHIR resource element to the OMOP ontology to determine the OMOP domain for which to map the FHIR resource’s supporting elements.
3. Aligning the supporting fields with the OMOP CDM table from the identified domain.
4. Handling any mapping gaps by falling back to a rigid pre-defined structural map between the FHIR resource type and an OMOP table, documenting the gap, and coordinating with the OHDSI THEMIS group on FHIR alignment needs and recommendations.

# Lessons learned from actual FHIR to OMOP transformations

Many initiatives assume that the source data can be aggregated into a data lake and used generically regardless of context. This may be true for general analyses involving “clinical and administrative assertions” which describe only one notion. For example, an atomic condition on a problem list captured during an encounter may include hypertension. There is no other qualifier or modifier to further describe the condition such as primary hypertension or secondary hypertension.

Defining or at least loosely identifying a set of study cohorts has several advantages. It helps to scope and prioritize the FHIR-to-OMOP mapping as well as provide a use case to test the mapping and ETL logic.

FHIR base resources have very few constraints and consequently allows for transmission of sparse and semantically lossy data. As a result, it is challenging for FHIR-to-OMOP ETL developers to ensure in a computable way whether a FHIR server already populated with resources contains the necessary data elements.

# Practical Examples and Case Studies

## Out of Scope

We recommend the following elements not be mapped for PHI reasons:

* Observation.category
* From Patient resource type:
  + Patient and patient contact names
  + Patient residential addresses other than postal/zip code.
  + Telecom (e.g.: phone numbers, email)
  + Medical record numbers
  + Government-issued identifiers (e.g.: social security number, driver’s license number, etc.)

Such lower priority FHIR elements are considered FHIR-specific metadata, or protected health information (PHI) which are likely to be excluded by Institutional Review Board (IRB) restrictions on human subject data use for observational research.

## General Patterns

This section summarizes structural and semantic mapping patterns that apply to multiple FHIR resources and is not specific to a clinical specialty or domain.

### What not to map

Some FHIR elements might be excluded or lower priority when mapping to the OMOP CDM. These include but are not limited to:

* The meta element in all FHIR resources.
* The text element in all FHIR resources.

For example, the meta and text values in the FHIR resource below would not be mapped to OMOP CDM:

"[resourceType](http://hl7.org/fhir/R4/condition.html#Condition)" : "Condition",

"[id](http://hl7.org/fhir/R4/condition.html#Condition.id)" : "us-core-condition-depression-jenny-m",

"[meta](http://hl7.org/fhir/R4/condition.html#Condition.meta)" : {

"[profile](http://hl7.org/fhir/R4/datatypes.html#Meta#Meta.profile)" : [

[🔗](http://hl7.org/fhir/us/core/STU6.1/StructureDefinition-us-core-condition-problems-health-concerns.html) "http://hl7.org/fhir/us/core/StructureDefinition/us-core-condition-problems-health-concerns"

]

},

"[text](http://hl7.org/fhir/R4/condition.html#Condition.text)" : {

"[status](http://hl7.org/fhir/R4/datatypes.html#Narrative#Narrative.status)" : "generated",

"[div](http://hl7.org/fhir/R4/datatypes.html#Narrative#Narrative.div)" : "<div xmlns=\"http://www.w3.org/1999/xhtml\"><p class=\"res-header-id\"><b>Generated Narrative: Condition us-core-condition-depression-jenny-m</b></p><a name=\"us-core-condition-depression-jenny-m\"> …etc… </a></p></div>"

}

### Title: The FHIR concept map is equivalent to the source-to-concept-map table in OMOP CDM.

**Pattern Description:** Code system does not exist as a standard OMOP vocabulary

The implications are in the ETL and whether the implementer will first query a FHIR terminology server to parse or reference OMOP CDM staging tables for source to concept mapping.

**Examples:**

1. Code systems OMB Race and Ethnicity category codes do not have OMOP concept ids.
2. Many FHIR-specific code systems for which HL7 is the steward. A list of FHIR internal and external code systems can be found in the [*HL7 Terminology Publication*](https://terminology.hl7.org/index.html).

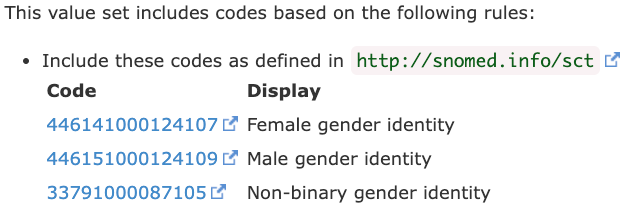
**Recommendation Options:**

|  |  |  |
| --- | --- | --- |
| **Option** | **Pros** | **Cons** |
| Request to add the code system as an OMOP standard vocabulary. | No FHIR concept map needed. | Slow process. |
| OMOP ETL performs cross-walk mapping;  Populates source\_to\_concept\_map | Less mapping burden on FHIR implementation - preserves existing concepts | More mapping burden on the OMOP ETL;  Process: Submit the concept map to populate source\_to\_concept\_map table; |
| Create a FHIR concept map and implementation sends the OMOP concept\_id | Less mapping burden on OMOP ETL | More mapping burden on the FHIR implementation |
| Represent as an uncoded term with no OMOP concept\_id. | No lost data, even if it is unstructured. | Data lossiness; potential errors of omission if cohort definition does not handle non-standard terms. |
| Request addition to the HL7 Healthcare Terminology Authority (HTA). | Some alignment through vocabulary | Slow process. |

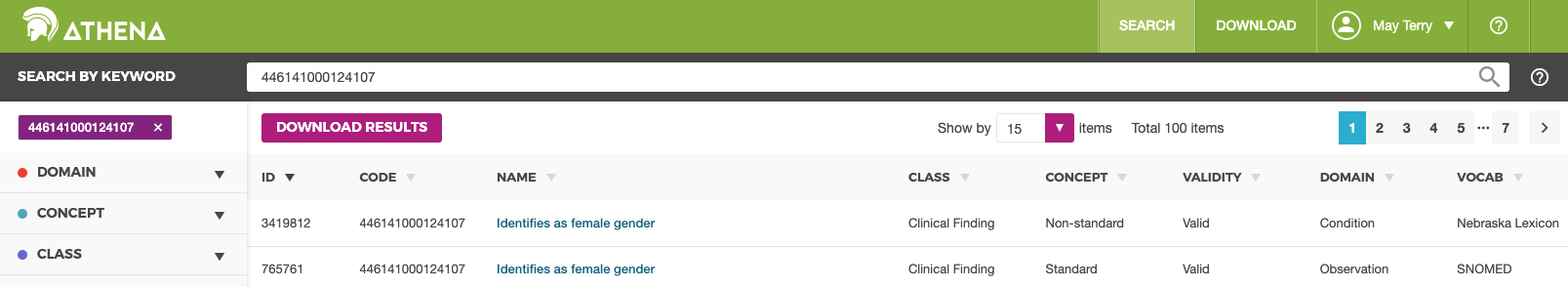
Comment: <*refer to THEMIS or the Book of OHDSI section on handling non-standard concepts*>. As a best practice, identify if the non-standard concept has a mapping to an OMOP standard concept. This can be searched in Athena.

**Example:**

Scenario: you are mapping from FHIR to OMOP for gender identity, which has the following value set:

* <https://terminology.hl7.org/5.5.0/ValueSet-gender-identity.html>
* 

Term: Identifies as female gender with the [SNOMED code 446141000124107](https://browser.ihtsdotools.org/?perspective=full&conceptId1=446141000124107&edition=MAIN/2024-07-01&release=&languages=en&latestRedirect=false) has two concepts



In the context of SNOMED-CT, this is a standard concept linked to a domain of Observation. However, note that if you were using Nebraska Lexicon as your source, then the concept is non-standard with a domain of Condition. You would need to use the OHDSI OMOP ontology-related tables to 1) check if there is a non-standard to standard OMOP concept.



### Aligning FHIR and OMOP statuses

**Pattern Description:** The notion of a *status* in FHIR is not the same as a *status* in the OMOP CDM.

**Examples:**

1. FHIR Condition.clinicalStatus is bound to the HL7 [condition-clinical](http://terminology.hl7.org/CodeSystem/condition-clinical) code system which contains codes that have no concept equivalent in OMOP.
2. FHIR MedicationRequest.status is bound to the HL7 [medicationrequest-status](https://build.fhir.org/valueset-medicationrequest-status.html) code system and contains codes that have no concept equivalent in OMOP.
3. We talked about loading only status completed to omop

**Recommendation Options:**

|  |  |  |
| --- | --- | --- |
| **Option** | **Pros** | **Cons** |
| OMOP CONDITION\_OCCURRENCE table adds a new field for condition status. | No FHIR concept map needed. |  |

### Aligning Pre- and Post-coordinated concepts

**Pattern Description:**

Semantic alignment between FHIR and OMOP vocabularies is not always as simple as asking for new concepts and vocabularies. **Term binding** to elements could involve complexities of pre- and post-coordination.

There is no ideal best practice for the decision, but there are trade-offs. Use of highly pre-coordinated concepts simplify the structural data model, however with a downside of a combinatorial explosion of pre-coordinated concepts.

Several OMOP-related projects tend to rely on pre-coordinated standard OMOP concepts. FHIR has many individual elements with more atomic concepts binding to the element.

|  |
| --- |
| ***\*\*Consensus for pre- and post-coordinated concepts is critical for OMOP because OHDSI’s reliance on the OMOP Ontology to determine the appropriate OMOP CDM table to map data.\*\**** |

**Recommendation Options:**

The table below lists variations in FHIR-to-OMOP concept mappings:

|  |  |  |
| --- | --- | --- |
| **FHIR-to-OMOP Concept mapping pattern** | **Example** | **Additional Remarks** |
| Non-coded to non-coded | FHIR: Patient.birthDate  OMOP: PERSON.year\_of\_birth | No FHIR concept map needed. |
| Non-coded to coded |  |  |
| Coded to coded |  |  |
| Post-coordinated pattern 1: non-coded/coded to coded |  |  |
| Post-coordinated pattern 2: coded/coded to coded |  |  |
| Pre-coordinated to pre-coordinated |  |  |

### Capturing and maintaining new concepts which do not exist today in any standard

**Pattern Description:** TBD

**Examples:** TBD

**Recommendation Options:**

Codeable concepts from code systems which do not exist in OMOP will be mapped as free-text.

1. Race and ethnicity concepts that are way more detailed than what exists today.

### Certain constructs do not exist in OMOP but are required by FHIR.

**Pattern Description:** TBD

**Examples:**

1. OMOP does not have a notion of Condition.clinicalStatus and FHIR has a 1..1 cardinality for the Condition resource. It is repetitive of previous page

**Recommendation Options:**

|  |  |  |
| --- | --- | --- |
| **Option** | **Pros** | **Cons** |
| OMOP ETL maps every status to *completed*. Assume that in OMOP handling of observational studies, all CONDITION\_OCCURRENCE and DRUG\_EXPOSURE activities have been completed. Do not map information that the statuses is not Completed. | No FHIR concept map needed. | Potentially changes intent. |
| Do not map status and fix the FHIR extension data-absent-reason = *unsupported*. | Reduces ambiguity with an explicit assertion that an equivalent field did not exist in OMOP. | None. |

### Mapping FHIR extensions

FHIR contains extensions which further qualify or modify a specific base resource. They can be standard (as determined by HL7 FHIR) or custom (as specified in a FHIR profile within a given IG).

Use the same principal and methodology to map extensions.

## FHIR [Patient](http://hl7.org/fhir/patient.html), [Person](http://hl7.org/fhir/person.html), and [Practitioner](http://hl7.org/fhir/practitioner.html) Patterns

Description: The FHIR Patient, Person, and Practitioner resources have common constructs which identify a specific human individual, including the following:

* .name
* .telecom
* .gender

### Pattern: Handling person names

The OMOP CDM PERSON table has no fields to indicate a person’s name. Instead, a de-identified person\_id is provided on-purpose to address PHI constraints. Consequently, it is implied that there is a separate and secure cross-tabulation between the de-identified person ID and their PHI (name, address, contact information, etc.).

### FHIR.human-name exists

If a human name exists, then map the following FHIR elements as follows:

One convention in FHIR-to-OMOP is to create an OBSERVATION record for which the concept\_id is the SNOMED code for “*person name”*. In this case, the link between that OBSERVATION record back to PERSON is the foreign key for person\_id.

FHIR however does all multiple person names with a qualifier for different types of names (legal, birth, etc). Further investigation is needed to identify where such qualifiers are stored. FHIR has a separate element for this qualifier and it is likely to be a pre-coordinated concept in OMOP instead. For example: OMOP could have an OBSERVATION with the concept\_id for maiden name ([OMOP id: 764223](https://athena.ohdsi.org/search-terms/terms/764223)).

### Pattern: Handling Patient.provider references

The FHIR Patient resource has 0..\* whereas the OMOP CDM PERSON table 1) has only one field for a provider\_id, 2) further restricts the provider to be “*the last known general practitioner of the person. If there are multiple providers, it is up to the ETL to decide which to put here*.”

It is not uncommon for a patient to have multiple providers in their care, or to be part of a physician group where several providers may be the PCP for a patient depending on availability.

We recommend that unless there is only one primary care physician assigned to the patient, that this field not be populated. We find that most use cases have no need for this field and that specific provider information is more relevant to the OMOP CDM event tables than the provider\_id field in the PERSON table.

### Pattern: Handling Patient.managingOrganization

This pattern is managed the same as provider references.

## FHIR [Condition](https://hl7.org/fhir/r4/condition.html) Patterns <WIP to be continued>

The FHIR [Condition](https://hl7.org/fhir/R4/condition.html) resource has a loose definition of “*resource is used to record detailed information about a condition, problem, diagnosis, or other event, situation, issue, or clinical concept that has risen to a level of concern.*”

### Condition.category

Condition category does not require mapping since it is supporting information and there is no equivalent in OMOP CDM. By default, Condition.category lists 2 options: problem-list or encounter-diagnosis.

{

"resourceType": "Condition",

"id": "example",

"category": [

{

"coding": [

{

"system": "http://terminology.hl7.org/CodeSystem/condition-category",

"code": "encounter-diagnosis",

"display": "Encounter Diagnosis"

},

{

"system": "http://snomed.info/sct",

"code": "439401001",

"display": "Diagnosis"

}

]

}

], etc…

Note: This element however has an *extensible* binding, which allows for the implementer to add any other concept as long as it is not equivalent to the [Condition Category Codes](https://hl7.org/fhir/R4/valueset-condition-category.html) value set.

* Condition.recorder

### Condition.clinicalStatus

FHIR Condition.clinicalStatus is bound to the HL7 [condition-clinical](http://terminology.hl7.org/CodeSystem/condition-clinical) code system which contains codes that have no concept equivalent in OMOP.

### Patient-stated Conditions

We assume that FHIR conditions that have an asserter who is a diagnositican will be mapped to the CONDITION\_OCCURRENCE table.

For OMOP, patient-reported conditions should go to the OBSERVATION table.

Example Scenario: A patient’s caregiver completes a patient reported outcome form stating their father (the patient) has fever.

In FHIR, this is represented as a Condition resource with the following elements:

Condition.code = SNOMED-CT code: TBD “fever”

Condition.asserter = Reference(Patient)

What this means is that

## FHIR Observation Patterns <WIP to be continued>

<brief description of what FHIR Observeration resources are meant to cover, plus scope>…and general considerations of mapping FHIR Observations to OMOP.

The [FHIR Observation](https://www.hl7.org/fhir/observation.html) resource is a general construct which support diagnosis, monitor progress, determine baselines and patterns and even capture demographic characteristics, as well as capture results of tests performed on products and substances.

The OMOP CDM further qualifies this to either a MEASUREMENT or an OBSERVATION table. Because of the level of detail, further analysis and processing are needed to determine which of the two tables is more appropriate. Moreover, FHIR Observations contain components which may also result in linked records within the OMOP CDM.

The following sections provide further guidance and tips on how to do this.

### Map by Category

The FHIR Observation.category code is helpful in determining the OMOP domain id where there is no OMOP standard concept for Observation.code.

|  |  |  |  |
| --- | --- | --- | --- |
| **Code** | **Display** | **Definition** | **Map to OMOP table** |
| social-history | Social History | Social History Observations define the patient's occupational, personal (e.g., lifestyle), social, familial, and environmental history and health risk factors that may impact the patient's health. | Observation |
| vital-signs | Vital Signs | Clinical observations measure the body's basic functions such as blood pressure, heart rate, respiratory rate, height, weight, body mass index, head circumference, pulse oximetry, temperature, and body surface area. | Measurement |
| imaging | Imaging | Observations generated by imaging. The scope includes observations regarding plain x-ray, ultrasound, CT, MRI, angiography, echocardiography, and nuclear medicine. | Procedure |
| laboratory | Laboratory | The results of observations generated by laboratories. Laboratory results are typically generated by laboratories providing analytic services in areas such as chemistry, hematology, serology, histology, cytology, anatomic pathology (including digital pathology), microbiology, and/or virology. These observations are based on analysis of specimens obtained from the patient and submitted to the laboratory. | Measurement |
| procedure | Procedure | Observations generated by other procedures. This category includes observations resulting from interventional and non-interventional procedures excluding laboratory and imaging (e.g., cardiology catheterization, endoscopy, electrodiagnostics, etc.). Procedure results are typically generated by a clinician to provide more granular information about component observations made during a procedure. An example would be when a gastroenterologist reports the size of a polyp observed during a colonoscopy. | Observation |
| survey | Survey | Assessment tool/survey instrument observations (e.g., Apgar Scores, Montreal Cognitive Assessment (MoCA)). | ? |
| exam | Exam | Observations generated by physical exam findings including direct observations made by a clinician and use of simple instruments and the result of simple maneuvers performed directly on the patient's body. | ? |
| therapy | Therapy | Observations generated by non-interventional treatment protocols (e.g. occupational, physical, radiation, nutritional and medication therapy) | procedures |
| activity | Activity | Observations that measure or record any bodily activity that enhances or maintains physical fitness and overall health and wellness. Not under direct supervision of practitioner such as a physical therapist. (e.g., laps swum, steps, sleep data) | Observation |

### Pattern: Grouping of FHIR Observation components into OMOP

**Example**: [GenomicVariant](http://hl7.org/fhir/us/mcode/StructureDefinition-mcode-genomic-variant.html) is a FHIR Observation resource which contains multiple Observation components. These components are considered as one group associated with that one observation. For example:

[GenomicVariant::Observation.component:geneStudied](http://hl7.org/fhir/us/mcode/StructureDefinition-mcode-genomic-variant-definitions.html#Observation.component:geneStudied)

[GenomicVariant::Observation.component:dnaChange](http://hl7.org/fhir/us/mcode/StructureDefinition-mcode-genomic-variant-definitions.html#Observation.component:genomicDNAChange)

[GenomicVariant::Observation.component:aminoAcidChange](http://hl7.org/fhir/us/mcode/StructureDefinition-mcode-genomic-variant-definitions.html#Observation.component:genomicDNAChange)

etc…(all other components).

If a genomic variant is considered an OMOP MEASUREMENT record, then how does one link groups of measurements?

|  |  |  |
| --- | --- | --- |
| **Option** | **Pros** | **Cons** |
| Create a primary MEASUREMENT record for the top-level Variant that is only a central id for all other FHIR components in the Observation. Each component, a MEASUREMENT, will populate MEASUREMENT.measurement\_event\_id with the measurement\_id of the abstract central MEASUREMENT record. | Creates the grouper. | Explosion of MEASUREMENT records in one table. At least 8 records for components to one master MEASUREMENT record that is the variant grouper. |

## FHIR Medication Patterns <WIP to be continued>

**Pattern Description:** TBD

**Examples:** TBD

**Recommendation Options:**

If a medication is patient-stated, then it should be mapped to the OBSERVATION table

## FHIR ServiceRequest Patterns <WIP to be continued>

**Pattern Description:** TBD

**Examples:** TBD

**Recommendation Options:**

## FHIR Procedure Patterns <WIP to be continued>

# Advanced Mapping Challenges

## Handling complex mapping scenarios

### Bidirectional Transformations

### Rare Data Elements

## Managing discrepancies <WIP to be continued>

## Maintaining data fidelity <WIP to be continued>

# Ongoing Updates and Community Contributions

The cookbook is currently a work in progress and will evolve over time to incorporate new insights and community feedback. We encourage readers to contribute to future versions of the cookbook in several ways:

* Contacting the authors by email listed in the Contact Information section of this document.
* By entering in issues via our Github FHIR to OMOP Cookbook repository: <https://github.com/CodeX-HL7-FHIR-Accelerator/fhir2omop-cookbook/issues>
* Configuring [Github notifications](https://docs.github.com/en/account-and-profile/managing-subscriptions-and-notifications-on-github/setting-up-notifications/configuring-notifications) for new updates to the Cookbook.

# Contact Information

Please contact the authors for further information or questions:

* May Terry, MITRE - [mayT@mitre.org](mailto:mayT@mitre.org)
* Guy Livne, Kineret – Ministry of Health, Israel - [guy.livne@moh.gov.il](mailto:guy.livne@moh.gov.il)
* Qi Yang, IQVIA - [qi.yang1@iqvia.com](mailto:qi.yang1@iqvia.com)

# Appendix A: Abbreviations

| **Abbreviation** | **Description** |
| --- | --- |
| CDM | Common Data Model |
| ETL | Extract, Transform, Load |
| FHIR | Fast Healthcare Interoperability Resources |
| HL7 | Health Level 7 |
| IG | Implementation Guide |
| OMOP | Observational Medical Outcomes Partnership (OMOP) |
| THO | HL7 Terminology |

# Appendix B: Mapping recommendations for specific scenarios

|  |  |  |
| --- | --- | --- |
| **Guidance Recommendation** | **Relevant FHIR profiles and/or OMOP tables** | **Additional Remarks** |
| Set all OMOP \_type\_concept\_id fields translated from FHIR to concept\_id = 32880 (*Standard algorithm*) | All OMOP Clinical Data tables, except PERSON[[2]](#footnote-2). |  |
| Only FHIR statuses that are *completed* or *final* will be mapped to OMOP |  | The OMOP CDM assumes that interventions have been performed. |

# Appendix C: Comparing FHIR and OMOP Frameworks

## Modeling in FHIR

FHIR is a framework with a flexible organization or categorization of elements.

The following position statements frame the organization and content of this cookbook:

1. HL7 FHIR and OHDSI are frameworks for which the FHIR information model and the OMOP CDM are only one part.
2. Both FHIR and OHDSI frameworks contain data models, message standards, process and tools aim to address FAIR principles of being findable, accessible, interoperable, and reusable.
3. Only FHIR statuses that are *completed* or *final* will be mapped to OMOP, the OMOP CDM assumes that interventions have been performed.

## Framework Design Assumptions

It is important to understand the foundational assumptions of the HL7 FHIR and OHDSI frameworks in order to effectively create a meaningful and sustainable alignment among both communities.

A summary of the comparisons between both is displayed below.

|  |  |  |
| --- | --- | --- |
| **Characteristic** | **FHIR** | **OMOP** |
| Primary purpose[[3]](#footnote-3) | Provider/patient/organizational data exchange. | A persistent data store optimized for large-scale analytics |
| Information model | Technology-agnostic. Logical models do not assume a specific data store. | Tied to a relational database schema. |
| Semantic Model | A collection of acknowledged HL7 internal and external terminologies.  There is no overarching semantic model which binds the same concepts from multiple terminologies or that ties concepts together through multiple relationships. | Ontology-centric containing 1) standard vocabularies with aligned concepts to an overarching OMOP concept id, and 2) non-standard vocabularies.  Standard concepts are assigned a domain id which determines which OMOP table to populate. |
| Messaging | Primarily, a REST-based API with representations in JSON, XML, or TTL. Ideally messaging-agnostic with support for other messaging formats such as SOAP or Document-based. | Records-based through SQL and R-based libraries. |

## General Design Principles - OMOP CDM

The [Book of OHDSI Chapter 4](https://ohdsi.github.io/TheBookOfOhdsi/CommonDataModel.html) describes the general design principles. Here we elaborate its implications to FHIR.

Optimized for observational studies using large scale analytics. This drives the information modeling design choices which potentially introduces lossiness from the source due to the constraints of institutional review boards (IRBs) and a set of data elements which are critical for research, and consequently which limit the type of data that can be shared in research.

**OMOP focuses on the concepts that were already completed**. For example, medications that were ordered but not administered might not exist in DRUG\_EXPOSURE.

**OMOP is mostly a** **storage model**: primarily an entity-relationship model adopting entity-attribute-value relationships most suitable for relational databases where accessibility is “typically achieved through the SQL interface[[4]](#footnote-4)”.

**OMOP aims to be *data-source agnostic***. Metadata specific to the data source should not be stored in OMOP. FHIR-contextual data should be left out of OMOP. For example:

The FHIR Profile metadata Condition.meta.profile

**OMOP is “*semantic first*”**, meaning that it relies heavily on its ontology to decide the entities where the mappings occur. For OMOP, each Standard Concept has a unique Domain assignment, which defines which table they are recorded in. Even though the correct Domain assignment is subject for debate in the community, this strict Domain table field correspondence rule assures that there is always an unambiguous location for any code or concept. Reference the OMOP CDM Conventions for further information.

**OMOP has limited cardinality for some tables**. For example, the OMOP PERSON table has only one provider\_id row.

**OMOP is currently a “last mile” data model**. This data is usually populated in an aggregated data store and not re-shared or re-translated.

**OMOP concepts are “closed-world”**. The OMOP Ontology presumes that all of the concepts needed for an observational study already exist. While the OMOP CDM schema allows population of values as string in the event that an OMOP concept id does not exist, doing this risks errors of omission for existing cohort definitions.

## General Design Principles - FHIR Framework

FHIR is a framework that includes a standardized but extensible information model based on resources, standardized operations, and tooling. Although some FHIR solutions exist with the intent of being a persistent data store, much of the focus on FHIR is in *data exchange and transport* among systems, and in some cases, bidirectionally. However, as FHIR evolves towards research-oriented use cases, reference implementations which were initially designed as a translational data hub are now evolving towards being a persistent data store.

**FHIR is currently treated as a transport standard** (although this is quickly evolving). While FHIR is popular in its adoption and some institutions use it as a persistent store, FHIR is mostly an exchange standard.

**FHIR lacks a single cohesive ontology**. The FHIR Community has a Healthcare Terminology Authority (HTA) which acknowledges and manages an identification of internal and external vocabularies. However, the FHIR framework itself does not have a comprehensive terminology which relates concepts across all acknowledged terminologies. When combined with a loose semantic binding for FHIR base resources, this potentially leads to a mapping of non-standard or non-existent concepts in the OMOP Ontology. Because of this, FHIR profiles can be used to further constrain code systems bound to a specific element. Historically, SNOMED and UMLS comes closest to providing an publicly-available ontology similar to what OMOP has. Some commercial terminology/content vendors offer a proprietary solution in which codes are harmonized to one overarching concept (e.g.: 3M, IMO).

**FHIR allows for several ways to structurally represent a clinical data construct**.

For example, a [FHIR MedicationRequest](http://hl7.org/fhir/R4/medicationrequest.html) resource can represent the coded medication in two ways:

1. As a codeable concept directly in the MedicationRequest resource
2. As a reference to a separate [Medication](http://hl7.org/fhir/medication.html) resource

Without any constraints specified in a profile, a data source can choose either method of representing the medication concept. From an ETL perspective, this means that, unless there is a FHIR profile which constrains the MedicationRequest resource to only one way to represent the code, a FHIR-to-OMOP translator will need to check and process both methods of representation.

**The FHIR framework allows for customized representations of its base and core resources in order to meet specific business or clinical requirements**. These customizations are represented in structured artifacts as “[FHIR profiles](https://www.hl7.org/fhir/profiling.html)”, FHIR Operations, and FHIR Conformance Rules tailored for a specific use case. FHIR specifiers can then aggregate these structured artifacts into a FHIR implementation guide (IG) which can then be shared with the FHIR implementation community. As an example, the [minimum Common Oncology Data Elements (mCODE) FHIR IG](http://hl7.org/fhir/us/mcode/) contains FHIR profiles, operations, and conformance statements that derive from FHIR base resources.

# Appendix D: Resources and References for further reading

## FHIR References

|  |  |
| --- | --- |
| **Reference** | **Comments** |
| [Introducing HL7 FHIR](http://hl7.org/fhir/summary.html) | An executive summary of the HL7 FHIR specification. |
| [FHIR Specification](http://hl7.org/fhir/) | A next generation standards framework created by HL7. FHIR combines the best features of HL7's [v2](http://www.hl7.org/implement/standards/product_brief.cfm?product_id=185), [HL7 v3](https://www.hl7.org/implement/standards/product_brief.cfm?product_id=186), and [Clinical Document Architecture (CDA)](http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7) product lines while leveraging the latest web standards and applying a tight focus on implementability. |

## OHDSI References

|  |  |
| --- | --- |
| **Reference** | Comments |
| [Book of OHDSI](https://ohdsi.github.io/TheBookOfOhdsi/) | This book aims to be a central knowledge repository for OHDSI, and it focuses on describing the OHDSI community, OHDSI data standards, and OHDSI tools. |
| [OHDSI Glossary](https://ohdsi.github.io/TheBookOfOhdsi/Glossary.html) | Definitions for common terms used in OHDSI documentation. |
| [OMOP Common Data Model](https://ohdsi.github.io/CommonDataModel/) | An open community data standard, designed to standardize the structure and content of observational data and to enable efficient analyses that can produce reliable evidence. The current version is CDM v5.4 |

1. https://www.ohdsi.org/ohdsi-hl7-collaboration/ [↑](#footnote-ref-1)
2. <https://ohdsi.github.io/CommonDataModel/cdm54.html#Clinical_Data_Tables> [↑](#footnote-ref-2)
3. FHIR has multiple purposes. The most mature use-case for FHIR is listed in this table. [↑](#footnote-ref-3)
4. Book of OHDSI, p26. [↑](#footnote-ref-4)