dHealth 2023
B. Pfeifer et al. (Eds.)

doi:10.3233/SHTI230004

From Authoring to Evaluating an Electronic Health Quality Measure - Applying Logic to FHIR® with CQL for Calculating Immunization Coverage

Anna M. LIN[a,1], Anja SCHWAB[a], Reza ABOLHASSNI[b] and Stephan WINKLER[a]
[a] *Bioinformatics Research Group,*
University of Applied Sciences Upper Austria, Hagenberg, Austria
[b] *Smile Digital Health, Toronto, Canada*

Abstract. Background: Current monitoring and evaluation methods challenge the healthcare system. Specifically for the use case of immunization coverage calculation, person-level data retrieval is required instead of inaccurate aggregation methods. The Clinical Quality Language (CQL) by HL7®, has the potential to overcome current challenges by offering an automated generation of quality reports on top of an HL7® FHIR® repository. Objectives: This paper provides a method to author and evaluate an electronic health quality measure as demonstrated by a proof-of-concept on immunization coverage calculation. Methods: Five artifact types were identified to transform unstructured input into CQL, to define the terminology, to create test data, and to evaluate the new quality measures. Results: CQL logic and FHIR® test data were created and evaluated by using the different approaches of manual evaluation, unit testing in the HAPI FHIR project, as well as showcasing the functionality with a developed user interface for immunization coverage analysis. Conclusion: Simple, powerful, and transparent evaluations on a small population can be achieved with existing open-source tools, by applying CQL logic to FHIR®.

Keywords. Clinical Quality Measures, Immunization Coverage, Clinical Quality Language CQL, Health Level 7 HL7

1. Introduction

1.1. Problem Statement

Health organizations around the world spend billions of dollars on monitoring and evaluation (M&E). The Global Fund [1] recommends that 5-10% of their funds for fighting HIV, Tuberculosis, and Malaria, should be awarded towards M&E. The current M&E practices challenge the system by diverting health workers from providing clinical care, high costs for professionals and data aggregation tools as well as the lack of data quality and comparability. [2]

Current methods to monitor immunization coverage include (1) inaccurate estimations by aggregating the total number of administered doses and (2) conducting surveys of representative households, with the disadvantages of high costs and non-timely information on an immunization program's performance. [3]

[1] Corresponding Author:
Anna Lin, University of Applied Sciences Upper Austria, Hagenberg, Austria, anna.lin@fh-hagenberg.at

1.2. Background

In 2020, Health Level 7 (HL7®) introduced CQL [4] as a mostly normative specification to express and share clinical knowledge. The FHIR Quality Measure Implementation Guide [5] describes how CQL can be combined with the FHIR Clinical Reasoning module [6] to represent electronic Clinical Health Measures (eCQMs). The on-demand retrieval of patient-level information from an electronic health record (EHR) enables continuous and up-to-date monitoring.

1.3. Objective

This paper describes how eCQMs can be authored and evaluated by using CQL and FHIR®, and presents a proof-of-concept (POC) on a population's immunization coverage analysis based on the *Routine Schedule* of the Publicly Funded Immunization Schedules for Ontario [7] ("the schedule"). The three user groups targeted by this POC are:

- *Patient*: current immunization status (fully immunized: yes/no)
- *Physician*: immunization status of each patient and coverage of all their patients
- *Public Health*: immunization coverage of an entire population

The POC should support the following, to overcome the issues described in Section 1.1:

- On-time, automatic person-level data retrieval from a FHIR-based EHR
- Reusable measures that satisfy the requirements of all three user groups
- Modifiable evaluation dates for simple comparability of scores

2. Methods

Five artifact types were identified for creating a CQL quality measure, as outlined in Figure 1. The following sections provide a methodological description for each type.

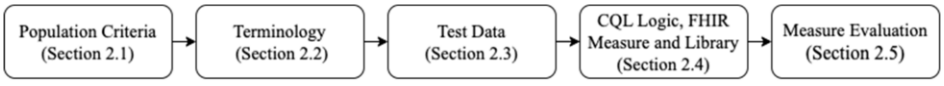

Figure 1. Five artifact types need to be created when authoring and evaluating a CQL quality measure.

2.1. Design of the Population Criteria

The first step is the transformation of unstructured input to structured definitions for electronic processing. The schedule [7] consists of a table displaying visual elements to indicate the timing of vaccine doses and eligibility criteria such as age or pregnancy status. Based on these eligibility criteria, we divided the schedule into three categories:

- *Routine*: vaccinations that are due at a predefined age
- *Seasonal*: annual vaccination, previous immunizations irrespective (influenza)
- *Pregnancy*: vaccinations that are due in the context of the risk factor *pregnancy*

For evaluating immunization coverage, a *proportion measure* [8] is applied. This measure calculates a score by dividing the number of patients who satisfy all quality criteria (*NUM*) by the number of eligible patients (*DENOM*). The Initial Population (*IP*) pre-defines a set of patients that are considered for calculation initially. Table 1 depicts the identified population criteria for each category of the schedule.

Table 1. Population criteria are defined for each immunization measure category.

	Routine	**Seasonal**	**Pregnancy**
IP	Age > 2 months	Age ≥ 6 months	Pregnant female > 10 years
DENOM	IP	IP	IP
NUM	Fully routine vaccinated	Valid influenza vaccination in current influenza season	One dose of the Tdap vaccine during the current pregnancy

2.2. Specification of the Terminology

Two types of FHIR® ValueSets are required: (1) For each defined antigen, which is referred to as "vaccine" in the schedule, one ValueSet containing all codes of valid vaccine formulas is needed. Existing ValueSets can be retrieved from the US ValueSet Authority Center (VSAC). (2) For evaluating a pregnancy status, two more ValueSets are needed, containing the LOINC codes for pregnancy status and weeks of gestation.

2.3. Creation of FHIR® Test Data

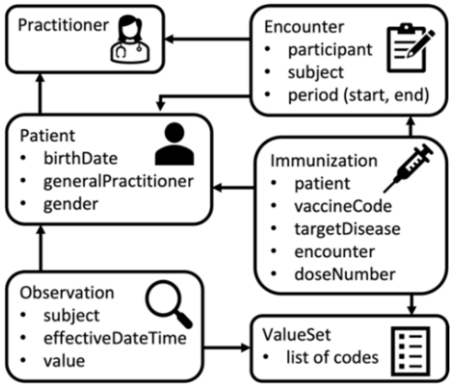

Figure 2. Relevant FHIR resources and interactions.

Figure 2 depicts the FHIR® resources and elements used for test data creation. This is an extended version of the minimum data set for person-level data sharing defined by IHE for an immunization use case [2]. A *Patient* receives an *Immunization* during an *Encounter* that is performed by a *Practitioner*. A patient's primary relationship to a physician is represented by the element *generalPractitioner*. *Observations* with specified codes indicate pregnancies. All test data of this POC is based on FHIR® R4 version `4.0.1`.

2.4. Implementation of the CQL logic, the related FHIR® Library and Measure

The population criteria of Table 1 translate to one CQL expression for each population. As a tool for authoring CQL, VS Code with its freely available CQL plugin "Clinical Quality Framework" offers a variety of functionality that helps implement and unit test the CQL expressions without a need for a separate FHIR® server.

The finalized CQL content can be published and evaluated on a FHIR® server, embedded in a FHIR® *Library* as an encoded *content* element of type *text/cql*. This Library is referenced by a FHIR® *Measure* resource as specified in [5]. Table 2.4.2-1 of [2] summarizes the essential elements of such FHIR® Measure and Library resources.

2.5. Evaluation of the Measure

The HAPI FHIR project (https://hapifhir.io/) is an Open Source Global Good implementation of the HL7® FHIR® standard. It supports the operation $evaluate-measure in a fully functioning server setup. The list below summarizes the steps for setting up and using the HAPI FHIR JPA server starter project to evaluate an eCQM:

1. Ensure that CQL is enabled according to the instructions in the README file
2. Ingest FHIR® data (i.e.ValueSes for terminology definitions as well as test data)
3. Ingest the CQL dependencies, such as the FHIR® asset Libraries *FHIRHelpers*
4. Ingest the FHIR® Measure and Library resources for the newly created eCQM
5. Call the FHIR® Operation $evaluate-measure and interpret the MeasureReport

3. Results

A broad range of different results were created from the work on this POC. Test data and CQL logic has been created, tests were carried out in the HAPI FHIR project with various evaluation requests, and a user interface (UI) has been developed.

3.1. FHIR® Test Data and CQL

A total of 1482 FHIR® resources was created, with a total of 85 test patients divided up among the three identified immunization measure categories. We created multiple test patients for each age group (i.e. column in the PDF schedule) to cover different scenarios of fully-, partly-, and not-at-all-immunized patients. Additionally, there are patients that do not satisfy the eligibility criteria to qualify for the IP. All test resources are created with an evaluation date of *16.09.2022* in mind.

Three separate CQL libraries were defined - one for each immunization measure category - containing the logic according to the population criteria as depicted in Table 1. Named CQL expressions were created within the CQL library to represent the specified populations. The CQL expression for the IP of the *routine* category is depicted in Listing (1). The entire CQL library for the routine category can be extracted from the content of the FHIR® Library in the HAPI FHIR unit test (see Section 3.2).

```
define "Initial Population":
          AgeInMonthsAt(date from start of "Measurement Period") >= 2   (1)
```

3.2. Testing in HAPI FHIR

The immunization use case including the resulting test data and CQL logic has been utilized to test the current implementation of the CQL module in the HAPI FHIR open-source project (https://hapifhir.io/).

A Java unit test for the *routine* measure with different evaluation scenarios has been defined (class *CqlMeasureEvaluationR4ImmunizationTest*). This unit test asserts that all test data for routine immunization is loaded and evaluated successfully.

In addition to the unit testing, the HAPI FHIR JPA server starter project has been used for manual testing. Evaluation with different dates, user groups, and measure categories passed testing, except for evaluations on large populations: An error occurred when evaluating an IP consisting of >20 patients. The error has been reported. The progress on a fix can be followed here: https://github.com/hapifhir/hapi-fhir/issues/4025

3.3. Identified HTTP Requests for Evaluation

For one immunization measure category, the requirements of all three user groups that are targeted by this POC can be addressed by using a single CQL library and FHIR Measure. Query parameters as depicted in the HTTP requests in Table 2 allow for filtering patients. With the parameter *subject* defined, the measure is executed for only the specified single patient. This satisfies the need of a (1) single patient user. The *practitioner* parameter is defined as a pre-filter of patients, to restrict measure evaluation to for patients with a primary relationship to the specified practitioner, as required by the (2) physician user. If no parameter is defined, a summary report for the entire population is generated as required by the (3) public health user.

The parameters *periodStart* and *periodEnd* define the evaluation date of the measure and allow for easy comparisons over time. For immunization coverage calculation, both start and end parameters contain the same date due to ambiguous results by aging patients when using an evaluation window of more than one day.

Table 2. Standardized HTTP GET requests for evaluating the routine measure, split up by user group.

User Group	Patient User	Physician User	Public Health User
GET Request according to the FHIR® specification	`[server]/Measure/Routine/$evaluate-measure` **`?subject=`**`Imm-pat-5&periodStart=2022-09-16&periodEnd=2022-09-16`	`[server]/Measure/Routine/$evaluate-measure` **`?practitioner`**`=Imm-prac-3&periodStart=2022-09-16&periodEnd=2022-09-16`	`[server]/Measure/Routine/$evaluate-measure?periodStart=2022-09-16&periodEnd=2022-09-16`

3.4. Interpretation of the FHIR® Measure Report

A FHIR® MeasureReport is generated once the evaluation process is completed. For the proportion measure, the most relevant element is the *measureScore,* which stores the percentage of immunization coverage. For a report on a single patient, the score can either be 0.0 or 1.0, indicating incomplete coverage (0.0), or full coverage (1.0). Scores in between are relevant for the evaluation of multiple patients. For each population (i.e. IP, DENOM and NUM), the total number qualifying patients based on the population criteria is returned in the *count* element of the MeasureReport. When working with the report type *subject-list*, in addition to the total number of patients, a list of references to qualifying patients can be retrieved for each population.

3.5. User Interface

We built a user interface (UI) to illustrate the user interactions as well as the information that can be presented to the user with a set-up as described in Section 2.5. Figure 3 depicts the UI for a public health user. The three immunization measure categories are represented by three different FHIR® Measures that are evaluated separately in the background for an entered evaluation date. The total number of patients in the IP differs for each category, despite equal underlying data. This is due to different population criteria, that allow for patients as young as 2 months old to be part of the IP for routine immunization, whereas for the seasonal category, only patients >6 months are considered. The pregnancy category is even more restrictive for consideration in the IP, since only patients with a corresponding pregnancy Observation are considered eligible.

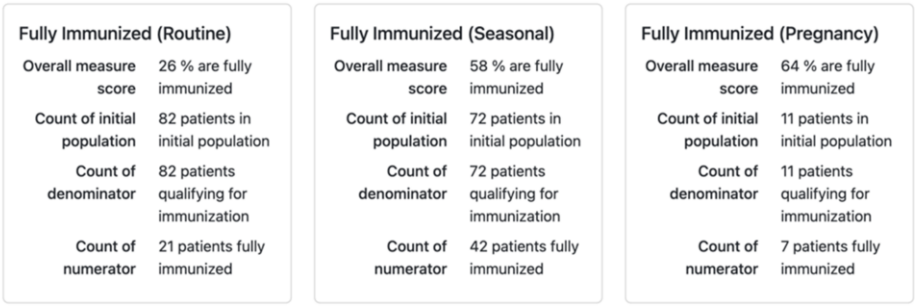

Figure 3. User interface of the public health view to analyze immunization coverage.

4. Discussion

This POC was created to analyze the capabilities of CQL and HAPI FHIR for the calculation of immunization coverage. Decisions were made without clinical expertise.

The ValueSets for vaccine codes were downloaded from the US-based VSAC, whereas the utilized schedule refers to Canadian legislation, which causes discrepancies.

A powerful technical enhancement would be the usage of *stratifiers*. which can divide measure populations into separately scored segments of interest. Stratifiers are not yet supported by HAPI FHIR but are on the roadmap with an upgrade of the CQL engine.

Funding

We are very grateful that this work has been funded by two parties: (1) *Digital Square*, a PATH-led initiative funded by USAID, the Bill & Melinda Gates Foundation, and a consortium of other partners. (2) *Smile Digital Health* as a company research partner.

References

[1] The Global Fund, https://www.theglobalfund.org/en/, last access: 11.1.2023.
[2] IHE Quality, Research and Public Health Planning Committee, Extracting Indicators from Person Level Data, https://www.ihe.net/uploadedFiles/Documents/QRPH/IHE_QRPH_White_Paper_CQL.pdf, last access: 11.1.2023.
[3] S. V. Sodha, V. Dietz, Strengthening routine immunization systems to improve global vaccination coverage, *Br Med Bull* **113**(1):5-14 (2015)
[4] HL7 International, Clinical Quality Language Release 1, https://cql.hl7.org, last access: 11.1.2023.
[5] HL7 International, Quality Measure Implementation Guide (STU3), https://hl7.org/fhir/us/cqfmeasures/, last access: 16.3.2023.
[6] HL7 International, FHIR Clinical Reasoning, http://www.hl7.org/fhir/clinicalreasoning-module.html, last access: 11.1.2023.
[7] Queen's Printer for Ontario, Publicly Funded Immunization Schedules for Ontario, https://www.health.gov.on.ca/en/pro/programs/immunization/docs/Publicly_Funded_ImmunizationSchedule.pdf, last access: 11.1.2023.
[8] IHE Quality, Research and Public Health Technical Committee, Electronic Clinical Quality Measure (eCQM) Standards Landscape, https://www.ihe.net/uploadedFiles/Documents/QRPH/IHE_QRPH_WP_eCQM_Standards.pdf, last access: 11.1.2023.