FHIR – The Future of Interoperability. Now

FHIR for Populations: Understanding the FHIR Bulk Data API Certification Criteria (Part 1)

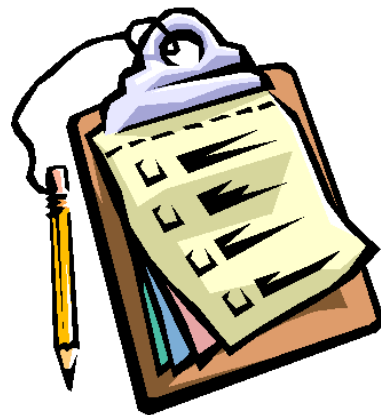Diego Kaminker diego@hl7.org

HL7 DCSIO

Agenda – April 20

- Session Structure
- Part 1
  - The FHIR Toolbox: FHIR Core Spec & IGs + …
  - Scenarios for the FHIR Toolbox
  - Population Health: BULK FHIR / CQL + Measures
  - Intro to Bulk FHIR
  - Bulk FHIR in Regulation
- Part 2
  - Discussion Zone: Example Scenarios
  - Questions/Answers

FHIR Tool Box

Domain/Country Specific Content, Behavior and Terminology

example: US CORE

FHIR Core Spec

Implementation Guides

SMART-on-FHIR

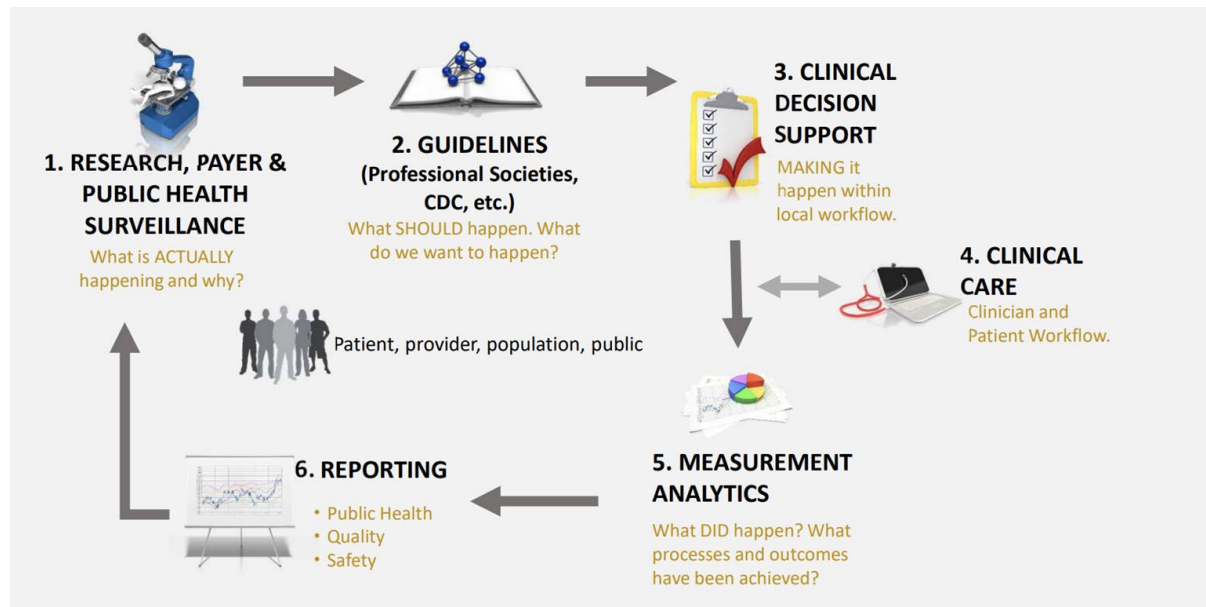Syntax Structure API Methods

CDS-HOOKS

Bulk FHIR

April 18 10-10:25

Today!

The FHIR Toolbox

The stack of FHIR family tools from HL7 supporting a **learning health care system**

- HL7 FHIR Core Spec
- HL7 FHIR Base IGs
- HL7 FHIR IGs
- Smart-on-FHIR
- CDS-Hooks
- Bulk FHIR Data Access
- CQL

1. RESEARCH, PAYER & PUBLIC HEALTH SURVEILLANCE
What is ACTUALLY happening and why?

2. GUIDELINES (Professional Societies, CDC, etc.)
What SHOULD happen. What do we want to happen?

3. CLINICAL DECISION SUPPORT
MAKING it happen within local workflow.

4. CLINICAL CARE
Clinician and Patient Workflow.

Patient, provider, population, public

5. MEASUREMENT ANALYTICS
What DID happen? What processes and outcomes have been achieved?

6. REPORTING
- Public Health
- Quality
- Safety

HL7 International

FHIR Tool Box: Knowledge

How can we apply medical knowledge in daily practice?

How can we overcome limitations of our EHRs and use/integrate other apps or services?

Which are the different scenarios?

Which FHIR tools apply?
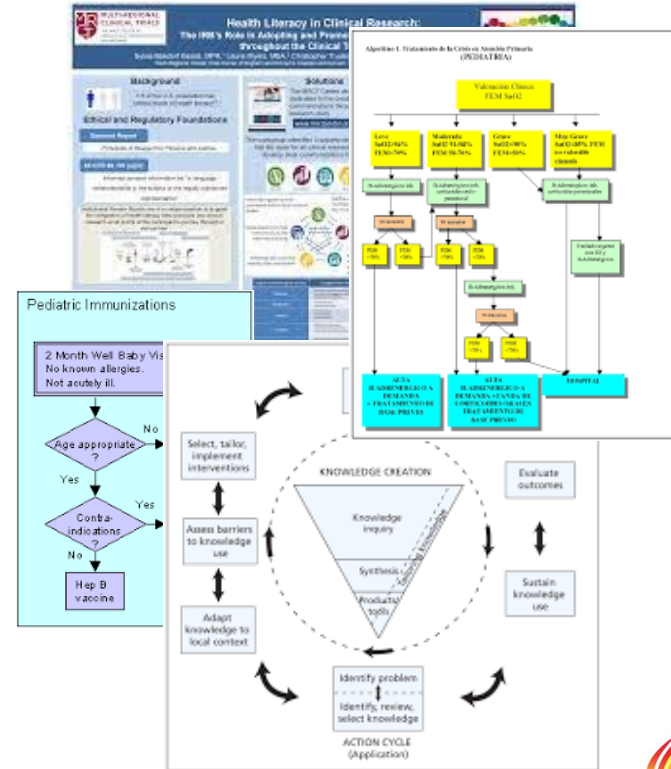
Scenarios

When do we want to apply knowledge?

Individual Patient

(We are attending a patient, or the patient is using a patient portal)
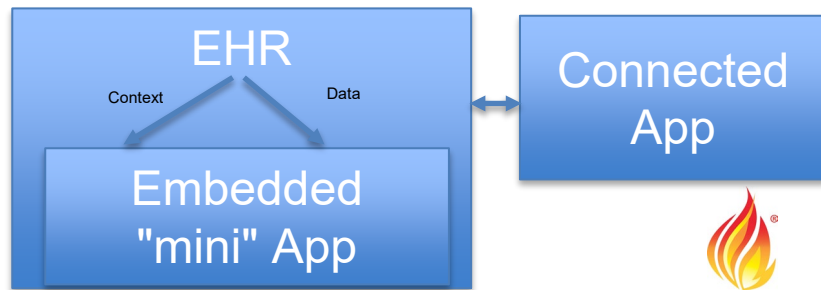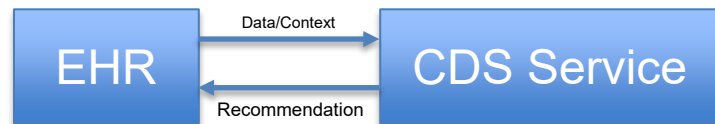
➔ **Independent Knowledge (CDS) Service**

We want to provide any EHR with a service: given context and data, apply knowledge and recommend an action course

➔ **Embedded/Shared User Interface**

We want to show information in a different way, integrated with the EHR/Patient Portal

CDS HOOKS™

EHR → Data/Context → CDS Service
CDS Service → Recommendation → EHR

SMART

EHR
Context   Data
Embedded "mini" App
↔ Connected App

HL7 International

Scenarios

Population Health

We want to analyze data from a (big) set of patients

➔ **Data Extraction**

      We want to extract data from the EHR for research, or apply discovery algorithms, or to calculate quality / clinical / epidemiological measures
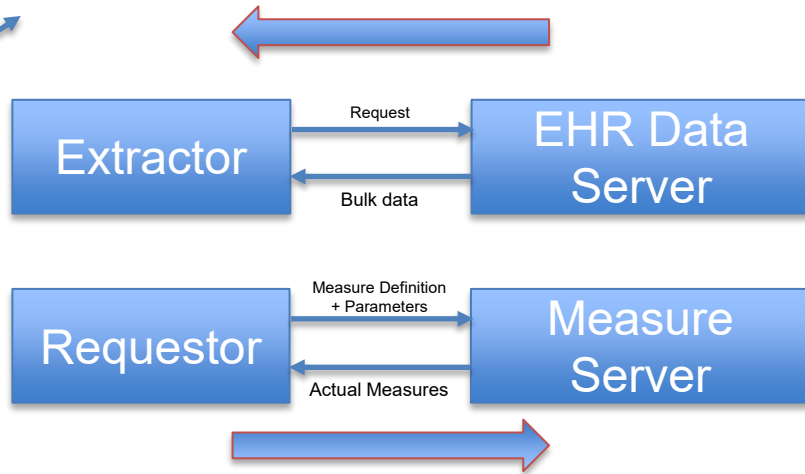
Also, for initial load of a native FHIR server

➔ **Population Health Measures**

      We want to use a service to calculate measures on a given population -> (numerator / denominator)

FHIR BULK
I ask for the data, and then process it

Extractor	Request → ← Bulk data	EHR Data Server

Requestor	Measure Definition + Parameters → ← Actual Measures	Measure Server

(I ask a FHIR Server to calculate)

CQL

HL7 International

Data Extraction

"Big" Data Extraction

We need **clinical/financial data about patients**, to do some kind of analysis: **quality measures, public health reporting, data mining, ML training, etc.**

➜ **Extractor:** requesting some data / parameters for extraction (which data?)

➜ **Data Server:** prepares the data and alert when the data is available (the process is asynchronic)

We are talking about SEVERAL (HUNDRED) THOUSANDS of records
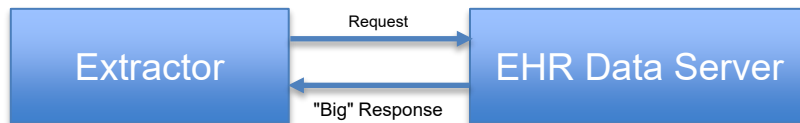
"Big" Data / "Extraction" examples

- "Recover vitals, demographics, labs of all patients with 'diabetes' in the last 2 years"

- "Recover all procedures for patient covered by XYZ network in the last 3 months"

- "Recover all medication, clinical notes and demographic data for all COVID19 patients"

HL7
International

POPULATION HEALTH

Data Extraction

Problems?

➜ How to guarantee security
➜ Async Processing
➜ Efficient format
➜ One API for each EHR?
➜ Different format for each extraction?

Different API for the same goal
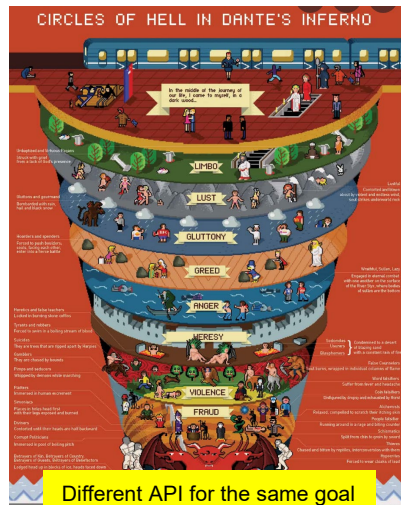
Who is interested?

- EHR Vendors / Data Exporters
- Public Health/Payors/Researchers
(EXT: Extractors)

EHR 1:Easy! Call this API
EHR 2:We cannot do that!
EHR 3: Give us 4 months

EXT 1:Please provide with a CSV file with these fields
EXT 2:Create an XLS file with these columns
EXT 3: Call this service with this XML for each case

HL7 International

FHIR BULK DATA

STU 1 12/2020

STU 2 11/2021

https://hl7.org/fhir/uv/bulkdata/

Efficient Data Exchange	ND-JSON
Security Model	Smart Back-End Services Authorization http://docs.smarthealthit.org/authorization/backend-services/
Asynchronous Data Exchange	FHIR $export API
Clinical Data Exchange	FHIR Resources (+Appropiate data model:IG)

Bulk FHIR

- Register your App

  - When registering the server requires a key-pair (Public/Private) and provides:
    - Client_ID
    - Auth URL
    - FHIR Bulk Data Server URL

Authentication

- Create a JWT with iss/sub = client_id, aud = url de auth, digitally signed with our private key

- The 'scope' is 'what I want to access'. Pretty open. But must be as small as possible (remember BOLA)

Conexión con el endpoint de autorización

(SMART Backend Services OAuth Profile)

```python
# Armo el jwt con el claim utilizando el client id y la clave privada
jwt_claims = {
    "iss": client_id,
    "sub": client_id,
    "aud": token_endpoint,
    "exp": datetime.datetime.utcnow() + datetime.timedelta(minutes=5),
    "jti": uuid.uuid4().hex
}

client_assertion = jose.jwt.encode(
    jwt_claims,
    signing_key,
    algorithm='RS384',
    headers={"kid": signing_key["kid"]}
)

token_request = {
    "scope": "system/*.read",
    "grant_type": "client_credentials",
    "client_assertion_type": "urn:ietf:params:oauth:client-assertion-type:jwt-bearer",
    "client_assertion": client_assertion
}
```

```python
[18] ## Solicito el token de acceso
token_response = requests.post(token_endpoint, data=token_request)
```

Scope:What I am requesting?

HL7 International

Kick-off request

- Standardized Operation **($export)**
- Which kind of resources we want
- Asynchronous : **Prefer: respond-async**
- Export all patients, a group of patients, or all the resources.

    - *[FHIR Server Base]*/Patient/$export

    - *[FHIR Server Base]*/Group/[group id]/$export

    - *[FHIR Server Base]*/$export

Groups are arbitrary and can be pre-negotiated

HL7
International

Kick-off request parameters

- $export parameter (v1)   *=Mandatory*

_outputFormat	Only **ndjson**
_since *	Last update date
_type *	Resource classes, comma separated
_typeFilter	Optional Filtering criteria for resources

New in V2 (optional)
_typeFilter: *(criteria)*
_elements: *(only these)*
Patient *(sp.patients ref)*
includeAssociatedData (last/all *provenance*)

```
#Lanzamos la solicitud con una operación $export
kickoff_request_headers = {
    'Prefer': 'respond-async',
    'Accept': 'application/fhir+json',
    'Authorization': 'Bearer ' + token_response.json()["access_token"]
}
kickoff_request_params = {'_type':'Observation,Condition,Patient'}
kickoff_response = requests.get(fhir_endpoint + '/Patient/$export', params=kickoff_request_params, headers=kickoff_request_headers)
```

Kick-off response

It will return **202 (Accepted)** and an URL where we can ask when our 300,000,000 records will be ready for download…

```
{'Server': 'Cowboy', 'Connection': 'keep-alive', 'X-Powered-By':
'Express', 'Content-Location': 'https://bulk-
data.smarthealthit.org/fhir/bulkstatus/cfe81d727b7f2e08e35ad77efdefe6ad'
, 'Content-Type': 'application/json; charset=utf-8', 'Content-Length':
'644', 'Etag': 'W/"284-F3AUtGqq00JHYRvr+vui14KUsQE"', 'Date': 'Sun, 14
Nov 2021 23:03:33 GMT', 'Via': '1.1 vegur'}
```

HL7
International

Polling

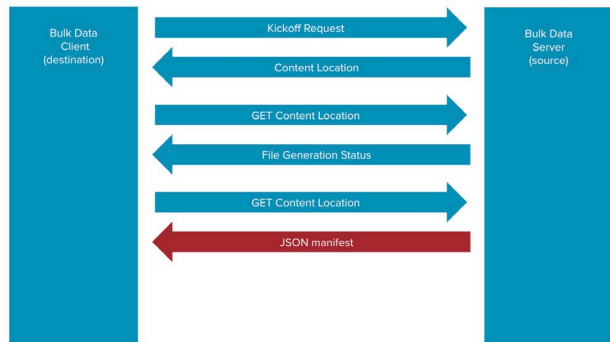We ask in Content-Location until we get a 200 instead of 202
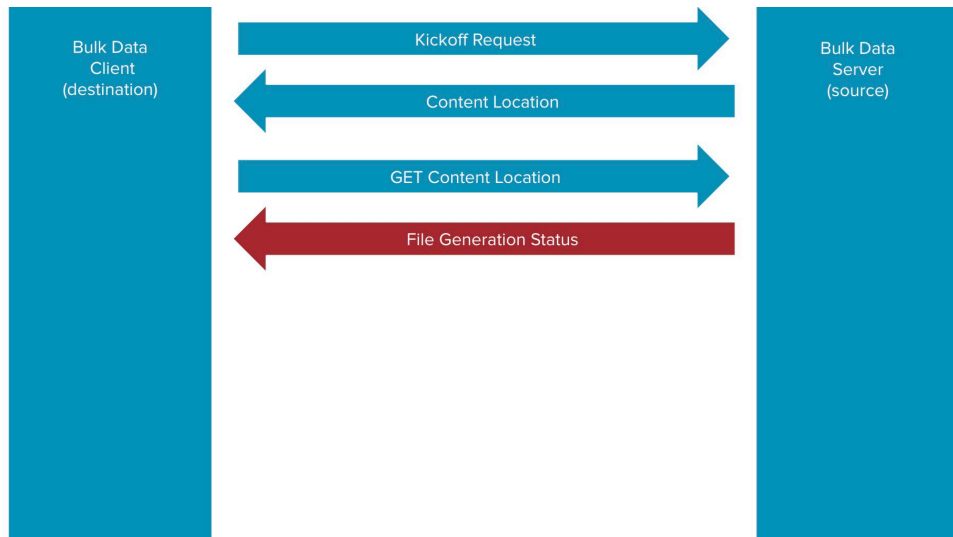
Status: 202 Accepted
X-Progress: "50% complete"
Retry-After: 120

HL7 International

Polling

And when we get 200…

Response code: 202…
Response code: 200
Response body: { "transactionTime": "2021-11-14T23:03:33.337Z",
"request": "https://bulk-data.smarthealthit.org/eyJlcnIiOiIiLCJwYWdlIjoxMDAwMCwiZHVyIj...SI6MSwic3R1Ijo0LCJkZWwiOjB9/fhir/Patient/$export?_type=Observation%2CCondition%2CPatient"
, "requiresAccessToken": true,
"output":
  [{ "type": "Condition", "count": 639,
"url": https://bulk-data.smarthealthit.org/eyJpZCI6ImNmZTgxZDcyN2I3ZjJlMDhlMzVhZD...2Zmc2V0IjowLCJsaW1pdCI6NjM5LCJzZWN1cmUiOnRydWV9/fhir/bulkfiles/1.Condition.ndjson} …}]

The manifest tells us how many resources per file and the url for the files

HL7
International

FHIR resources
(But not just ANY FHIR resource : constrained!)

Structure: FHIR Resources
Specific profile so we know what to expect
on content/terminology

Efficient: ndjson

n resources in the same
file. No Spaces or CR/LF
http://ndjson.org/

FHIR core
Patient.identifier: 0..*
Patient.name 0..*
Patient.gender:0..1

'Our' FHIR
Patient.identifier: 1..*
Patient.name: 1..*
Patient.gender: 1..*
Observation.code.system
=LOINC

HL7
International

And then…?

The standarization "succeds" when we get our ndjson files

- We can upload our FHIR resources to any JSON / FHIR native repository (Google BigQuery - example) and then use them

- Can be also ELK Elastic/Logstash/Kibana

- **$import** operation: early proposal https://github.com/smart-on-fhir/bulk-import/blob/master/import.md

```
SELECT

json_extract      ( Observation.json      , '$.    subject.reference      ')
paciente    ,

json_extract      ( Observation.json      , '$.    effectiveDateTime      ')
fecha  ,

json_extract      ( coding.value      , '$.code') test,

json_extract      ( Observation.json      , '$.    valueQuantity.value      ')
resultado

FROM

Observation
```

HL7
International

But how? BULK FHIR to SQL Example

- https://bit.ly/3xmChS1
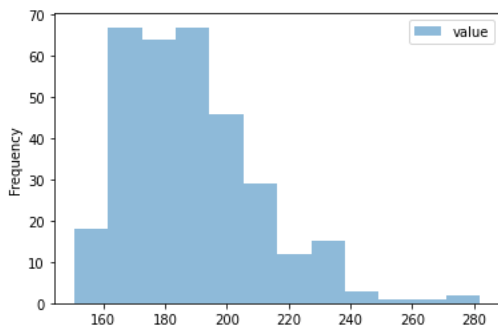
- Obtains Patient / Observations for all patients in the EHR
- Create an Histogram for Total Cholesterol

EHR → *Bulk FHIR* → Data Analyst

HL7 International

The Standard

- 2019 – HL7 Publishes the first version of the Bulk Data Access FHIR Implementation Guide (STU1)

- 2021 – Second version of Bulk Data Access (STU2) published
- http://hl7.org/fhir/uv/bulkdata/STU2/index.html

Regulation (a)

(1) ONC's [Cures Act Final Rule](#) added a new certification criterion to the ONC Health IT Certification Program at §170.315(g)(10) for standardized APIs for patient and population services.

(2) ONC also adopted a suite of regulatory provisions that address health IT developers' business practices associated with certified API technology through the API Condition of Certification. Health IT developers seeking certification to this [criterion](#) are required to implement an API using the HL7 FHIR standard to make patient data available to users.

Regulation (b)

- The criterion also requires developers to build capabilities for standardized export of large quantities of health information according to the Bulk Data IG.

- Support of STU1 is mandatory, but STU2 can be also supported

- Adoption of "bulk data" APIs by certified health IT products represents the culmination of more than five years of work to enable more efficient access to large sets of health data.

- Support for FHIR API for Patient Access
- Support for FHIR Bulk API for Provider/Payor Access

Implementers Tools (I)

- [Bulk Data Server](#)– A configurable reference implementation of a bulk data server conforming to the Bulk Data IG.

- [Bulk Data Client](#)– A sample client application using a command-line interface to test bulk data servers.

Implementers Tools (II)

- [Bulk Data Test Suite](#)– A web-based online tester for bulk data servers that contains a test suite and test runner. The test suite has been integrated into the [Inferno](#) test tool.

- [Bulk Data Tools](#)– A library for working with and converting bulk data into different formats (json, ndjson, csv, tsv)

Discussion (1)

https://bcda.cms.gov/

Beneficiary Claims Data API

Provides FHIR-formatted bulk data files to an ACO for all of the beneficiaries eligible to a given Shared Savings Program ACO. BCDA does not require individual beneficiary authorization but does allow a process for patients to opt out of data sharing.

BCDA's Parameters

Parameters are options that can be passed to the endpoint to influence the response. Parameters can be used to filter or select for certain desired data. We would like to highlight two query parameters used in BCDA.

1. The type parameter

The _type parameter allows you to request different Resource Types from the API. Instead of receiving data from all three Resource Types when no _type parameter is specified, you will be able to use the _type parameter to submit one or more Resource Types. The API will then produce data from the specified Resource Types.

To learn more about the Resource Types see the FHIR Resource Types section above.

2. The since parameter

The _since parameter allows you to apply a date parameter to your bulk data requests. Instead of receiving the full record of historical data every time you request data from an endpoint, you will be able to use _since to submit a date. The API will then produce claims data from the bulk data endpoints that have been loaded since the entered date.

To learn more about the differences using _since between the /Patient and /Group endpoints see the FHIR endpoints section above.

BCDA Version 2 (V2)

Version 2 of the Beneficiary Claims Data API (BCDA V2) is available as of Summer 2021.

Discussion (2)

- What if?

Importing via .csv file Event-Level COVID-19 Vaccination Form- Residents - LTC Component

Table 1: NHSN Event-Level COVID-19 Vaccination Form- Residents Import File Format

Field	Requirement	Values	Format	Description of Field
orgID	Required	-	must be a whole number	Must be a valid NHSN Facility ID (organization identifier)
resID	Required	-	Character (15)	Resident identifier - a unique identifier for the individual, assigned by your facility
dob	Required	MM/DD/YYYY	Datetime	Resident Date of Birth
Resgender	Required for all records beginning 10/24/2022 *	F M O	Character (1)	Resident Gender F – Female M – Male O – Other/Unknown
resethnicity	Required for all records beginning 10/24/2022 *	HISP NOHISP DEC UNK	Character (6)	Resident Ethnicity HISP – Hispanic or Latino NOHISP – Not Hispanic or Latino DEC – Declined to respond UNK – Unknown
resrace	Required for all records beginning 10/24/2022 *	AMIN ASIAN AAB NH-PI WHITE DEC UNK	Character	Resident Race: AMIN – American Indian/Alaskan native ASIAN – Asian AAB – Black or African American NH-PI Native Hawaiian/Other Pacific Islander WHITE – White DEC – Declined to respond UNK- Unknown

resAdmitDate	Required	MM/DD/YYYY	Datetime	Resident Admit Date
resDischDate	Optional	MM/DD/YYYY Must be >= resAdmitDate	Datetime	Resident Discharge Date
resGName	Required	-	Character (30)	Resident First Name
resSurname	Required	-	Character (30)	Resident Last Name
dose1Date	Conditionally required (each record must contain At least ONE status- This means each record much be classified into at least one of the main categories, such as partial or complete primary series	MM/DD/YYYY	Datetime	Dose 1 vaccination date

Bulk FHIR!

Our discussion should be limited to:

Which patients?
Which FHIR resources?
Which elements?
Which criteria?