# Introduction to FHIR Bulk Data

National Institutes of Health (NIH), Office of the Director (OD), Office of Data Science Strategy (ODSS)

June 13, 2023

FHIR® is the registered trademark of Health Level Seven International (HL7). Use of the FHIR trademark does not constitute an HL7 endorsement of this work shop.

# What is FHIR Bulk Data?

- HL7 specification, published in November 2021

- [https://hl7.org/fhir/uv/bulk_data/](https://hl7.org/fhir/uv/bulk_data/)

From the spec:

> Providers and organizations accountable for managing the health of populations often need to efficiently access large volumes of information on a group of individuals. For example, a health system may want to periodically retrieve updated clinical data from an EHR to a research database.

# Why do we need FHIR Bulk Data?

From the spec:

> Existing FHIR APIs work well for accessing small amounts of data, but large exports can require hundreds of thousands of requests.

FHIR Bulk Data provides a simpler, standard mechanism to extract large amounts of FHIR-format data from an EHR.

# Why now?

The Office of the National Coordinator for Health Information Technology (ONC)'s Health IT Certification Program requires adoption of FHIR Bulk Data APIs:

- "ONC Supports Adoption and Implementation of Bulk Data APIs"

- §170.315(g)(10) Standardized API for patient and population services

- FHIR Bulk Data group-export required for certification deadline December 2022 for USCDI data elements
- US Core FHIR Implementation Guide

# Aside: EHI export

- ONC's program also requires population-level data export ("EHI export"):

§170.315(b)(10) Electronic Health Information export

> *Patient population electronic health information export*. Create an export of all the electronic health information that can be stored at the time of certification by the product, of which the Health IT Module is a part.

§170.315(b)(10) fact sheet (pdf)

> must be available to their customers December 31, 2023

- *Can* be implemented with FHIR Bulk Data, but does not necessarily have to be FHIR Bulk Data
  - FHIR(ish) example: athenaPractice (née GE Centricity) https://docs.mydata.athenahealth.com/fhir-r4/ehiexport.html, which uses NDJSON, FHIR resources, and some custom resources defined in this Implementation Guide.

# Vendor support

- Widespread vendor support for FHIR Bulk Data export of groups of patients thanks to December 2022 ONC certification deadline

- List of vendors with §170.315(g)(10) certification: https://chpl.healthit.gov

  - Filter for Certification Criteria → 2015 Cures Update Certification Criteria → 170.315 (G)(10): Standardized Api for Patient and Population Services (Cures Update)

- Examples: Epic, Cerner

# FHIR Bulk Data specifications

Implementations use these specs:

- [https://hl7.org/fhir/uv/bulkdata/export.html](https://hl7.org/fhir/uv/bulkdata/export.html) defines the request flow for retrieving Bulk Data from a FHIR server

- [https://www.hl7.org/fhir/smart-app-launch/backend- services.html further describes the](https://www.hl7.org/fhir/smart-app-launch/backend-services.html) authorization/authentication flow

# Types of Bulk Data requests

- Patient, to obtain resources related to all Patients

- Group, to obtain resources associated with a particular [Group](Group)

- System, to obtain all resources, whether or not they are associated with a patient

# High level FHIR Bulk Data request steps

- Get an access token from the FHIR server

- Request: GET
  [https://fhir.example.org/Group/[id]/$export](https://fhir.example.org/Group/[id]/$export)

- Receive the requested FHIR data ("NDJSON" format)

  Can be split into multiple files by the server to support arbitrarily large data sets

# Characteristics of research that may benefit from FHIR Bulk Data

- Need for direct access to (identified) EHR data for a specific group of patients
  - Note that there is an [open source library](#) from Microsoft for de- identifying FHIR Bulk Data
- The necessary data elements are part of USCDI, or otherwise included in EHR's FHIR implementation
- Real-time access is not needed (other FHIR APIs are better for this)
- One-off data pull or regular updates
- Multi-site research without existing shared CDM

# Non-EHR uses of FHIR Bulk Data

From this FHIR DevDays 2022 presentation:

- CMS ACO Beneficiary Claims Data (pilot)

- CMS Data at the Point of Care (pilot)

- CMS Claims to Part D S ponsors (pilot)