Introduction

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, a provider may want to send clinical data on a roster of patients to their ACO to calculate quality measures, or an EHR may want to access claims data to close gaps in care. In most cases, access to these bulk-data exports is pre-authorized between the data holder and the data requester. The data exchange involves extracting a specific subset of fields from the source system, mapping the fields into a structured file format like CSV, and persisting the files in a server from which the requester can then download them into the target system. This multi-step process increases the cost of integration projects and can act as a counter-incentive to data liquidity.

Existing FHIR APIs work well for accessing small amounts of data, but large exports can require hundreds of thousands of requests. The FHIR Bulk Data Implementation Guide outlines a standardized, FHIR based approach for exporting bulk data from a FHIR server to a pre-authorized client.

The Boston Children’s Hospital (BCH) SMART Health IT team acted as a subcontractor to HL7 to develop this guide and create and support an open-source, developer focused FHIR bulk data reference server implementation. We engaged in several 2018 FHIR Connectathon events to test the bulk data server and implement changes to the specification as guided by experts within the community. As part of this subcontract, the BCH team also wrote a white paper defining a recommended approach for secure, compliant access to consented EHR data, describing use cases, requirements, and business needs for bulk data access using the FHIR API specification.

FHIR Bulk Data Meeting & White Paper

In December 2017, on behalf of the Office of the National Coordinator for Health Information Technology (ONC), The Boston Children’s Hospital Computational Health Informatics Program and SMART hosted a Population Level Data Export / FLAT FHIR Meeting to discuss standardizing bulk data exports from EHR systems and data warehouse environments. The
meeting brought together key stakeholders from government, industry and academia to understand and guide the technical roadmap and regulatory environment for population level data export from health information systems using FHIR.

Following this meeting, together with HL7 and the ONC, the BCH SMART Health IT team published a white paper, “Push Button Population Health Data: Extending the HL7® FHIR® Standard to Support Bulk Data Export”, which concluded that, with a shift toward value based care, the healthcare system needs better data analytics and population health management tools. Major health systems are already using FHIR APIs to export population-level data and aggregate data from multiple sources, in order to support innovative new tools and solutions. A FHIR bulk data API will allow for liquidity of population-level data, which means lower costs, increased competition in the health IT market, and a spur in innovation. Whether the data is being used for reporting quality measures, machine learning, disease registries, or data mining to identify high-risk patients, the ability to combine data from multiple sources in a standardized way would be a major next step toward the learning health care system that has been envisioned.

The SMART and FHIR standards have already been adopted by major EHR vendors and health systems throughout the United States, and SMART apps are being used today to help clinicians better utilize the data and information available to make decisions at the point of care. Using web and mobile applications written against the SMART on FHIR standards, which have adopted the profiles outlined in the Argonaut Implementation Guide, patients and providers are accessing and using health data to improve the health and lives of individuals.

SMART continues to work with the ONC, HL7 and the FHIR community to extend the FHIR standard to support bulk data export, leveraging existing FHIR and SMART work to improve the liquidity of population level data.

**Reference Implementation**

A reference server that implements the FHIR Bulk Data API is live at [https://bulk-data.smarthealthit.org/](https://bulk-data.smarthealthit.org/). SMART also posted the code at [https://github.com/smart-on-fhir/bulk-data-server](https://github.com/smart-on-fhir/bulk-data-server) so users can install and run the server locally. It has a number of innovative features such as the ability to produce sample data files containing millions of FHIR Resources (it loops through a base set of data and alters the ids and references on the fly), simulate server errors such as an expired token when testing clients, and configure parameters like the time it takes to return a response so client apps can ensure they respond correctly to slow servers (Figure 1). In addition, SMART has built a reference client that works with any server supporting the FHIR Bulk Data API Implementation Guide and posted the code at [https://github.com/smart-on-fhir/sample-apps-stu3/tree/master/fhir-downloader](https://github.com/smart-on-fhir/sample-apps-stu3/tree/master/fhir-downloader) (Figure 3). At FHIR Connectathons, the SMART reference client was tested with servers from Epic, Cerner and Care Evolution.
### Launch Options

**Authentication:**
- JWKS URL (recommended)
- JWKS

Your backend service should expose an endpoint that returns the JWKS as JSON. Also, make sure it only hosts your public keys.

**Resources per File:**
- 10,000

### Advanced

**Access Token Lifetime:**
- 15 minutes

**Simulate Error for Testing:**
- None

**Database Size:**
- 100 Patients

**Simulated file generation duration:**
- 10 Seconds

### Launch Configuration

**FHIR Server URL:**
- [https://bulk-data.smarthealthit.org/try/bulk-data.smarthealthit.org](https://bulk-data.smart-data.smarthealthit.org/try/bulk-data.smarthealthit.org)

**Authentication URL:**
- [https://bulk-data.smarthealthit.org/auth/token](https://bulk-data.smart-data.smarthealthit.org/auth/token)

**Client ID:**

---

Figure 1: SMART bulk data server reference implementation
Figure 2: Integrated web-based client application
The initial live version of the reference implementation was used by participants at the January 2018 FHIR Connectathon, with updated versions used at the May 2018 Connectathon, FHIR Developer Days in June 2018 and the September 2018 Connectathon.

**Bulk Data Implementation Guide**

Over the course of 2018, the Bulk Data Implementation Guide was expanded and refined based on community feedback and a security review sponsored by the Argonaut Project. Significant improvements include:

**FHIR Export Operation**
- Created completion data structure to return metadata on operation
- Incorporated a system wide $export operation to encompass resources like ValueSet
- Added “Error” and “Count” properties to completion response
- Added experimental _typeFilter query parameter to limit data being exported

**SMART Backend Services Authorization**
- Updated recommended access token lifetime to five minutes
- Clarify client registration mechanisms including public keys
  - Clarify rules for servers verifying authentication JWTs
- Added error handling expectations (per OAuth 2 base spec)
The Draft FHIR Bulk Data Implementation Guide is available at
https://github.com/smart-on-fhir/fhir-bulk-data-docs/

Discussion: Success & Uptake

There is strong community interest in the use of FHIR for bulk data access. The standard will be
balloted through HL7 in 2019, and the bulk data specification is already getting significant
uptake. The Centers for Medicare and Medicaid Services is piloting the use of a FHIR Bulk Data
Server to exchange claims with Medicare ACOs.

Google is establishing a bulk data server as well.