Journal of the American Medical Informatics Association, 28(6), 2021, 1284–1287
doi: 10.1093/jamia/ocab028
Advance Access Publication Date: 1 March 2021
Brief Communications

Brief Communications

A landscape survey of planned SMART/HL7 bulk FHIR data access API implementations and tools

James Jones[1], Daniel Gottlieb[1,2], Joshua C. Mandel[1,3], Vladimir Ignatov[1], Alyssa Ellis[1], Wayne Kubick[4], and Kenneth D. Mandl[1,2,3]

[1]Computational Health Informatics Program, Boston Children's Hospital, Boston, Massachusetts, USA, [2]Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, USA, [3]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA, and [4]Health Level Seven International, Ann Arbor, Michigan, USA

***Corresponding Author:** Kenneth Mandl, 300 Longwood Avenue, Boston, MA 02115, USA; Kenneth_Mandl@Harvard.edu

Received 6 November 2020; Editorial Decision 29 January 2021; Accepted 3 February 2021

ABSTRACT

The Office of National Coordinator for Health Information Technology final rule implementing the interoperability and information blocking provisions of the 21st Century Cures Act requires support for two SMART (Substitutable Medical Applications, Reusable Technologies) application programming interfaces (APIs) and instantiates Health Level Seven International (HL7) Fast Healthcare Interoperability Resources (FHIR) as a lingua franca for health data. We sought to assess the current state and near-term plans for the SMART/HL7 Bulk FHIR Access API implementation across organizations including electronic health record vendors, cloud vendors, public health contractors, research institutions, payors, FHIR tooling developers, and other purveyors of health information technology platforms. We learned that many organizations not required through regulation to use standardized bulk data are rapidly implementing the API for a wide array of use cases. This may portend an unprecedented level of standardized population-level health data exchange that will support an apps and analytics ecosystem. Feedback from early adopters on the API's limitations and unsolved problems in the space of population health are highlighted.

Key words: medical informatics, health information system, applications

INTRODUCTION

A rule published by the Office of the National Coordinator of Health Information Technology (ONC) implements the application programming interface (API) provisions of the 21st Century Cures Act.[1] The rule requires that certified health information technology (IT) support 2 open APIs and authorization frameworks which were designed to enable a health app ecosystem and access to population datasets, developed by the SMART (Substitutable Medical Applications, Reusable Technologies) Health IT team and standardized by Health Level Seven International (HL7). The first, SMART on FHIR (Fast Healthcare Interoperability Resources),[2] provides patient- and provider-facing data access to apps for individual patients. The second, the SMART/HL7 FHIR Bulk Data Access API,[3] affords access to population-level data.[4]

At the time of the survey, compliance with the standardized API functionality for bulk export in the rule was required by May 2022. It has since been extended to December 2022.[5] We sought to measure early progress toward implementations of the API—both among organizations taking part in the ONC Health IT Certification Program[6] as well as others not regulated under the rule—to understand use cases, timelines and barriers.

MATERIALS AND METHODS

Design

This was a survey study.

Participants

We reached out to 37 organizations known to be implementing the SMART/HL7 FHIR Bulk Data Access API[3] in existing or new systems. Within each organization, we identified representatives who had detailed knowledge of the technological approaches undertaken. The initial list of organizations was generated from participants in the 2019 Meeting to Advance Push Button Population Health,[7] and expanded by soliciting suggestions for additional respondents from these representatives and other key stakeholders. In some cases, our initial point of contact directed us to another representative as the optimal respondent within their organization.

Survey content

The survey (Supplementary Appendix) solicits information about implementation progress and deployment of FHIR servers and tools implementing and advancing the SMART/HL7 FHIR Bulk Data Access Implementation Guide.[3] Initial versions of the survey were revised with input from experts in academia, open standards, and American National Standards Institute standards bodies. Questions were split between implementations focused on delivering bulk data (servers) and implementations focused on requesting, receiving, and consuming the data (clients), while offering organizations the ability to answer both sets of questions if appropriate. The survey asked for information on standardized features of the API likely to be supported and on timelines for any implementations planned ahead of the regulated timeline for certified health IT.

Survey process

Initial contact was made through email, requesting respondents either complete the survey online or set up a 15- to 30-minute slot to answer the questions over the phone. For the first month, weekly email reminders were sent, dropping off to biweekly after the first month until the target of at least 20 respondents was achieved. To promote uninhibited sharing of information, it was made clear that individual organizations and product names would not be identified in reports.

RESULTS

Survey results

Of the 37 organizations approached, 22 (59%) responded. Three were surveyed by phone, 17 filled out the survey online, and 2 responded online during a screen share.

Seven organizations were developing both FHIR bulk data servers and clients based on the API, 12 were developing servers alone, and 3 were developing clients alone. Respondent organizations were categorized as follows: payors (n = 4), research institutions (n = 3), cloud vendors (n = 5), FHIR tooling developers (n = 4), electronic health record (EHR) vendors (n = 4), and other purveyors of health IT platforms (n = 2).

When asked to identify use cases for implementing bulk data, only EHR vendors indicated that the primary motivation was to meet regulatory requirements. Figure 1 shows how many times the 9 most common healthcare use cases were chosen, including: value-based care (11 times), public health reporting (10 times), and data sharing—with payors (10 times), researchers (10 times), accountable care organizations (9 times), across health information exchanges (9 times), and across sites within an organization (8 times). Machine

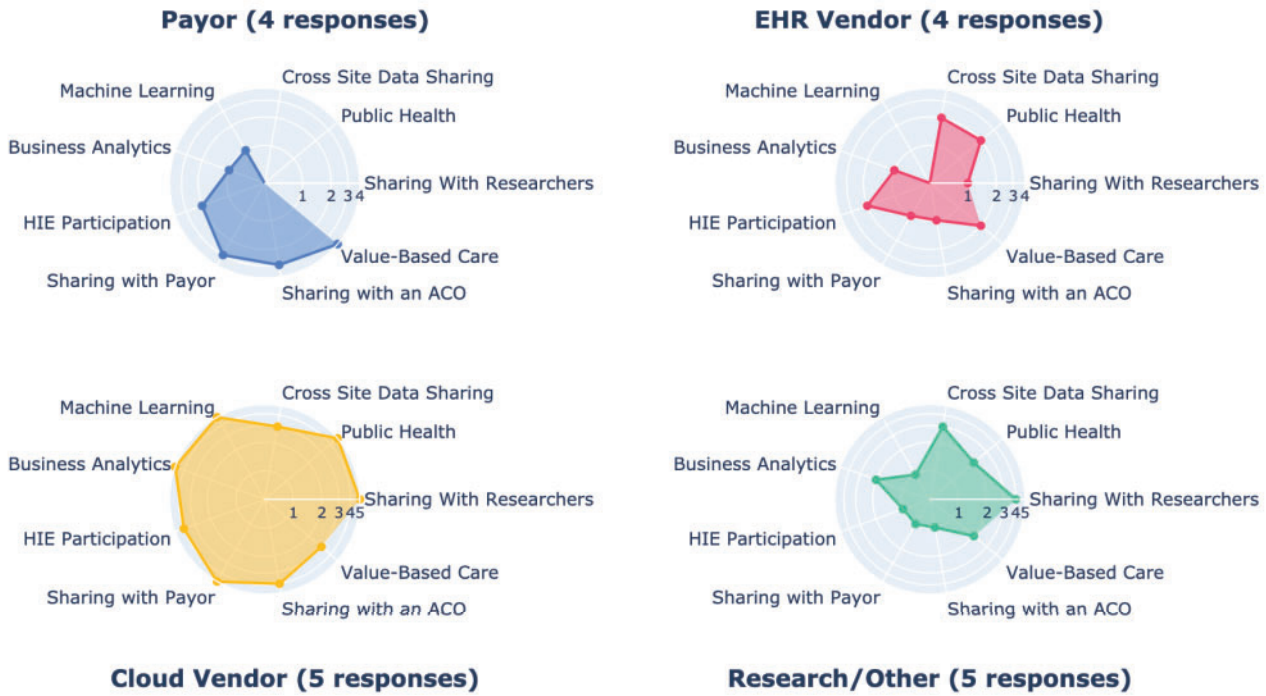

Figure 1. Respondent-selected bulk data use cases by organization type. Polar charts logging number of times each use case was selected by developers from each organization type: payor, electronic health record vendor, cloud vendor, and research/other. Omitted are 4 responses from bulk FHIR (Fast Healthcare Interoperability Resources) tooling organizations, as no respondent in that category selected any specific use cases. ACO: accountable care organization

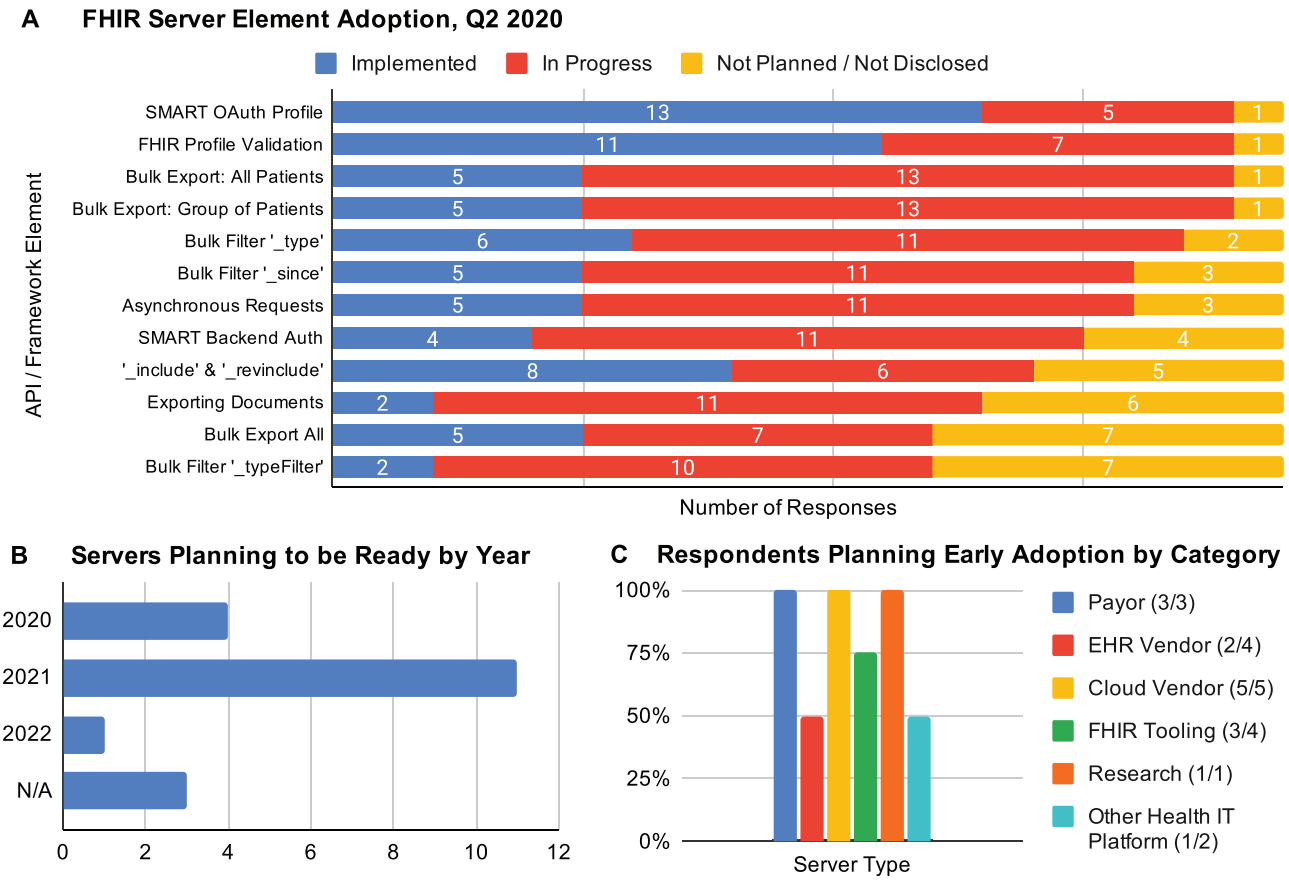

Figure 2. FHIR (Fast Healthcare Interoperability Resources) bulk data early adoption landscape. (A) Progress toward individual application programming interface (API) and framework elements of 19 servers. (B) Planned timelines for bulk data implementations to be ready for interoperability. (C) Breakdown by category of the observed ratios between servers that expressed plans to complete implementations before May 2022 and those that did not. Raw positive and total responses are represented in the legend as numerator and denominator. EHR: electronic health record; IT: information technology; N/A: no answer given.

learning (7 times) and business analytics (9 times) were also of interest to cloud vendors, researchers, and some payors. In general, researchers, payors, and EHR vendors had divergent motivations for implementing the bulk data API, with some overlap. Cloud vendors selected the full range of common use cases. The 4 respondents developing FHIR tooling did not identify any particular use cases besides advancing the standard itself and are omitted from the figure.

FHIR Bulk Data server adoption

Server-specific questions were answered by 19 organizations. Progress and plans toward individual elements of the FHIR Bulk Data specification were requested, with respondents able to select "Currently supported," "Planning to support by May 2022," and "Not planning to support." Answers are shown in Figure 2. Notably, 15 of the 19 server implementers reported plans to be ready ahead of the deadline. Features commonly expected for early rollout include the bulk data operations for exporting all patients and individual groups, along with basic filtering on resource type and last-updated timestamp. The SMART backend authorization framework and asynchronous request pattern are in scope for 15 of the 19 servers, while the optional "_typeFilter" query parameter and exporting all server contents were flagged as out of scope or too costly by 7 respondents. Further, 7 implementers are planning to support uploading data directly to cloud buckets where possible, and support for serializing data in the Parquet format is in scope in 5 implementations (in addition to the required NDJSON format).

Notable implementation hurdles listed in free text were (1) hardware limitations and logistics for moving large datasets, including error handling, processing time, deciding where to split large files, and how best to load them; (2) managing granular access, particularly in federated systems and where the user requesting a bulk export needs to be explicitly identified for audit purposes; and (3) de-identifying data stored in documents and free text fields when leveraging the exported data for some use cases. To reduce data hosting costs, 2 respondents expressed intentions to fulfill requests from within their own organization via streams of data, rather than pointing to generated files.

Of the 14 developer teams with experience previously implementing OAuth for apps, 4 had already implemented the SMART backend authorization workflow and 6 had implemented FHIR's asynchronous request pattern to enable nonblocking requests of bulk data. None of the 5 organizations without experience implementing SMART on FHIR apps considered their Bulk Data implementations complete.

DISCUSSION

Though the initial SMART on FHIR standard took 11 years from conception to regulation,[1,8] SMART/HL7 FHIR Bulk Data Access

has moved much more rapidly, taking only 2 years from conception to regulation under ONC's rule in mid-2020. Encouragingly, we discovered that many organizations not required to implement the API under the regulation are planning for rollout, even ahead of the original 2022 date. Additionally, developer teams appear to have an easier lift to implement various aspects of bulk data if they were already supporting patient and practitioner-facing apps through SMART on FHIR. In free-text sections of the survey, respondents identified implementation challenges for feedback into the specification, some of which can be addressed ahead of the 2022 deadline: patient group management and efforts to limit oversharing data beyond minimum necessary amounts are being iterated upon by the standards community. Federated access control and de-identification tools may merit more attention.

One limitation of this study is its focus on organizations that had previously expressed interest in adopting the bulk FHIR specification; extrapolation of results to organizations not active in the FHIR development community is difficult. Another limitation is that projections for functionality and timelines given by the surveyed organizational representative are only estimates and subject to change.

It remains to be seen how the industry will adapt to the availability of standardized population-level healthcare data coming out of these APIs, particularly as open source tooling and standardized methods to import and manage bulk data files after downloading is an area of ongoing development. The marked interest in this space by cloud vendors suggests that bulk data capabilities will offer a common path for accessing cloud-based health data at scale, even if downstream data processing and analytics are not yet standardized across cloud platforms.

Implementation of SMART/HL7 bulk FHIR access is off to a strong start, with many implementations on track ahead of the regulatory requirement. Using EHR data to drive healthcare delivery and research will require affordable computation, storage, and transfer. As the landscape evolves and community-driven updates are integrated into the FHIR APIs and guidance, repeating this survey periodically may provide useful benchmarks, especially if the compliance deadline were to be extended again. The resources, costs, and charges for moving and storing data by EHR and cloud vendors will have a large impact on the robustness of the health information economy. Monitoring these factors periodically will be important as well.

FUNDING

This work was supported by 90AX0022/01-00 and 90AX0019/01-01 from the Office of the National Coordinator of Health Information Technology and by 1U01TR002623 from the National Center for Advancing Translational Sciences/National Institutes of Health.

AUTHOR CONTRIBUTONS

All authors contributed significantly to the conception and design of the study and drafting and revising the analysis and text, and approve of the final version of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online

DATA AVAILABILITY STATEMENT

The data underlying this article will be shared on reasonable request to the corresponding author.

CONFLICT OF INTEREST

Boston Children's Hospital receives corporate philanthropic support for KDM's laboratory from SMART Advisory Committee members that include the American Medical Association, the BMJ Group, Eli Lilly and Company, First Databank, Google Cloud, Hospital Corporation of America, Microsoft, Optum, Premier Inc, and Quest Diagnostics. KDM is an advisor to Medal, Inc and has advised Merck on use of real-world evidence. KDM is the chief architect for Microsoft Healthcare. DG is Principal, FHIR and Healthcare Data Standards, for Central Square Solutions. WK is CTO of Health Level Seven International.

REFERENCES

1. Office of the National Coordinator of Health Information Technology in 45 CFR Parts 170 and 171, RIN 0955–AA01 (Department of Health and Human Services, Federal Register; 2020).
2. Mandel JC, Kreda DA, Mandl KD, *et al.* SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc* 2016; 23 (5): 899–908.
3. FHIR Bulk Data Access. https://hl7.org/fhir/uv/bulkdata/index.html Accessed October 1, 2020.
4. Mandl KD, Gottlieb D, Mandel JC, *et al.* Push Button Population Health: The SMART/HL7 Bulk Data Access Application Programming Interface. *NPJ Digit Med* 2020; 3: 151.
5. ONC's Cures Act Final Rule. https://www.healthit.gov/curesrule/download Accessed November 2, 2020.
6. About The ONC Health IT Certification Program. https://www.healthit.gov/topic/certification-ehrs/about-onc-health-it-certification-program Accessed February 25, 2021.
7. SMART Health IT and Health Level 7 Meeting to Advance Push Button Population Health: SMART/HL7 Bulk Data Export/FLAT FHIR. http://smarthealthit.org/wp-content/uploads/SMART-2019_FHIR-Bulk-Data-Meeting_final.pdf Accessed October 1, 2020.
8. Mandl KD, Mandel JC, Murphy SN, *et al.* The SMART Platform: early experience enabling substitutable applications for electronic health records. *J Am Med Inform Assoc* 2012; 19 (4): 597–603.