Journal of the American Medical Informatics Association, 30(3), 2023, 485–493
https://doi.org/10.1093/jamia/ocac246
Advance Access Publication Date: 22 December 2022
Research and Applications

Research and Applications

Introducing HL7 FHIR Genomics Operations: a developer-friendly approach to genomics-EHR integration

Robert H. Dolin[1], Bret S.E. Heale[2], Gil Alterovitz[3,4], Rohan Gupta[5], Justin Aronson[6], Aziz Boxwala[1], Shaileshbhai R. Gothi[7], David Haines[8], Arthur Hermann[9], Tonya Hongsermeier[1], Ammar Husami[10,11], James Jones[12], Frank Naeymi-Rad[13], Barbara Rapchak[13], Chandan Ravishankar[13], James Shalaby[1], May Terry[14], Ning Xie[15], Powell Zhang[16], and Srikar Chamala[17,18]

[1]Elimu Informatics, El Cerrito, California, USA, [2]Humanized Health Consulting, Salt Lake City, Utah, USA, [3]Brigham and Women's Hospital, Boston, Massachusetts, USA, [4]Harvard/MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, Massachusetts, USA, [5]Shri Mata Vaishno Devi University, Katra, Jammu and Kashmir, India, [6]Northwestern University, Chicago, Illinois, USA, [7]Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, Florida, USA, [8]Leap of Faith Technologies, Libertyville, Illinois, USA, [9]Department of Health IT Strategy & Policy, Kaiser Permanente, Pasadena, California, USA, [10]Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA, [11]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA, [12]Computational Health Informatics Program, Boston Children's Hospital, Boston, Massachusetts, USA, [13]Leap of Faith Technologies, Libertyville, Illinois, USA, [14]MITRE Corporation, McLean, Virginia, USA, [15]Biomedical Cybernetics Laboratory, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA, [16]Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, [17]Keck School of Medicine, Department of Pathology, University of Southern California, Los Angeles, California, USA and [18]Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, Los Angeles, California, USA

Corresponding Author: Robert H. Dolin, MD, Elimu Informatics, 1709 Julian Ct, El Cerrito, CA 94530, USA; bdolin@elimu.io

Received 27 September 2022; Revised 16 November 2022; Editorial Decision 23 November 2022; Accepted 2 December 2022

ABSTRACT

Objective: Enabling clinicians to formulate individualized clinical management strategies from the sea of molecular data remains a fundamentally important but daunting task. Here, we describe efforts towards a new paradigm in genomics-electronic health record (HER) integration, using a standardized suite of FHIR Genomics Operations that encapsulates the complexity of molecular data so that precision medicine solution developers can focus on building applications.

Materials and Methods: FHIR Genomics Operations essentially "wrap" a genomics data repository, presenting a uniform interface to applications. More importantly, operations encapsulate the complexity of data within a repository and normalize redundant data representations—particularly relevant in genomics, where a tremendous amount of raw data exists in often-complex non-FHIR formats.

Results: Fifteen FHIR Genomics Operations have been developed, designed to support a wide range of clinical scenarios, such as variant discovery; clinical trial matching; hereditary condition and pharmacogenomic screening; and variant reanalysis. Operations are being matured through the HL7 balloting process, connectathons, pilots, and the HL7 FHIR Accelerator program.

Discussion: Next-generation sequencing can identify thousands to millions of variants, whose clinical significance can change over time as our knowledge evolves. To manage such a large volume of dynamic and complex data, new models of genomics-EHR integration are needed. Qualitative observations to date suggest that

freeing application developers from the need to understand the nuances of genomic data, and instead base applications on standardized APIs can not only accelerate integration but also dramatically expand the applications of Omic data in driving precision care at scale for all.

Key words: clinical genomics, application programming interface, electronic health record, clinical decision support, HL7 FHIR

OBJECTIVE

"It is more important to know what sort of person has a disease than to know what sort of disease a person has."

Hippocrates

As we strive to practice medicine with a precision that Hippocrates never imagined, enabling clinicians to glean insights and formulate individualized clinical management strategies from the sea of ever-expanding molecular data and knowledge remains a daunting task, but one of fundamental importance. Solutions for integrating actionable genomic data into the electronic health record (EHR) are only now emerging. Soon, routine genomics integration will mean far more than pushing a static PDF report into the EHR, and we can expect to see real-time clinical decision support (CDS) with rules dynamically updated to reflect current knowledge. A barrier has been the development effort required to overcome the hurdle of bioinformatic calculations. Here, we summarize efforts towards a new paradigm in genomics-EHR integration, via a standardized suite of genomics operations (aka genomics APIs). This approach holds promise to obviate the need for developers to be experts in bioinformatic calculations and to dramatically simplify the required application development effort to bring concise contextually relevant genomic findings and recommendations to clinicians at the point of care.

BACKGROUND AND SIGNIFICANCE

Genomics-EHR integration challenges

Patients' genomic testing results are commonly integrated into their EHR through custom PDF reports designed by the testing laboratory. These reports include select genomic test findings, interpretation knowledge, and patient care recommendations reflecting the lab's most current internal curation process. While state-of-the-art, these textual reports are neither ideal for clinicians nor for CDS[1,2]—they contain only a slice of key variants and a point-in-time snapshot of interpretations; they are difficult and time-consuming to review; clinicians are not assisted in evaluating relevant interactions mentioned in the reports when making decisions; and they do not provide structured data needed for CDS guidance or analytics. To address these limitations, EHR vendors are enhancing their products in anticipation of incorporating structured genomic findings,[3] and HL7 Version 2 messaging and HL7 FHIR Genomics reporting standards are maturing.[4,5] Large research projects such as eMERGE[6] and CSER[7] are exploring the use of FHIR Genomics, and HL7 FHIR is gaining wide traction, as are apps based on the SMART-on-FHIR platform for broader clinical applications.[8–10] Adoption of the FHIR Genomics implementation guide provides a playbook for how to provide genomic data computably, helping to solve some of the issues.

But even with HL7 FHIR Genomics, EHR integration of genomic data remains challenging for many reasons:

- Data volume: Genomic data are voluminous and complex.[1,2,11,12] Next-generation sequencing (NGS) can identify thousands to millions of variants; significant bioinformatics knowledge and computational resources can be required to make sense of the data;
- Limited interoperability: Raw sequencing data formats and EHR data standards differ, with only a limited number of clinical informaticists having a detailed understanding of bioinformatics data specifications, which can lead to costly and resource intensive efforts to master bioinformatic calculations to normalize data. Furthermore, genomics data, even when encoded in HL7 FHIR format, pose variability in variant representation that creates a barrier to effective search and aggregation, for such activities as matching a patient's genetic profile with knowledge bases.[13] The variability leads to additional normalization challenges to be overcome during development of applications that would use genomic data to inform clinical, operational or research efforts;
- Rapid evolution of knowledge: The implications of genomic data are rapidly changing as our knowledge matures, resulting in these commonly utilized PDF reports quickly becoming outdated. Analysis of data from the NIH ClinVar genomic data repository shows that from 2016 to 2018, the number of known pathogenic or likely pathogenic variants in 66 genes of high clinical relevance[14] went from 10 137 to 18 718[15,16] reflecting over eight thousand new or recategorized clinically significant variants over the course of merely 2 years. The amount of variation and complexity considered in clinical interpretation of test results will only increase, especially with the emergence of polygenic risk scores.[17]

As will be described, FHIR Genomics Operations are designed to overcome these challenges.

Why FHIR?

HL7 Fast Healthcare Interoperability Resources (FHIR) is a next-generation interoperability standard designed to enable health data, including clinical and administrative data, to be quickly and efficiently exchanged. Based on common World Wide Web technologies and core application programming interface (API) capabilities, coupled with base semantic resources that enable easy exchange of conditions, medications, laboratory observations, and more, HL7 FHIR has gained rapid acceptance on a global scale as an innovative standard for enabling health data interoperability.

What is more, and particularly relevant in the context of the work described in this report, FHIR describes a mechanism for extending basic FHIR query capabilities through the creation of "Operations." FHIR Operations are a standardized way to extend the RESTful FHIR API's Create/Read/Update/Delete actions and enable use cases where servers play an active role in formulating responses, where the intended purpose is to cause side effects such as the creation of new resources, and for data normalization to abstract away from variability in data representation. Many FHIR specifications supplement the standardization of data structures with the addition of FHIR Operations that define advanced API capabilities.

The HL7 FHIR Genomics standard[5] defines FHIR representations for a range of genomic data structures (eg, variants, haplotypes, and variant implications), enabling a standards-based communication of simple and structural variants, germline and somatic variants, pharmacogenomic star alleles, HLA typing, and other findings generated from sequencing, chip technology, cytogenetic analysis, along with variant annotations and interpretations. The work described in this report extends the HL7 FHIR Genomics reporting standard with a suite of FHIR Genomics Operations. As will be described, these Operations extend the FHIR Genomics standard and basic FHIR search capabilities in order to simplify developer access to potentially complex and voluminous data structures.

Why operations?

To manage such a large volume of dynamic and complex results, many institutions are exploring the storage of genomic data outside or alongside the EHR, using a genomic data server, also referred to as a Genomic Archiving and Communication System (GACS).[18,19] A GACS stores sequence data generated from a sequencing laboratory and is analogous in many ways to a picture archiving and communication system (PACS), which stores image files that are not suitable to store directly in an EHR. The US Office of the National Coordinator's (ONC) Sync for Genes project emphasizes the need for pilots that test the use of FHIR for GACS integration with EHRs.[20] Just as the Digital Imaging and Communications in Medicine (DICOM) standard enabled the widespread adoption of PACS, there is a need for a standard set of APIs by which clinical applications can access and parse rich and complex genomic data.[21,22] FHIR-based APIs are being developed and promoted by leading federal agencies such as ONC,[23] NIH,[24] and the NHGRI.[25] However to date there are no FHIR-based APIs specifically designed to meet the needs of clinical genomics applications. Alterovitz et al[22] found that the use of APIs based on FHIR "allows for simple implementation, small payload sizes, and intuitive nonduplicative retrieval of data". Swaminathan et al[26] found that FHIR Genomics APIs "hold the promise to make building of complicated genomics applications easier with downstream constructive effects on healthcare." FHIR Genomics Operations meet that promise by reducing the overhead and necessity of normalizing variability in variant representation by API users (such as application developers of both native EHR or non-native EHR CDS) to reduce both runtime and developmental cost. Today's leading academic healthcare systems looking to incorporate genomics CDS into their native EHR environments face many barriers.[27] Standardizing and commoditizing genomic operations are a foundation for equity in precision medicine.

MATERIALS AND METHODS

FHIR genomics operations

FHIR Genomics Operations[28] (Figure 1) are based on the premise that genomic data, in FHIR format and/or some other format (eg, VCF format), are stored in a repository, either in or alongside an EHR, possibly along with phenotype annotations. The FHIR Genomics Operations essentially "wrap" the repository, presenting a uniform interface to applications, regardless of internal repository data structures. More importantly, FHIR operations can encapsulate the complexity of data within a repository and normalize redundant data representations. This is particularly relevant in genomics, where a tremendous amount of raw data exists in often-complex non-FHIR formats, and where variability in variant representation is a known barrier to matching a patient's genetic profile against knowledge bases—for clinical trials matching, for prediction of genetic predisposition to disease, for precision medication selection, genetic risk calculation.

A common use case driving FHIR Genomics Operations is the notion of an application (eg, a SMART-on-FHIR clinical genomics app, a CDS application, an EHR screen) needing specific genotype or phenotype information for a patient or a population in a form or format that does not already exist on the server. Applications have diverse needs, such as matching a cancer patient to available clinical trials based on identified somatic variants; screening for actionable hereditary conditions; identifying a risk for adverse medication reactions based on pharmacogenomic (PGx) variants; optimizing the dose of a medication; and updating a patient's risk as knowledge of their variants evolves. A goal for FHIR Genomics Operations is to ultimately support any and all of these clinical scenarios. Currently, these require an ability to bioinformatically calculate from among a myriad of variant representations, both in data stores and in knowledge sources. Genomic operations that normalize the data store reduce the complexity by giving a single representation of variation data to work with, freeing industry and academic solution developers to focus on building high-performance, user-friendly, scalable genomic CDS solutions. In the future, additional standard operations can be added supporting genomics research, variant calling and annotation, and knowledge base lookups.

We categorize FHIR Genomics Operations along two orthogonal axes (Table 1)—subject versus population, and genotype versus phenotype. For example, the "find-subject-variants" operation is categorized as a "subject" and a "genotype" operation, that retrieves genotype information for a single subject; whereas the "find-population-tx-implications" is categorized as a "population" and a "phenotype" operation, that retrieves a count or list of patients having specific phenotypes (such as being intermediate metabolizers of clopidogrel). The metadata operation retrieves metadata about the genomic studies that generated the data.

The current suite of FHIR Genomics Operations is shown in Table 2. Fifteen FHIR Genomics Operations are designed to support a wide range of clinical scenarios, such as simple and structural variant discovery; clinical trial matching; screening for actionable hereditary conditions; variant prioritization and filtering; identifying a risk for adverse medication reactions based on PGx variants; and updating a patient's risk as knowledge of their variants evolves.

General characteristics of all FHIR Genomics Operations include genomic region-based searching and an expectation of server-side normalization and genomic liftover. These can be extremely costly (in runtime and development of logic) to calculate due to the multiplicity of variation representation. Adding further to the burden, as our understanding of genomics grows, we continue to recognize the clinical significance of variations not only within protein-coding gene exons, but also in introns and intergenic regions. To enable a consistent approach to query, whether for variants within or outside a gene, certain operations provide a "ranges" query parameter (eg, to query for variants in the APC gene, one can call the "find-subject-variants" operation with a parameter of ranges="NC_000005.10:112707497-112846239").

With regards to variation representation, variants can be represented in different ways, and a "variation" can be relative to a number of different references. For instance, these representations synonymously refer to the same variant:

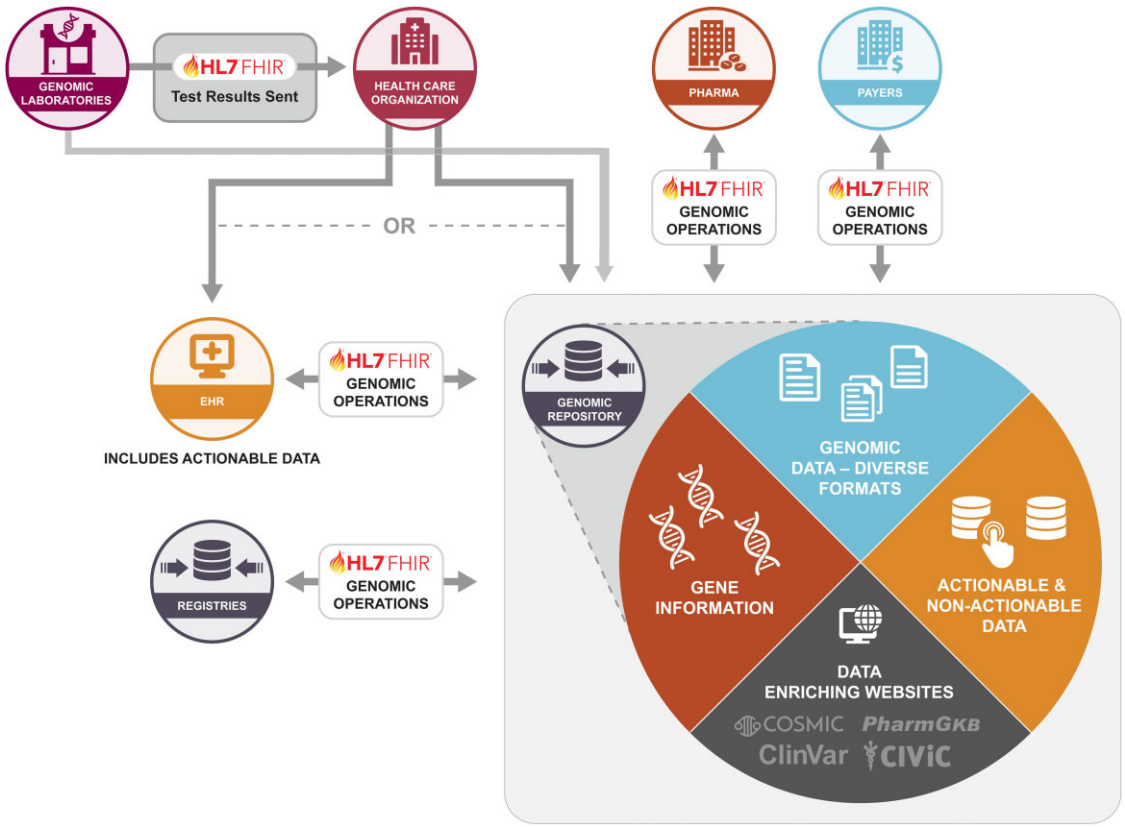

Figure 1. FHIR Genomics Operations. Operations essentially "wrap" a genomic data repository, presenting a uniform interface to applications. (Image courtesy of HL7 FHIR Accelerator CodeX™.)

Table 1. Scope of FHIR Genomics Operations

	Subject operations	Population operations
Genotype operations	Retrieve simple/structural variants, haplotypes • find-subject-variants • find-subject-specific-variants • find-subject-structural-intersecting-variants • find-subject-structural-subsuming-variants • find-subject-haplotypes • find-subject-specific-haplotypes	Identify cohorts with simple/structural variants, haplotypes • find-population-specific-variants • find-population-structural-intersecting-variants • find-population-structural-subsuming-variants • find-population-specific-haplotypes
Phenotype operations	Retrieve diagnostic and therapeutic implications • find-subject-tx-implications • find-subject-dx-implications	Identify cohorts with diagnostic and therapeutic implications • find-population-tx-implications • find-population-dx-implications
Metadata operations	Retrieve sequencing study metadata • find-study-metadata	

- NM_001195798.2: c.12G>A
- NM_001195803.2: c.12G>A
- NC_000019.9: g.11200236G>A
- NG_009060.1: g.5180G>A
- NC_000019.10:11089559: G: A

One requirement of the FHIR Genomics Operations is that one can query using any of these representations, and variants matching any of its synonyms will be returned regardless of how they are formatted/represented/stored in a server. This achieves the goal of simplifying the data access efforts of the application developers, data analysts, business analysts, or researchers. By adopting the standard FHIR Genomics Operations, a system achieves an accessible, reliable, consistent representation of their institution's genomic data.

In addition to variant data access and interoperability, phenotype operations return diagnostic implications (eg, a patient's variant is known to be associated with a hereditary condition) and therapeutic implications (eg, a patient's genotype is known to affect the metabolism of a particular drug) of a subject or a population's variants. The inclusion of phenotype operations within the suite of FHIR Genomics Operations enables the development of solutions for a variety of use cases such as ACMG screening,[14] PGx screening, clinical trial matching, risk scoring, clinical management guidance, and more. Operations can return previously instantiated implications (eg, those that came in via a static lab report) and/or dynamically computed implications (eg, those computed on the fly using an associated knowledge base). Decoupling the update of knowledge

Table 2. Suite of FHIR Genomics Operations

Operation	Description
find-subject-variants	Determine if simple variants are present that overlap range(s).
find-subject-specific-variants	Determine if specified simple variants are present.
find-subject-structural-intersecting-variants	Determine if structural variants are present that overlap range(s).
find-subject-structural-subsuming-variants	Determine if structural variants are present that fully subsume a range.
find-subject-haplotypes	Retrieve haplotypes/genotypes for specified genes.
find-subject-specific-haplotypes	See if specified haplotypes/genotypes are present.
find-subject-tx-implications	Retrieves genetic therapeutic implications for variants/haplotypes/genotypes.
find-subject-dx-implications	Retrieves genetic diagnostic implications for variants.
find-population-specific-variants	Retrieve count or list of patients having specified variants.
find-population-structural-intersecting-variants	Retrieve count or list of patients having structural intersecting variants in specified regions.
find-population-structural-subsuming-variants	Retrieve count or list of patients having structural subsuming variants in specified regions.
find-population-specific-haplotypes	Retrieve count or list of patients having specified genotypes/haplotypes.
find-population-tx-implications	Retrieve count or list of patients having therapeutic implications.
find-population-dx-implications	Retrieve count or list of patients having diagnostic implications.
find-study-metadata	Retrieve metadata about sequencing studies performed on a subject.

bases from the reporting of sequencing results enables the most current annotations and CDS guidance to be provided to clinicians. In addition to freeing an API data recipient from the cost of developing and performing normalizing bioinformatic calculations, adoption of the phenotype operations can simplify access to clinically curated knowledge sources (whether the curation is local, based on laboratory reports, or consumed from knowledge sources such as ClinVar, PharmGKB, or CIViC directly or via CDS services intermediaries). A standard approach in the phenotype operations provides an unprecedented opportunity to streamline the interaction between local clinical curation of variant knowledge, static lab implications, and extension by additional knowledge bases, allowing one to keep pace as knowledge evolves. For instance, if one wants to deliver the lab's impression of the effect of a variant on a PGx interaction, coupled with local PGx pharmacy specialist interpretations, then the "find-subject-tx-implications" phenotype operation is the source from which it is possible to consume that amalgamation. And, given its nature as a standard, one can expect that programmed-solutions implemented at one hospital can more seamlessly be reused at another, or that intermediary cloud-based genomic CDS vendors can provide phenotype operations as a SaaS[29] solution. The phenotype operations are built on the variation (genotype) focused operations. Together they provide a platform for CDS and analytics.

Reference implementation

To make it easier to capture the value of Genomic Operations, we have produced a full open-sourced reference implementation[30] of the FHIR Genomics Operations using synthetic or publicly available and anonymized data.

The reference implementation user interface[31] is based on Swagger, and allows anyone to visualize and interact with the APIs using only a web browser. All source code is available on the reference implementation's github site.[30] Detailed documentation is provided on the site's wiki. In particular, the "Getting Started" page describes patient data and knowledge data used to drive this implementation of the Operations. You can experiment with predefined queries (see postman collection) or create your own queries based on available data; the "Examples and Exercises" page provides scenarios that demonstrate various capabilities of each Operation. Exercises are provided for those that want to test their understanding; the "Software Utilities" page describes software utilities used primarily to help load data into the database that underlies the reference implementation (eg, vcf2json uses vcf2fhir[32] logic to import VCF data) and to support fast normalization (eg, by replicating portions of NCBI SPDI variation services[13]); the "Genomics Apps" page provides some simple genomics applications built using the FHIR Genomics Operations; the "Replicating the Reference Implementation" page describes how to fully replicate the reference implementation (code and data). The reference implementation is written in python and deployed on Heroku,[33] with data stored in MongoDB.[34]

RESULTS

Illustrative examples

A few examples (drawn from the "Examples and Exercises" page of the reference implementation described above) are provided here to further illustrate the use of the operations.

We are interested in patient HG00403's LDLR variants. The location of the LDLR gene on the GRCh37 reference is NC_000019.9:11200138-11244496, whereas the location on the GRCh38 reference is NC_000019.10:11089431-11133820. Given that operations encapsulate server-side genomic liftover, these two calls will return the exact same results:

```
https://fhir-gen-ops.herokuapp.com/subject-operations/genotype-operations/$find-subject-variants?subject=HG00403&ranges=NC_000019.9:11200138-11244496&includeVariants=true
```

```
https://fhir-gen-ops.herokuapp.com/subject-operations/genotype-operations/$find-subject-variants?subject=HG00403&ranges=NC_000019.10:11089431-11133820&includeVariants=true
```

We are interested to know if patient m123 has a particular APC variant. The lab that reported the variant only reports transcript level variants, not necessarily based on the MANE transcript.[35] Given that operations encapsulate server-side normalization, these three calls will return the exact same results:

```
https://fhir-gen-ops.herokuapp.com/subject-operations/genotype-operations/$find-subject-specific-variants?subject=m123&variants=NM_001127510.3:c.145A>T
```

https://fhir-gen-ops.herokuapp.com/subject-operations/genotype-operations/$find-subject-specific-variants?subject=m123&variants=NC_000005.10:112766334:A:T

https://fhir-gen-ops.herokuapp.com/subject-operations/genotype-operations/$find-subject-specific-variants?subject=m123&variants=NC_000005.9:g.112102032A>T

Patient CA12345 has metastatic non-small cell lung cancer. Biopsy shows two somatic variants felt to be oncogenic: NM_002524.5:c.182A>C (NRAS: p.Gln61Pro), NM_001354609.2:c.1799T>A (BRAF:p.V600E). Clinician now wants to determine if there are any molecularly guided medication treatment options for this patient. (Results show predicted resistance to dabrafenib, sensitivity to vemurafenib, and sensitivity to dabrafenib+trametinib.)

https://fhir-gen-ops.herokuapp.com/subject-operations/phenotype-operations/$find-subject-tx-implications?subject=CA12345&variants=NM_002524.5:c.182A>C, NM_001354609.2: c.1799T>A&conditions=https://disease-ontology.org|3908

Patient HG02657 has liver disease, and the patient's clinician suspects hemochromatosis. The clinician wants to see if patient HG02657 has any variants associated with hereditary hemochromatosis. (Results show presence of variant NC_000006.11:26091178:C:G, pathogenic for hemochromatosis type 1.)

https://fhir-gen-ops.herokuapp.com/subject-operations/phenotype-operations/$find-subject-dx-implications?subject=HG02657&conditions=https://www.ncbi.nlm.nih.gov/medgen|C3469186

Operations maturation process

The FHIR Maturity Model[36] can be used by implementers to assess the maturity level, and therefore the stability, of an artifact. Maturity levels advance from 0 through 5. FHIR Genomics Operations are currently at FHIR Maturity Level 0—they are draft, and will be subject to formal HL7 balloting in early 2023. We are targeting a FHIR Maturity Level 2 by end of 2023, meaning that the operations will have been balloted and tested through at least three pilots. Operations are maturing through connectathons, pilots, and the HL7 FHIR Accelerator program. In addition, three GA4GH-sponsored Google Summer of Coding projects[37,38] have allowed us to benefit from student (Rohan Gupta, Justin Aronson) developer contributions. We are not aware of any live production applications at this time. The most up to date information on the Operations can be found on the HL7 FHIR site.[5]

HL7 FHIR Connectathons[39] are collaborative hands-on FHIR integration testing events held triannually. FHIR Genomics Operations were first tested in 2018, as part of a PGx CDS scenario, and were most recently tested in May 2022 under a broad range of scenarios.[40]

The HL7 FHIR Accelerator program[41] is designed to assist communities and collaborative groups across the global health care spectrum in the creation and adoption of high quality FHIR Implementation Guides or other standard artifacts to move toward the realization of global health data interoperability. The CodeX Accelerator is focused on driving advances in cancer care.[42,43] Under the CodeX umbrella, the GenomeX[44] project is driving advances in molecular/precision medicine (not limited to cancer), and of particular relevance here, the GenomeX—Genomics Operations[45] Use Case will synthesize feedback from pilots, and will channel that feedback to the HL7 Clinical Genomics committee for iterative enhancements to the operations (Figure 2).

Pilot projects are a rich source of feedback. We generally think of two main types of pilots—those that implement operations against a genomic data repository, and those that use operations in an application (eg, an app or a CDS pipeline). Early pilots, such as the Face Sheet app (described here[32]) and the PGx CDS pipeline[46] demonstrated the feasibility of querying a genomic data store in real time. The reference implementation[30] was itself a sizable pilot, providing valuable insights into optimal operation design and normalization strategies. Implementation of the operations against an evolving genomics data model was performed by Flatiron Health, with the goal of testing the operations and evaluating their utility. This testing and discussion led to a deeper understanding of requirements including the scenario where patients are sequenced more than once. Under ONC's Sync for Genes Phase 5, our team is exploring a broad set of scenarios and applications.[47] A full description of Sync for Genes, Phase 5 will be the subject of a future manuscript, but scenarios range from a relatively simple return of genetic screening results to a more complex return of richly annotated variants to drive a filtering/prioritization workflow such as described in reference 48. An early finding from the Sync for Genes work is that computational annotation is more accurate when phase data are considered—so the operations were updated to also return phase data. This nicely illustrates how real world pilots lead to ongoing enhancements in FHIR Genomics Operations over time. Our team is currently working on a pilot for the use of Operations in a clinical trial matching workflow[49] where we have found that Operations, by returning canonical variants in a predictable format, greatly simplify CDS rule development; and as part of a National Human Genome Research Institute sponsored grant[50] looking to develop a comprehensive, medication decision support service (PillHarmonics™) enriched with PGx guidance, where we plan to assess the adequacy of phenotype operations responses. Additional pilots being considered under the GenomeX umbrella include implementing the Operations against different physical data storage paradigms (eg, an OMOP relational model,[51] a TileDB sparse array[52]), and testing the Operations in a variety of clinical genomics apps (eg, germline ACMG screening, tumor board dashboard, and PGx screening).

DISCUSSION AND CONCLUSION

NGS can identify thousands to millions of variants, whose clinical significance can change over time as our knowledge evolves. To manage such a large volume of (dynamic and complex) results, new models of genomics-EHR integration are needed.

Here we describe a genomics-EHR integration strategy based on a standard set of FHIR Genomics Operations that essentially wrap a genomics data repository or GACS server, presenting a uniform interface to applications, regardless of internal repository data structures. Housing genomic data in a separate genomic data server, wrapped by a set of FHIR APIs, in communication with the EHR and/or an intervening CDS engine, offer exciting possibilities for

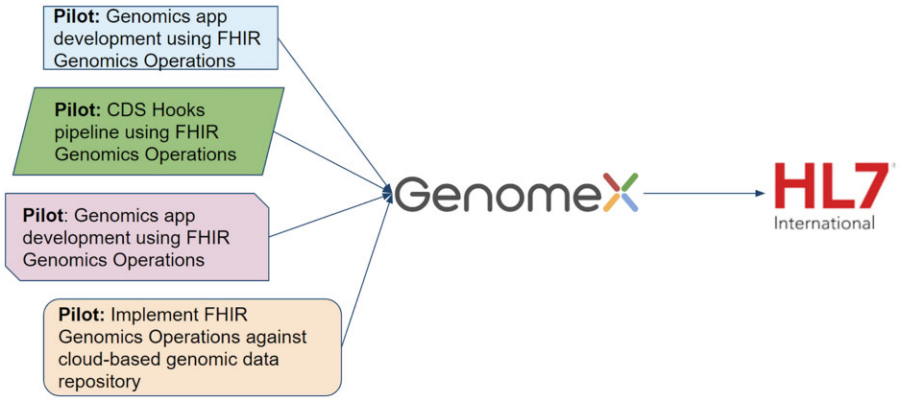

Figure 2. Cascading pilot feedback through GenomeX to enhance genomics operations. GenomeX FHIR Accelerator project will synthesize feedback from pilots, and will channel that feedback to the HL7 Clinical Genomics committee for iterative enhancements to the operations.

managing a person's entire genome, managing evolution in our understanding of a person's genome, and for provision of contextually relevant genomics findings and recommendations at the point of care. Experiments to date, through connectathons, FHIR Accelerator projects, and pilots, are providing a rich array of use cases against which the operations are being enhanced. There are many additional use cases (eg, molecular tumor board app) yet to be explored, many genomic concepts yet to be formally modeled in FHIR (eg, repeat expansions, gene fusions), and ambiguities in the base FHIR Genomics specification (eg, choice of phenotype code system) yet to be resolved. That said, qualitative observations to date suggest that freeing application developers from the need to understand the nuances of genomic data, and instead base applications on standardized APIs can not only accelerate integration but also dramatically expand the applications of Omic data in driving precision care at scale for all.

FUNDING

This work received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

AUTHOR CONTRIBUTIONS

RHD: Primary design of FHIR Genomics Operations and reference implementation; primary manuscript author. BSEH: Senior advisor on Operations design and implementation; substantial contribution to conception and design of Operations; contributed to manuscript. GA, AB, JS, TH: Advisor on Operations design and implementation; contributed to manuscript. RG, JA: Primary software developer; contributed to manuscript. SRG: Primary software oversight; contributed to manuscript. ArthurH: Championed the GenomeX Operations accelerator; contributed to manuscript. AmmarH: Senior advisor on Operations design and implementation; contributed to manuscript. FNR, BR, CR, DH, JJ: Contributed to conception and pilots; contributed to manuscript. MT: Co-championed the GenomeX Operations accelerator; contributed to Operations conception, design, and usability; contributed to manuscript. NX, PZ: Contributed to conception and pilots; contributed to manuscript. SC: Senior author; senior advisor on Operations design and implementation; substantial contribution to conception and design of Operations; contributed to manuscript.

CONFLICT OF INTEREST STATEMENT

None declared.

ACKNOWLEDGMENTS

What if you build a standard, and no one comes (to use it)? We are so fortunate to have not had that problem. Introduction of a new paradigm in genomics-EHR integration simply cannot happen without a collection of visionaries and those willing to study and pilot and improve upon an evolving set of specifications. While the authors would not list all of the folks involved in testing the operations, they would like to give special recognition to several groups. First, the authors gratefully acknowledge the HL7 Clinical Genomics committee, diligently and deliberately focused over the past several years on defining FHIR-based representations of genomic concepts. It would be impossible for an operation to return standards-based FHIR data where there are no FHIR Genomics standards to leverage. Second, they gratefully acknowledge the GA4GH Genomic Knowledge Standards group. Collaboration with this group led to, among others, significant enhancements in the FHIR-based representation of diagnostic and therapeutic implications, greatly enhancing the ability of operations to dynamically leverage ClinVar and CIViC knowledge. Third, the authors gratefully acknowledge Leap of Faith Technologies for their generous support of this work. Fourth, the authors would like to thank Flatiron Health, an independent subsidiary of Roche, and their employees Kim Peifer, and Shannon Lee for participating in an HL7 Connectathon (May 3–4, 2022, virtual) wherein testing of the evolving FHIR operations on the reference implementation using sample data was performed and assessed.

DATA AVAILABILITY

No original research data were generated in the course of the work described in this report. HL7 FHIR Genomics Operations are a public standard maintained by HL7. The most current Operations definitions are here (http://build.fhir.org/ig/HL7/genomics-reporting/operations.html). An open-source reference implementation of the HL7 FHIR Genomics Operations is here (https://github.com/FHIR/genomics-operations).

REFERENCES

1. Walton NA, Johnson DK, Person TN, *et al.* Genomic data in the electronic health record. *Adv Mol Pathol* 2019; 2 (1): 21–33.

2. Williams MS, Taylor CO, Walton NA, *et al.* Genomic information for clinicians in the electronic health record: lessons learned from the Clinical Genome Resource Project and the Electronic Medical Records and Genomics Network. *Front Genet* 2019; 10: 1059. doi:10.3389/fgene.2019.01059
3. Sutton J, Wix K, Amusan A. Experiences and lessons learned from a new functionality for pharmacogenomics CDS in the electronic health record. https://zerista.s3.amazonaws.com/item_files/1663/attachments/427873/original/amia_summit_2018_poster.pdf. Accessed November 5, 2020.
4. HL7 Standards Product Brief – HL7 Version 2.5.1 Implementation Guide: Laboratory Results Interface, Release 1 STU Release 3 – US Realm | HL7 International. https://www.hl7.org/implement/standards/product_brief.cfm?product_id=279. Accessed August 19, 2021.
5. HL7.FHIR.UV.GENOMICS-REPORTING\Home Page – FHIR v4.0.1. http://build.fhir.org/ig/HL7/genomics-reporting/index.html. Accessed July 18, 2022.
6. eMERGE Consortium. Harmonizing Clinical Sequencing and Interpretation for the eMERGE III Network. *Am J Hum Genet* 2019; 105: 588–605.
7. Wynn J, Lewis K, Amendola LM, *et al.* Clinical providers' experiences with returning results from genomic sequencing: an interview study. *BMC Med Genomics* 2018; 11 (1): 45.
8. Alterovitz G, Warner J, Zhang P, *et al.* SMART on FHIR Genomics: facilitating standardized clinico-genomic apps. *J Am Med Inform Assoc* 2015; 22 (6): 1173–8.
9. Warner JL, Rioth MJ, Mandl KD, *et al.* SMART precision cancer medicine: a FHIR-based app to provide genomic information at the point of care. *J Am Med Inform Assoc* 2016; 23 (4): 701–10.
10. 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.
11. Ohno-Machado L, Kim J, Gabriel RA, *et al.* Genomics and electronic health record systems. *Hum Mol Genet* 2018; 27 (R1): R48–R55.
12. Caraballo PJ, Bielinski SJ, St Sauver JL, *et al.* Electronic medical record-integrated pharmacogenomics and related clinical decision support concepts. *Clin Pharmacol Ther* 2017; 102 (2): 254–64.
13. Holmes JB, Moyer E, Phan L, *et al.* SPDI: data model for variants and applications at NCBI. *Bioinformatics* 2020; 36 (6): 1902–7.
14. Kalia SS, Adelman K, Bale SJ, *et al.* Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med* 2017; 19 (2): 249–55.
15. NCBI ClinVar gene-specific summary, 2016. NCBI ftp site. Accessed December 3, 2020.
16. NCBI ClinVar gene-specific summary, 2018. NCBI ftp site. Accessed December 3, 2020.
17. Topol E. Human genomics vs clinical genomics. Ground Truths. 2022. https://erictopol.substack.com/p/human-genomics-vs-clinical-genomics. Accessed September 23, 2022.
18. Starren J, Williams MS, Bottinger EP. Crossing the omic chasm: a time for omic ancillary systems. *JAMA* 2013; 309 (12): 1237–8.
19. Masys DR, Jarvik GP, Abernethy NF, *et al.* Technical desiderata for the integration of genomic data into electronic health records. *J Biomed Inform* 2012; 45 (3): 419–22.
20. Alterovitz G. Enabling clinical genomics for precision medicine via HL7 FHIR [published online November 2017]. *ONC Sync Genes Rep 20171–39.* https://www.healthit.gov/sites/default/files/sync_for_genes_report_november_2017.pdf. Accessed November 5, 2020.
21. Conway JR, Warner JL, Rubinstein WS, Miller RS. Next-generation sequencing and the clinical oncology workflow: data challenges, proposed solutions, and a call to action. *JCO Precis Oncol* 2019; 3: PO.19.00232. doi: 10.1200/PO.19.00232.
22. Alterovitz G, Heale B, Jones J, *et al.* FHIR Genomics: enabling standardization for precision medicine use cases. *Npj Genomic Med* 2020; 5: 1–4.
23. Standardized API for patient and population services. HealthIT.gov. https://www.healthit.gov/test-method/standardized-api-patient-and-population-services. Accessed August 5, 2022.
24. NOT-OD-19-122: Fast Healthcare Interoperability Resources (FHIR) Standard. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-19-122.html. Accessed August 5, 2022.
25. Schatz MC, Philippakis AA, Afgan E, *et al.* Inverting the model of genomics data sharing with the NHGRI genomic data science analysis, visualization, and informatics lab-space. *Cell Genomics* 2022; 2 (1): 100085.
26. Swaminathan R, Huang Y, Moosavinasab S, *et al.* A review on genomics APIs. *Comput Struct Biotechnol J* 2016; 14: 8–15.
27. Wake DT, Smith DM, Kazi S, *et al.* Pharmacogenomic clinical decision support: a review, how-to guide, and future vision. *Clin Pharma Therap* 2022; 112 (1): 44–57.
28. HL7.FHIR.UV.GENOMICS-REPORTING\Operations – FHIR v4.0.1. http://build.fhir.org/ig/HL7/genomics-reporting/operations.html. Accessed July 18, 2022.
29. Software as a service. Wikipedia. 2022. https://en.wikipedia.org/w/index.php?title=Software_as_a_service&oldid=1110254354. Accessed September 22, 2022.
30. HL7 FHIR Genomics Operations – Reference Implementation. 2022. https://github.com/FHIR/genomics-operations. Accessed July 18, 2022.
31. FHIR Genomics Operations – Reference Implementation (User Interface). https://fhir-gen-ops.herokuapp.com/. Accessed July 18, 2022.
32. Dolin RH, Gothi SR, Boxwala A, *et al.* vcf2fhir: a utility to convert VCF files into HL7 FHIR format for genomics–EHR integration. *BMC Bioinform* 2021; 22 (1): 104.
33. Cloud Application Platform. Heroku. https://www.heroku.com/. Accessed August 5, 2022.
34. MongoDB: The Developer Data Platform. MongoDB. https://www.mongodb.com. Accessed August 5, 2022.
35. Morales J, Pujar S, Loveland JE, *et al.* A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. *Nature* 2022; 604 (7905): 310–5.
36. FHIR Maturity Model. https://build.fhir.org/versions.html#maturity. Accessed July 19, 2022.
37. Google Summer of Code 2021. https://summerofcode.withgoogle.com/archive/2021/organizations/4716668999172096. Accessed September 26, 2022.
38. Google Summer of Code 2022. https://summerofcode.withgoogle.com/programs/2022/organizations/global-alliance-for-genomics-and-health. Accessed September 26, 2022.
39. Connectathons – FHIR – Confluence. https://confluence.hl7.org/display/FHIR/Connectathons. Accessed August 5, 2022.
40. 2022-05 Clinical Genomics: retrieving FHIR genetic data with operations – FHIR – confluence. https://confluence.hl7.org/display/FHIR/2022-05+Clinical+Genomics%3A+Retrieving+FHIR+Genetic+data+with+Operations. Accessed July 18, 2022.
41. HL7 FHIR ACCELERATOR™ Program | HL7 International. https://www.hl7.org/about/fhir-accelerator/. Accessed July 19, 2022.
42. Goel AK, Campbell WS, Moldwin R. Structured data capture for oncology. *JCO Clin Cancer Inform* 2021; 5: 194–201.
43. Osterman TJ, Terry M, Miller RS. Improving cancer data interoperability: the promise of the Minimal Common Oncology Data Elements (mCODE) Initiative. *JCO Clin Cancer Inform* 2020; 4: 993–1001.
44. Genomics – CodeX – Confluence. https://confluence.hl7.org/display/COD/Genomics. Accessed July 18, 2022.
45. GenomeX – Genomics Operations – CodeX – Confluence. https://confluence.hl7.org/display/COD/GenomeX+-+Genomics+Operations. Accessed July 18, 2022.
46. Dolin RH, Boxwala A, Shalaby J. A pharmacogenomics clinical decision support service based on FHIR and CDS hooks. *Methods Inf Med* 2018; 57 (S 02): e115–23.
47. Garcia SJ, Zayas-Cabán T, Freimuth RR. Sync for genes: making clinical genomics available for precision medicine at the point-of-care. *Appl Clin Inform* 2020; 11 (2): 295–302.

48. Austin-Tse CA, Jobanputra V, Perry DL, *et al*. Best practices for the interpretation and reporting of clinical whole genome sequencing. *NPJ Genomic Med* 2022; 7: 1–13.
49. A novel CDS strategy for molecularly-guided clinical trials matching. Elimu Inform. 2022. https://www.elimu.io/blog/a-novel-cds-strategy-for-molecularly-guided-clinical-trials-matching/. Accessed July 18, 2022.
50. PillHarmonics: an orchestrated medication clinical decision support service. NHGRI Grant 1R43HG011832-01A1. https://reporter.nih.gov/search/TI4lc-go4FUeum8BAsB-ftw/project-details/10382886. Accessed November 9, 2022.
51. OMOP Common Data Model – OHDSI. https://www.ohdsi.org/data-standardization/the-common-data-model/. Accessed September 7, 2022.
52. TileDB-VCF. 2022. https://github.com/TileDB-Inc/TileDB-VCF. Accessed September 7, 2022.