FHIR Genomics: Enabling standardization for precision medicine use cases

Gil Alterovitz
Bret Heale
James Jones
David Kreda
Fan Lin

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs
Authors
Gil Alterovitz, Bret Heale, James Jones, David Kreda, Fan Lin, Lei Liu, Xin Liu, Kenneth D. Mandl, David W. Poloway, Rachel Ramoni, Alex Wagner, and Jeremy L. Warner
FHIR Genomics: enabling standardization for precision medicine use cases

Gil Alterovitz1,2,3,4, Bret Heale2,3, James Jones1, David Kreda5, Fan Lin1,5, Lei Liu1, Xin Liu1, Kenneth D. Mandl1,2,6, David W. Poloway1, Rachel Ramoni7, Alex Wagner5 and Jeremy L. Warner6

The development of Fast Healthcare Interoperability Resources (FHIR) Genomics, a feasible and efficient method for exchanging complex clinical genomic data and interpretations, is described. FHIR Genomics is a subset of the emerging Health Level 7 FHIR standard and targets data from increasingly available technologies such as next-generation sequencing. Much care and integration of feedback have been taken to ease implementation, facilitate wide-scale interoperability, and enable modern app development toward a complete precision medicine standard. A new use case, the integration of the Variant Interpretation for Cancer Consortium (VICC) “meta-knowledgebase” into a third-party application, is described.

INTRODUCTION

Successful practice of precision medicine will depend upon knowledge-based interpretation of genomic variant data at the point of care, leading to drastically improved diagnosis, prognosis, and treatment selection.1–4 Variant data are identified both by traditional genetic panels and high-throughput sequencing, with many large genomic databases already housing non-compatible data structures, often employing codes based on different nomenclatures.5–7 It follows that an effective standard is needed to assimilate genomic results from various formats with other clinical data. Currently, most genetic test reports entered into the Electronic Health Record (EHR) are in PDF format.8 There is a great opportunity to expand functionality with the recently developed SMART platform utilizing the Fast Healthcare Interoperability Resources (FHIR) specification, enabling apps to be launched directly from within the EHR.

FHIR consists of multiple linkable and extendable data structure specifications called resources, modeling concepts in healthcare scenarios such as patients, conditions, and clinical observations and reports. These resources can be tailored to use cases by standardized sets of constraints and extensions called profiles. This approach has been praised both by large EHR vendors and major technology companies.9–12 Given this critical mass and building on the work of Alterovitz et al.,13 the authors and others in the Health Level Seven (HL7) community have enumerated many clinical genomics use cases in a domain analysis model, including clinical sequencing, cancer screening, pharmacogenomics, public health reporting, and decision support tools.14 These use cases and others required the introduction of additional specific data structures to expand the FHIR data model and prototype apps to showcase interoperability.

RESULTS

FHIR Genomics introduces

1. A new resource, MolecularSequence, for capturing non-interpretive (raw) sequencing data or pointing to it stored in an external repository as needed.

2. New profiles on the existing resources Observation, DiagnosticReport, ServiceRequest, Task, and FamilyMemberHistory to facilitate sharing genetic test results, including observed variants from reference sequences and their clinical implications and interpretations.

A full description of the FHIR Genomics artifacts and suggested usage can be found at https://www.hl7.org/fhir/genomics.html.

The genomics artifacts have been subject to ballot and are deemed ready for trial use implementation. International pilot evaluators from across the spectrum of stakeholders—including hospitals, universities, and vendors—have experimented with FHIR-based clinical genomics apps and use cases.13,15 Regular development and testing opportunities arise during international FHIR “Connectathon” events, held three times a year. These events create short feedback loops where research institutions, production EHR systems, and other developers join the community and test the specification for necessary interoperability and features. Further trial use and feedback will result in increasing maturity of the elements within FHIR as they move toward normative status, where adoption can occur without future iterative developments causing breaking changes.

A FHIR Genomics implementation guide has also received HL7 ballot feedback, leveraging standardized “components” within Observation, containing coordinated “codes” and “values.” Codifying individual concepts with this feature facilitates search operations and eases implementation. The guide leverages more profiles on top of the resources versioned in FHIR Release 4,
introducing a hierarchical structure of profiles and examples for many use cases.

**DISCUSSION**

During the domain analysis, it was observed that sequencing data may often need reanalysis, either when delegated to different organizations or determined necessary after updated interpretive information (e.g., on drugs and diseases) becomes available. Using FHIR for genomics data provides corresponding reusability; new resource instances can easily be linked to existing ones thanks to FHIR's JSON/XML-based architecture and RESTful application program interface (API). This allows for simple implementation, small payload sizes, and intuitive non-duplicative retrieval of data.

![Fig. 1 FHIR Genomics workflow use case. a A ServiceRequest instance; b a DiagnosticReport instance refers to Observations under "result"; c a MolecularSequence instance carries the reference sequence and variant information; d an Observation instance carries clinical interpretation; e additional Observations can carry further analysis information. Arrows depict the inter-resource pointers.](image)

Compared to other clinical data standards able to communicate genomic information, FHIR Genomics stands out for its inter-resource linkage capabilities and separation of interpretive and non-interpretive data. Before FHIR, HL7 v2 was the healthcare information exchange standard and was extended for genetics test reports with an implementation guide.17 However, criticism was made regarding the “Segment” structure—segments were not uniformly created and could lack unique identities, resulting in inefficiencies for data mining, downstream analysis, and complex interoperability. Concerns were also held against the rigid HL7 v3 standard,18 where one must largely implement the entire model to transmit data.

Apps can easily retrieve data using services such as SAMtools19 and export them into FHIR’s normative and widely adopted Observation resource using the defined genomic profiles. For communication with EHRs, clinicians, patients, and decision support engines, scaling this conversion (e.g., to all variants detected from whole-genome sequencing) may stress systems, so filtering for clinical actionability and/or translating on demand from a GACS implementation may be needed. Servers storing FHIR Genomics resources can interface with the research community.
through GA4GH APIs such as Beacon, though careful data security considerations must be made, as these servers may also contain additional sensitive clinical information.

FHIR Genomics has been shown to represent genomic data that suits the needs of current and upcoming clinical genomics use cases, and SMART technology continues to translate the potential of a single data standard into powerful precision medicine apps. As envisioned, the FHIR Genomics framework enables many value-added opportunities such as clinically integrated genomics knowledge-based apps and a translational bridge between research-oriented genomics data and precision medicine.

METHODS
To further probe the capabilities of the FHIR Genomics guidance, a prototype app was expanded to read and record genomic variants per updates to the FHIR R4 standard and implementation guide. In addition, cancer variants were linked to the Variant Interpretation for Cancer Consortium meta-knowledgebase API for context, with evidence classified by the AMP/ASCO/CAP somatic classification guidelines. A reference server was constructed using the Health Services Platform Consortium sandbox (https://sandbox.hspconsortium.org) where the app was successfully integrated with sample patient data. Additional supported use cases require storage and validation of information and terminologies cataloged by diverse efforts in the genomics research community, including ClinVar identifiers, HGVS nomenclature, HUGO Gene Nomenclature Committee identifiers, NCBI reference sequences, terms from the Sequence Ontology, clinical guidelines, and services, such as LOINC and SNOMED CT.

REPORTING SUMMARY
Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY
All relevant data are available from the authors without restriction. Guidance for the data structures needed for FHIR genomics can be found at http://hl7.org/fhir/R4/genomics.html and http://hl7.org/fhir/uv/genomics-reporting/index.html.

CODE AVAILABILITY
Code for the updated reference SMART-on-FHIR application is freely available on GitHub via https://github.com/smart-cancer-navigator/Application, with a public instance and instructions for linking to the public HSPC sandbox hosted at https://smart-cancer-navigator.github.io/app. Python code for a prototype reference server with sample FHIR genomics data is also available on GitHub at https://github.com/bcl-lab/FHIR-Genomics_v2.

Received: 12 February 2019; Accepted: 2 January 2020; Published online: 18 March 2020

REFERENCES

ACKNOWLEDGEMENTS
This work was funded in part by the SMART health IT project from the Office of the National Coordinator for Health Information Technology, the Precision Link Biobank at Boston Children’s Hospital, and U01 CA231840 from the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors would also like to thank the many members of the HL7 Clinical Genomics Work Group and its FHIR subgroup for their perspectives and contributions.

AUTHOR CONTRIBUTIONS
All authors contributed equally to the conceptualization, investigation, and methodology of the work, as well as draft preparation, reviewing, and editing. All authors approved the final manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information is available for this paper at https://doi.org/10.1038/s41552-020-0115-6.
Correspondence and requests for materials should be addressed to G.A.
Reprints and permission information is available at http://www.nature.com/reprints
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020