BRIEF COMMUNICATION   OPEN

Check for updates

FHIR Genomics: enabling standardization for precision medicine use cases

Gil Alterovitz [1,2 ✉], Bret Heale [3], James Jones[1], David Kreda[4], Fan Lin[1,5], Lei Liu[1], Xin Liu[1], Kenneth D. Mandl [1,6], David W. Poloway[1], Rachel Ramoni[7], Alex Wagner [8] and Jeremy L. Warner [9]

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.

npj Genomic Medicine (2020)5:13 ; https://doi.org/10.1038/s41525-020-0115-6

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

FIHR 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 FamilyMember-History 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,

---

[1]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, USA. [2]Harvard/MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, MA, USA. [3]Intermountain Healthcare, Salt Lake City, UT, USA. [4]Center for Biomedical Informatics, Harvard Medical School, Boston, MA, USA. [5]Xiamen University, School of Software, Xiamen, Fujian, China. [6]Departments of Pediatrics and Biomedical Informatics, Harvard Medical School, Boston, MA, USA. [7]Partners HealthCare, Boston, MA, USA. [8]Division of Oncology, Washington University School of Medicine, St. Louis, MO, USA. [9]Departments of Biomedical Informatics and Medicine, Vanderbilt University, Nashville, TN, USA ✉email: ga@alum.mit.edu
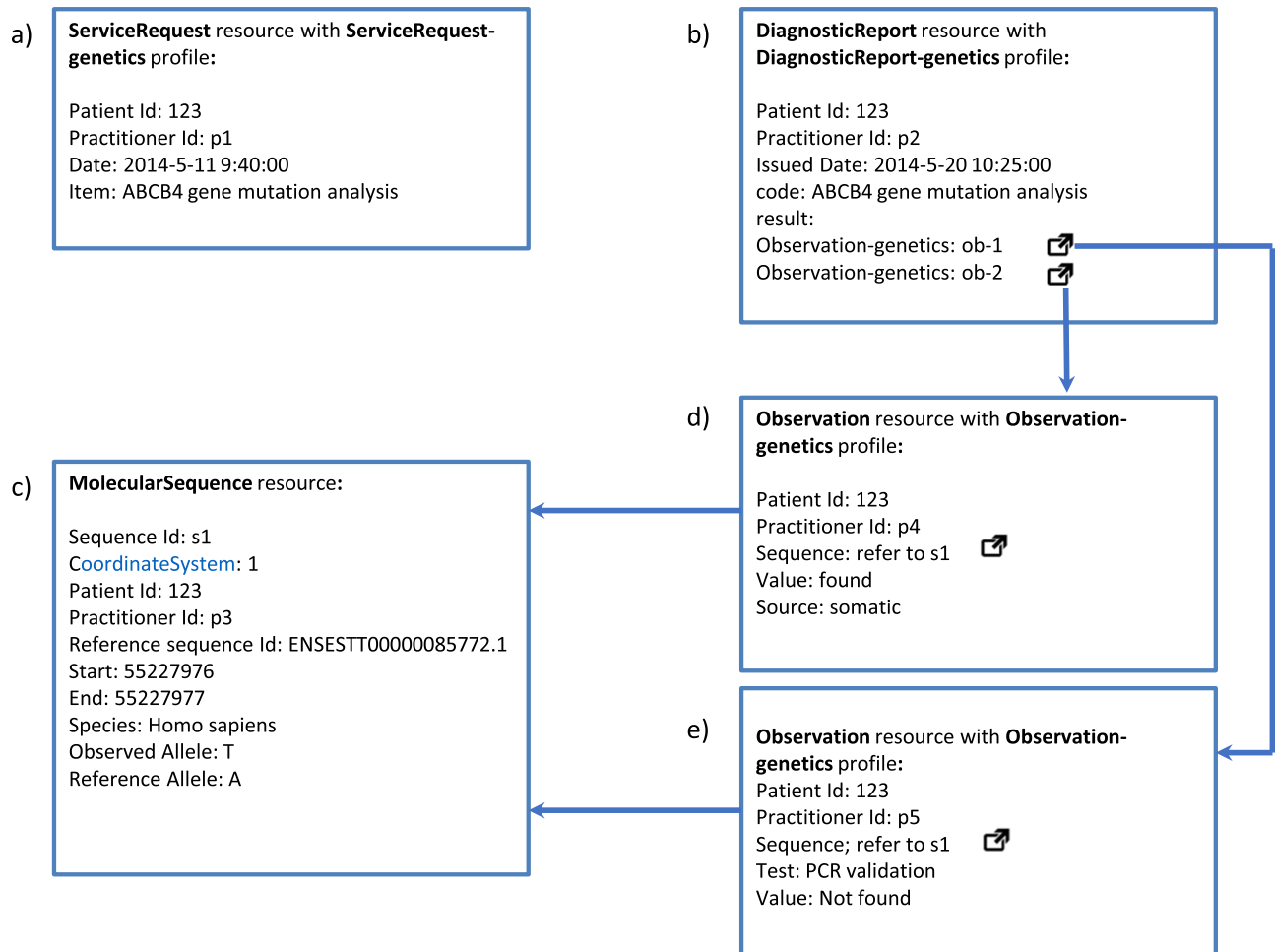
Fig. 1 FHIR Genomics workflow use case. a A *ServiceRequest* instance; **b** a *DiagnosticReport* instance refers to *Observations* under "result"; **c** a *MolecularSequence* instance carries the reference sequence and variant information; **d** an *Observation* instance carries clinical interpretation; **e** additional *Observations* can carry further analysis information. Arrows depict the inter-resource pointers.

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. Figure 1 shows an example depicting a common workflow of FHIR Genomics from an EHR perspective—an order is requested, and test results are reported.

Mirroring the common approach of a Picture Archiving and Communication System, it is proposed that EHRs will store only metadata—such as path, address, and IDs—and then retrieve raw sequencing data as needed, forming a Genomic Archiving Communication System (GACS). This approach takes into consideration the large size and limited clinical application of raw genomic data, as well as the trend of new infrastructure being released to the research community before clinical optimization.[16] Following this model, *MolecularSequence* enables GACS integration with its repository feature, which can carry URIs and identifiers needed to retrieve

sequencing data stored in specifications maintained by the Global Alliance for Genomics and Health (GA4GH), for example. A similar approach may also be taken with genomic knowledge-based artifacts as scientific evidence constantly increases. Following ref. [15], an FHIR Genomics app may take a patient's clinical context and observed variations and link this information to external databases to ease interpretation (see Fig. 2).

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 SAMtools[19] 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
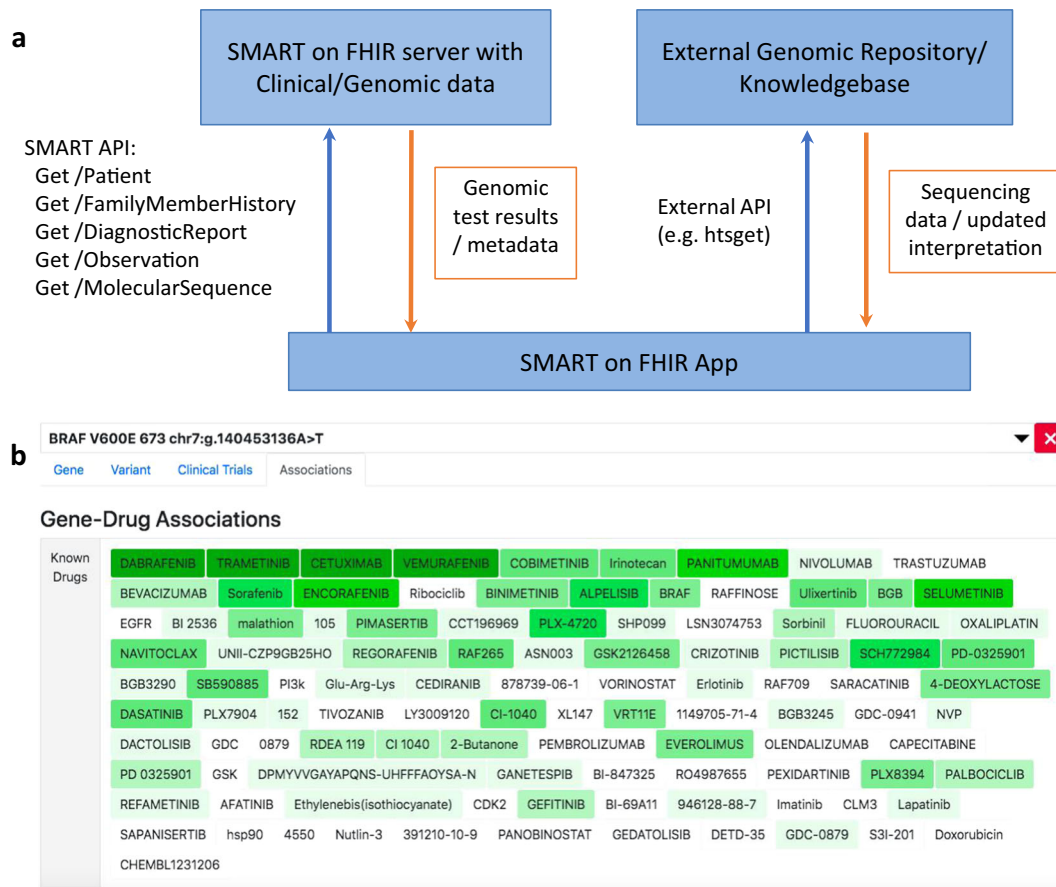
Fig. 2 **Example of SMART on FHIR application. a** Arrows show API calls integrating clinical and genomic data at the point of care with external information, for example, through the GA4GH streaming standard htsget.[33] **b** Screenshot of sample application showing associated drugs in the VICC meta-KB[22] for a given variant on a sample patient. Clicking a drug name provides a list of relevant publications, sorted by AMP/ASCO/CAP guideline levels of evidence.[23]

through GA4GH APIs such as Beacon,[20] 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.[21] In addition, cancer variants were linked to the Variant Interpretation for Cancer Consortium meta-knowledgebase API[22] for context, with evidence classified by the AMP/ASCO/CAP somatic classification guidelines.[23] 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,[24] HGVS nomenclature,[25] HUGO Gene Nomenclature Committee identifiers,[26] NCBI reference sequences,[27] terms from the Sequence Ontology,[28] clinical guidelines,[29,30] and services, such as LOINC[31] and SNOMED CT.[32]

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

REFERENCES

1. Metzker, M. L. Sequencing technologies — the next generation. *Nat. Rev. Genet.* **11**, 31 (2009).
2. Katsanis, S. H. & Katsanis, N. Molecular genetic testing and the future of clinical genomics. *Nat. Rev. Genet.* **14**, 415 (2013).

3. Wu, T. H., Hsiue, E. H., Lee, J. H., Lin, C. C. & Yang, J. C. New data on clinical decisions in NSCLC patients with uncommon EGFR mutations. *Expert Rev. Respir. Med.* **11**, 51–55 (2017).

4. Lu, S., Nand, R. A., Yang, J. S., Chen, G. & Gross, A. S. Pharmacokinetics of CYP2C9, CYP2C19, and CYP2D6 substrates in healthy Chinese and European subjects. *Eur. J. Clin. Pharmacol.* **74**, 285–296 (2018).

5. Tenenbaum, J. D., Sansone, S.-A. & Haendel, M. A sea of standards for omics data: sink or swim? *J. Am. Med. Inform. Assoc.* **21**, 200–203 (2014).

6. Overby, C. L. & Tarczy-Hornoch, P. Personalized medicine: challenges and opportunities for translational bioinformatics. *Personalized Med.* **10**, 453–462 (2013).

7. Voelkerding, K. V., Dames, S. A. & Durtschi, J. D. Next-generation sequencing: from basic research to diagnostics. *Clin. Chem.* **55**, 641 (2009).

8. Tarczy-Hornoch, P. et al. A survey of informatics approaches to whole-exome and whole-genome clinical reporting in the electronic health record. *Genet. Med.* **15**, 824–832 (2013).

9. Mandel, J. C., Kreda, D. A., Mandl, K. D., Kohane, I. S & Ramoni, R. B. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J. Am. Med. Inform. Assoc.* **23**, 899–908 (2016).

10. Farr, C. Apple will let you keep your medical records on your iPhone. CNBC. https://www.cnbc.com/2018/01/24/apple-coo-williams-says-new-health-record-beta-is-right-thing-to-do.html (2018) (accessed 3 Dec 2018).

11. Conn, J. Epic tries to ease outside app development. Modern Healthcare. https://www.modernhealthcare.com/article/20170222/NEWS/170229974/epic-new-app-program-tries-to-connect-ehr-networks (2017) (accessed 3 Dec 2018).

12. Mandl, K. D., Mandel, J. C. & Kohane, I. S. Driving innovation in health systems through an apps-based information economy. *Cell Syst.* https://doi.org/10.1016/j.cell.2015.05.001 (2015).

13. Alterovitz, G. et al. SMART on FHIR Genomics: facilitating standardized clinico-genomic apps. *J. Am. Med. Inform. Assoc.* **22**, 1173–1178 (2015).

14. HL7.org. HL7 Standards Product Brief - HL7 domain analysis model: clinical sequencing, release 1. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=446 (2017) (accessed 30 Jan 2019).

15. Warner, J. L. et al. SMART cancer navigator: a framework for implementing ASCO Workshop Recommendations to enable precision cancer medicine. *JCO Precis. Oncol.* https://doi.org/10.1200/PO.17.00292 (2018).

16. Aronson, S. J. & Rehm, H. L. Building the foundation for genomics in precision medicine. *Nature* **526**, 336–342 (2015).

17. HL7.org: HL7 version 2 implementation guide: clinical genomics. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=23 (2013) (accessed 30 Jan 2019).

18. HL7.org: HL7 Standards Product Brief - HL7 version 3: Reference Information Model (RIM). https://www.hl7.org/implement/standards/product_brief.cfm?product_id=77 (2017) (accessed 30 Jan 2019).

19. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

20. Fiume, M. et al. Federated discovery and sharing of genomic data using Beacons. *Nat. Biotechnol.* **37**, 220–224 (2019).

21. HL7.org: Genomics Reporting Implementation Guide. http://hl7.org/fhir/uv/genomics-reporting/index.html (2019) (accessed 18 Dec 2019).

22. Wagner, A. H. et al. A harmonized meta-knowledgebase of clinical interpretations of cancer genomic variants. Preprint at https://doi.org/10.1101/366856v2 (2018).

23. Li, M. M. et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J. Mol. Diagn.* **19**, 4–23 (2017).

24. Landrum, M. J. et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–D868 (2016).

25. den, Dunnen, J. T. et al. HGVS recommendations for the description of sequence variants: 2016 update. *Hum. Mutat.* **37**, 564–569 (2016).

26. Gray, K. A., Yates, B., Seal, R. L., Wright, M. W. & Bruford, E. A. Genenames.org: the HGNC resources in 2015. *Nucleic Acids Res.* **43**, D1079–D1085 (2015).

27. Pruitt, K. et al. in *The NCBI Handbook* (eds McEntyre, J. & Ostell, J.) Ch. 18 (National Center for Biotechnology Information, Bethesda, MD, 2002).

28. Eilbeck, K. et al. The Sequence Ontology: a tool for the unification of genome annotations. *Genome Biol.* **6**, R44 (2005).

29. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–423 (2015).

30. Caudle, K. et al. Standardizing terms for clinical pharmacogenetic test results: consensus terms from the Clinical Pharmacogenetics Implementation Consortium (CPIC). *Genet. Med.* **19**, 215–223 (2017).

31. McDonald, C. J. et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin. Chem.* **49**, 624–633 (2003).

32. Richesson, R. L., Andrews, J. E. & Krischer, J. P. Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. *J. Am. Med. Inform. Assoc.* **13**, 536–546 (2006).

33. Kelleher, J. et al. GA4GH Streaming Task Team, htsget: a protocol for securely streaming genomic data. *Bioinformatics* **35**, 119–121 (2019).

ACKNOWLEDGEMENTS

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/s41525-020-0115-6.

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.